

MAPPING MEAN ANNUAL AND OTHER
RAINFALL STATISTICS OVER SOUTHERN AFRICA

by

M.C. DENT, S.D. LYNCH AND R.E. SCHULZE

with contributions by A.W. Seed and H.M.M. Wills

Department of Agricultural Engineering

University of Natal

Pietermaritzburg 3201, South Africa

Agricultural Catchments Research Unit

Report No. 27

1987

Report to the Water Research Commission on the Project
"Revision of the Temporal and Spatial Distribution of
Precipitation Statistics in Southern Africa"

Head of Department	:	Professor P. Meiring
Projects Leader	:	Professor R.E. Schulze

WRC REPORT NO. 109/1/89
ISBN : 0-947447-16-4

00070044

ACKNOWLEDGEMENTS

The research described in this report resulted from a project funded by the Water Research Commission and entitled:

"Revision of the temporal and spatial distribution of precipitation statistics in Southern Africa."

The Steering Committee responsible for this project, consisted of the following persons,

Mr D.W.H. Cousens	Water Research Commission (Chairman)
Mr P.W. Weideman	Water Research Commission (Secretary)
Dr P.J.T. Roberts	Water Research Commission
Prof. P. Meiring	University of Natal
Prof. R.E. Schulze	University of Natal
Mr F.J. Kruger	Department of Environment Affairs
Mr J.F. Erasmus	Department of Agricultural and Water Supply
Dr P.T. Adamson	Department of Water Affairs
Mr F. Viljoen	Department of Environment Affairs

The financing of the project by the Water Research Commission and the contribution of the members of the Steering Committee is acknowledged gratefully.

This project of five years' duration was only possible with the cooperation of many individuals and institutions. The authors therefore wish to record their sincere thanks to the following:

Department of Environment Affairs, South African Weather Bureau

in particular Mr J.S. van Rhyn, for rainfall data on magnetic tape and for the assistance with the location of rainfall stations;

Department of Agriculture and Water Supply, Soil and Irrigation Research Institute

in particular Messrs J. Erasmus, J. Myburgh, J. Hoorn, D. Turner, S. Barrow and D. Rowsell, for rainfall data on magnetic tape and the Western Cape altitude grid;

University of Natal, Pietermaritzburg, Computer Services Division

for assistance and advice with computing facilities;

University of the Witwatersrand erstwhile Hydrological Research Unit

Prof. D.C. Midgley and Dr W.V. Pitman, for providing the South African grid of co-ordinates and altitudes;

Department of Environment Affairs, S.A. Forestry Research Institute

Mr J. Bosch and Mr F. Rogers for rainfall data in the Jonkershoek and Cathedral Peak areas;

South African Sugar Association Experiment Station

Dr G. Thompson, Mr M. Murdoch and Mr R. Harding for providing rainfall data;

Natal Parks Board, Organised Agriculture, Public of South Africa

for supplying rainfall data;

Department of Geography, University of Natal, Pietermaritzburg

for the loan of maps;

Mr A.W. Seed

whose input to this project as an M.Sc. Engineering student was most innovative and dynamic;

Mrs H.M.M. Wills

of the Department of Agricultural Engineering who was responsible for large parts of the computing on this project;

The Computing Centre for Water Research

a venture by ISM (Pty) Limited, the Water Research Commission and the University of Natal for provision of a computing facility dedicated to water research at the University of Natal, Pietermaritzburg. This facility was of great benefit to this project;

Dr R. Sparks, Prof. W. Zucchini, Prof. G. Pegram, Mr A. Lonergan

for advice and comment on some of the statistical aspects of monthly rainfall synthesis;

Mr S. Piper

Department of Land Surveying and Mapping, University of Natal, Durban, for providing the Natal altitude grid;

Messrs and Mesdames G. Angus, S. Dunsmore, J. Garland, B. Gaydon, W. George, M. Maharaj, E. Murgatroyd, S. Neuwirth, C. O'Mahoney, H. Tarboton, K. Temple, J. Whyte, K. Wiercx, N. Wijnberg

all of the Department of Agricultural Engineering and who worked on the project at various times with patience, diligence and tenacity;

Messrs and Mesdames J. Conning, K. Fismer, M. Gorven and A. Kure

staff of the Computing Centre for Water Research.

TABLE OF CONTENTS

	Page
LIST OF TABLES	I
LIST OF FIGURES	II
Chapter	
1 INTRODUCTION	1 - 1
1.1 Objectives	1 - 2
1.2 Importance of mean annual precipitation in hydrological applications	1 - 2
1.3 Mapping rainfall statistics	1 - 4
2 THE PHYSIOGRAPHIC DATA SETS AND INDICES	2 - 1
2.1 The importance of physiographic information in mapping precipitation	2 - 2
2.1.1 In other parts of the world	2 - 2
2.1.2 In South Africa	2 - 4
2.2 Need for establishment of a physiographic data base	2 - 5
2.3 Establishing the location of rainfall stations	2 - 6
2.4 The grid system of altitudes	2 - 9
2.5 Checking the altitude of grid points and rainfall stations	2 - 14
2.6 Exposure, roughness, aspect, rainbearing wind direction and continentality indices	2 - 16
2.6.1 Exposure index	2 - 16
2.6.2 Roughness index	2 - 18
2.6.3 Continentality index	2 - 18

2.6.4	Distance from a physiographic barrier	2 - 19
2.6.5	Aspect index	2 - 20
2.6.6	Rainbearing wind direction index	2 - 23
3	THE RAINFALL DATA FILE	3 - 1
3.1	Sources of rainfall data	3 - 1
3.2	Quality of the rainfall data	3 - 3
3.3	Checking the rainfall data	3 - 8
3.4	Managing the rainfall data	3 - 13
4	SYNTHESISING MONTHLY RAINFALL	4 - 1
4.1	Period and desired length of the rainfall record	4 - 4
4.2	Techniques for synthesising monthly rainfall	4 - 13
4.2.1	Background and concepts of synthesising	4 - 13
4.2.2	Selection of target:control station pairs	4 - 15
4.2.3	Synthesising models reported by Zucchini and Hiemstra (1984)	4 - 18
4.3	Optimal interpolation techniques	4 - 23
4.3.1	Universal Kriging	4 - 24
4.3.2	Ripley's method	4 - 25
4.3.3	Gandin's method	4 - 25
4.4	Multiple regression	4 - 26
4.5	Comparison between Gandin's method and multiple regression	4 - 28
4.6	Comparison between Gandin's method and the Zucchini and Hiemstra models	4 - 32
4.7	Replacement of some missing months by the mean value for that month	4 - 35

5	AUTOMATION OF THE REGRESSION PROCEDURE FOR RELATING MEAN ANNUAL PRECIPITATION TO PHYSIOGRAPHY	5 - 1
5.1	Procedural steps in the regression process	5 - 2
5.2	Examination and treatment of residuals	5 - 8
5.3	The response surface outside the predictor space	5 - 9
5.4	Weighting and transforming the regression variables	5 - 10
5.5	Discussion	5 - 11
6	MAPPING STATISTICS OF PRECIPITATION	6 - 1
6.1	Comparison of techniques of mapping	6 - 2
6.2	Confidence of estimates and estimates of error in mapped values of MAP	6 - 13
6.3	The composite regression/residual surface	6 - 19
6.4	Selection of stations for use in mapping	6 - 22
6.5	Example of the procedure for generating a MAP surface	6 - 28
6.6	Mapping monthly mean and variability of rainfall	6 - 35
7	HOMOGENEOUS RAINFALL REGIONS	7 - 1
7.1	Existing delimitations of rainfall regions	7 - 1
7.2	Method of delimitation of homogeneous rainfall regions	7 - 3
7.3	Problems at regional boundaries	7 - 6
7.4	Conclusions	7 - 13

8	DISCUSSION AND CONCLUSIONS	8 - 1
8.1	Merits of isohyetal maps vs digital images of rainfall	8 - 1
8.2	Other uses of the products of this study	8 - 2
8.3	Recommendations for future research	8 - 5
9	SUMMARY	9 - 1
9.1	Introduction and objectives	9 - 1
9.2	Need for a physiographic data set	9 - 3
9.3	Basic rainfall and physiographic data	9 - 4
9.4	Generated rainfall and physiographic data	9 - 5
9.5	Checking the rainfall and physiographic data	9 - 6
9.6	Automated management of the rainfall data base	9 - 7
9.7	Regression of MAP on physiographic variables	9 - 7
9.8	Techniques of mapping MAP	9 - 8
9.9	Mapping mean monthly and other rainfall statistics	9 - 8
9.10	Delimitation of homogeneous rainfall regions	9 - 9
9.11	Conclusions and recommendations for future research	9 - 10
9.11.1	Merits of isohyetal maps vs digital images of rainfall	9 - 10
9.11.2	Other uses of the products of this study	9 - 11
9.11.3	Recommendations for future research	9 - 13

10	REFERENCES	10 - 1
11	APPENDIXES	11 - 1
A	Upper limits of the percentage error of short term mean and standard deviation	11 - 1
B	Monthly rainfall synthesising techniques	11 - 25
B.1	Universal Kriging	
B.2	Ripley's method	
B.3	Gandin's method	
C	Recipients of data products	11 - 33
D	List of computer programs used in this study	11 - 35
E	Map of homogeneous rainfall region boundaries	

LIST OF TABLES

Table	Page
2.1 Rainbearing wind direction indices	2 - 25
3.1 Breakdown of the sources and location of rainfall data	3 - 2
3.2 Quality codes used in the monthly rainfall data file and the SAWB daily equivalents	3 - 6
3.3 Daily and monthly rainfall data for Jonkersnek and Virgin Peaks	3 - 11 3 - 12
3.4 Format of the monthly rainfall data for one station	3 - 14
3.5 Format of the file INDEX	3 - 15
3.6 Search and extraction request file	3 - 19
4.1 Upper limits of the percentage error of the short term mean: An example of the analysis of possible error for the Welding and Havenga region 20B	4 - 11
4.2 Percentage of missing months and years in the rainfall data set	4 - 14
4.3 Model test results, percentage of points in class 0,7 - 1,3 of ratio (synthetic:observed)	4 - 36
4.4 Effect on the MAP and the standard deviation of annual totals if a limited number of missing months are replaced by their observed means	4 - 38
6.1 Correlation matrices and adjusted r^2 from regressions in the Drakensberg and OFS/TVL study areas	6 - 10

LIST OF FIGURES

Figure	Page
2.1 Altitude grid data over Southern Africa (sources of information)	2 - 10
2.2 Points at which the altitude from the HRU data set differed by more than 100 m from altitudes extracted off the 1:50 000 maps	2 - 12
2.3 Illustration of the grid point altitude/ rainfall station altitude discrepancy	2 - 14
2.4 Classified image of altitudes in the south western Cape	2 - 17
2.5 Rainfall/altitude relationship along section from Bergville in the NNE to Mothelsessane in the SSW	2 - 20
2.6 Classified digital image of the aspect index in the south western Cape	2 - 22
2.7 Schematic diagram of the numerical classification system for aspect	2 - 23
2.8 Illustration of the development of rainbearing wind direction index	2 - 24
3.1 Source area for new rainfall stations from SIRI western Cape region	3 - 4
3.2 Location of the "new" and "old" long term (>15 years stations in an area of Natal/Orange Free State	3 - 5
3.3 Logic path of the data search and extraction program	3 - 17
3.4 Schematic illustration of the spatial depiction routine	3 - 21
4.1 Spatial distribution of rainfall stations in the south western Cape with record lengths of more than 10, 20 and 40 years	4 - 3

4.2	Cumulative frequency distributions of rainfall station altitude and grid point altitude for rainfall stations in the south western Cape with record lengths of more than 10, 20 and 40 years	4 - 5
4.3	Record lengths required to ensure that the mean annual rainfall estimates to within 10 per cent of the long term mean, 90 per cent of the time	4 - 10
4.4	Schematic presentation of the procedure to assess the Zucchini and Hiemstra (1984) models	4 - 19
4.5	The study area in relation to the Welding and Havenga (1974) rainfall regions for Natal	4 - 29
4.6	Ratio of synthetic:observed January rainfall for best subset multiple regression and Gandin's method	4 - 30
4.7	Ratio of synthetic:observed mean January rainfall for best subset multiple regression and Gandin's method	4 - 31
4.8	Ratio of synthetic:observed standard deviation of January rainfall for best subset multiple regression and Gandin's method	4 - 31
4.9	Location of the areas selected for testing Gandin's method against the models presented by Zucchini and Hiemstra (1984)	4 - 33
5.1	Regional boundaries used in the regression analysis	5 - 3
6.1	Location of the OFS/TVL and Drakensberg study areas	6 - 3
6.2	Three dimensional representation of the physiography of the Drakensberg study area	6 - 5
6.3	Three dimensional representation of the physiography of the OFS/TVL study area	6 - 5

6.4	Frequency distribution of grid point and station point altitudes in the Drakensberg study area	6 - 6
6.5	Frequency distribution of grid point and station point altitudes in the OFS/TVL study area	6 - 6
6.6	Percentage difference between the estimates of MAP using interpolation (I) and regression (R) in the OFS/TVL study area	6 - 8
6.7	Percentage difference between the estimates of MAP using regression (R) and interpolation (I) in the Drakensberg study area	6 - 9
6.8	Profile analysis to illustrate the performance of the regression and interpolation techniques in the Drakensberg	6 - 11
6.9	Observed versus estimated MAP using regression in the Drakensberg areas	6 - 11
6.10	Pictorial representation of a linear regression model	6 - 14
6.11	MAP in the Babanango area of Natal predicted by the least squares regression technique (Units: mm x 100)	6 - 15
6.12	The 90 per cent confidence limits of predicted MAP in the Babanango area of Natal, using least squares regression (Units: mm)	6 - 16
6.13	Profile of MAP surface along the Section AA' in Figure 6.14 (Study area: Venda)	6 - 20
6.14	Isohyets of MAP, obtained by multiple regression only (Study area: Venda)	6 - 20
6.15	Residual surface of MAP (Study area: Venda)	6 - 21
6.16	Regression surface adjusted according to the residual surface (Study area: Venda)	6 - 21

6.17	Cumulative frequency distributions of rainfall station altitude and grid point altitude for record lengths of more than 10, 20 and 40 years in the Natal Drakensberg	6 - 24
6.18	Existing isohyets of MAP in the Frankfort area	6 - 26
6.19	Isohyets of MAP in the Frankfort area, drawn using ploynomial interpolation	6 - 27
6.20	Location of the south western Cape region used in this study	6 - 28
6.21	Three dimensional plot of altitudes in a part of the south western Cape region	6 - 29
6.22	Location of the 28 stations (>20 years) used in this area for the analysis of MAP in the south western Cape region	6 - 29
6.23	Isohyets of MAP by direct interpolation between the station MAPs in the south western Cape region (Units: mm x 100)	6 - 30
6.24	Isohyets of MAP resulting from the regression technique in the south western Cape region (Units: mm x 100)	6 - 31
6.25	Residual surface by interpolation from the station point residuals in the south western Cape region (Units: percentage)	6 - 32
6.26	The 90 per cent confidence limits of prediction of MAP using least squares regression, expressed as a percentage of the predicted value in the south western Cape region (Units: mm x 100)	6 - 33
6.27	Regression and residual surfaces of MAP combined in the south western Cape region (Units: mm x 100)	6 - 34
7.1	Classified image of altitude and the location of stations with >10 years of record for part of the south western Cape	7 - 4

7.2	The detailed delimitation of rainfall regions in the south western Cape	7 - 7
7.3	Illustrations of the problem of merging rainfall surfaces at regional boundaries	7 - 9
7.4	Merging of two surfaces of MAP in which the altitude term is more strongly represented in the region to the east of AA'	7 - 11
7.5	Continuity at the boundary after adjusting the altitude term in the region to the east of AA'	7 - 12

CHAPTER 1

INTRODUCTION

Mean annual precipitation¹ (MAP) is perhaps the most widely used variable in hydrological design, water resources planning and agrohydrology. In the past two decades one of the basic sources of information on MAP in Southern Africa² has been the 1:250 000 average rainfall map series compiled and drawn by the Hydrological Research Division, Department of Water Affairs in 1965 from data obtained from the South African Weather Bureau. Since then the rainfall records have lengthened by more than 20 years. Furthermore, techniques of analysis and computerised mapping have become more sophisticated than in the 1960's, as a direct consequence of considerably enhanced computer capabilities.

This chapter provides an overview of the objectives of the project in addition to a review of the importance of mapping rainfall statistics. Previous maps of rainfall statistics, in particular MAP, in Southern Africa are mentioned. An outline of the data sets and techniques used and reported in this project is presented briefly as are some of the important by-products of this study.

¹ The terms precipitation and rainfall will be used synonymously throughout this document to designate rainfall, since fog is not measured at sufficient sites to warrant inclusion in this study and the contribution of snow is not important to total precipitation in Southern Africa.

² Southern Africa is taken to include the Republic of South Africa, the TBVC States, as well as Lesotho and parts of Swaziland for the purposes of this document.

1.1 Objectives

The engineering/hydrology fraternity requested the Water Research Commission to initiate research into the revision of the 1:250 000 average annual rainfall map series. This project to revise the temporal and spatial distribution of statistics in Southern Africa, was therefore motivated to the Water Research Commission and commenced in July 1982. The project had the following objectives:

- (a) the re-evaluation of the distribution of mean annual and mean monthly precipitation over Southern Africa,
- (b) an investigation into the variability and other statistics of precipitation over Southern Africa,
- (c) the automated mapping of these precipitation statistics, and
- (d) a re-evaluation of the regions of homogeneous precipitation distribution.

1.2 Importance of mean annual precipitation in hydrological applications

The importance of mean annual precipitation as an index in hydrological applications may be gauged by the numerous examples mentioned below. MAP is an index which is used to characterise overall climate and moisture status of a catchment. It has a major influence on both soil conditions and their drainage characteristics (Bedient, Huber and Heaney, 1978) and is a dominant factor influencing type and condition of vegetation. Pitman (1980) used MAP as a basis for the regionalisation of SWA/Namibia into homogeneous meteorological zones. He also related the frequency of occurrence of severe storms and amount of rainfall per event to MAP. Midgley and Pitman (1969) used MAP in the estimation of Mean Annual Runoff (MAR). Midgley and Pitman (1978) incorporated MAP in their depth-duration-

frequency relationship for point rainfall in South Africa. In this capacity alone MAP must surely be used many hundreds of times per year in Southern Africa. Pitman (1973) expressed the rainfall input to his well known monthly streamflow estimation model in terms of MAP. Pitman and Stern (1981) used MAP in the same role in SWA/Namibia. Seeber (1983) used MAP in an extreme-value expression for the determination of the cumulative distribution of point rainfall rate in Southern Africa in a model for use in the telecommunications industry. Schulze (1984) used the streamflows from Pitman, Middleton and Midgley (1981) and related MAR to MAP through regression analysis in 21 runoff regions in Natal to then map MAR using MAP as the only predictor. Midgley and Pitman (1969) showed that an underestimation of 10 per cent in a catchment MAP of 1 300 mm would underestimate MAR by 26 per cent in the Drakensberg. Boughton (1981) found that MAR varied from 100 mm to 530 mm for an MAP variation from 850 mm to 1700 mm in the Upper Condamine River Basin, Australia and that variations of 20 per cent in MAR for 10 per cent change in MAP are common.

In the field of agrohydrology, MAP has been used extensively. Jones (1982) expressed Eragrostis curvula yield as a function, inter alia of MAP. Brockett (1982) used MAP and other variables in order to obtain a function for kikuyu yield. MAP and other variables were used by Jones et al (1980) to express a minimum percentage fodder-bank accumulation. Schulze (1983) used recommendations from the Department of Agriculture and Fisheries (1981) and then developed regional regression equations which included MAP in order to enable the mapping of average first burning dates of veld in Natal. Nield and Boshell (1976) incorporated MAP as one of their criteria for determining optimum growing areas for pineapples. Schulze (1983) obtained correlation coefficients typically of 0,84 when he related first planting dates for maize to MAP in Natal.

MAP, altitude, temperature and duration of soil moisture deficit were used by Schönau (1982) in order to determine criteria for optimum growth of various timber species.

1.3 Mapping rainfall statistics

The areal estimation of long term rainfall, from point measurements, is an important pre-requisite to the use of this fundamental hydrological variable in, for example, distributed modelling and agricultural land use planning. It is therefore necessary to estimate and then extrapolate and interpolate this variable. According to Hall and Barclay (1975) the analyst must resort to sophisticated statistical procedures to overcome data deficiencies unless he is satisfied with a mere index of rainfall in the assessment of water resources. The technique of trend surface analysis was selected and automated to perform this function in combination with the residual rainfall surface which resulted from the regression analysis.

In this project the task of mapping the temporal and spatial statistics of rainfall in Southern Africa, required an extensive data set of rainfall and physiographic variables. The creation and/or collation, checking and management of these data was a fundamental prerequisite to the successful completion of this project. Daily or monthly rainfall records from 9409 stations form the precipitation data set. Since physiography plays an important role in influencing the spatial distribution of long term average rainfall in many areas, a comprehensive set of primary and generated physiographic data was also created. The generated physiographic data were indices of continentality, exposure, distance to a mountain barrier, surface roughness and aspect. Such data are pre-requisites to performing trend surface analyses of MAP in areas which have a complex terrain and a sparse raingauge network. These data were assembled onto a grid at a resolution of one minute of a

degree. The size of this data set dictated that most of the functions of checking, manipulation and management be computerised into highly automated systems. However, patient and careful human scrutiny was also required.

An investigation into previous research involving mapping MAP over South Africa revealed several studies in addition to the currently used 1:250 000 average annual rainfall map series. However most of these were confined to selected areas. For example, Whitmore (1968) investigated the relationship between MAP and locality and site factors such as altitude, aspect, continentality and longitude in the southern Cape. Schulze (1975) mapped MAP at selected catchments at Cathedral Peak and he later mapped MAP for the entire Natal Drakensberg area, using trend surface analysis (Schulze, 1979). Hughes (1982) presented a relationship between MAP and physiographic variables for a section of the south west Cape coast. A map of MAP for Natal was prepared by Schulze (1983), using trend surface analysis over parts of Natal. Myburgh and Hoon (1985) produced maps of MAP and mean monthly rainfall at a scale of 1:250 000 for the winter rainfall region shown in Figure 3.1. Many of the techniques, particularly trend surface analysis, and ideas proposed by the abovementioned were used in this study.

The temporal distribution of rainfall at monthly level and variability of the annual rainfall are also important hydrological indices. Therefore in this project the mean monthly precipitation digital images have been generated on the same spatial resolution as the MAP, namely 1 minute of a degree. The variability of MAP expressed in terms of the coefficient of variation, the skewness coefficient and the 75th percentile of the annual rainfall are also produced at the same resolution in digital form.

The extensive data file of MAP and physiographic variables provided a good basis for the delimitation of 712 homogeneous precipitation regions. This time consuming and detailed exercise was performed manually using human intelligence and pattern recognition techniques.

The data set has also been most useful in a number of other applications. One of the objectives of this project was to automate the mapping of these precipitation statistics, this in view of the magnitude of the task and the possible need to revise such maps in the future. A number of valuable by-products and data management techniques were thus developed, the use and significance of which go well beyond the objectives of this project. Examples of these by-products are a monthly rainfall data file which was developed with data from a number of sources and is now frequently being sought by the same institutions which supplied the data, since this data set has all the other institutions' data included in one data file. The grid of altitudes and other physiographic data which were gathered or generated for this study, is also a much sought after by-product, which has been used to produce preliminary maps, inter alia, of monthly temperatures and evaporation in Southern Africa. The detailed regionalisation of rainfall has been used in two other major current studies viz. on regional soil moisture deficits (Dent, Schulze and Angus, 1987) and storm runoff estimates (Schmidt and Schulze, 1987a; 1987b).

Following on this introductory chapter the next three chapters elaborate on the data sets used, viz those of physiography rainfall. Thereafter the techniques of multiple regression for trend surface analysis and residual analysis followed by mapping procedures are explained. Finally the delimitation of homogeneous rainfall regions is discussed and possibilities for future research based on the data and findings of this project are reviewed.

CHAPTER 2

THE PHYSIOGRAPHIC DATA SETS AND INDICES

The influence of physiography on the spatial distribution of MAP has been known for a long time. According to Bleasdale and Chan (1973) this appreciation had, however, been based largely on subjective, qualitative assessment in many instances. The existing 1:250 000 average annual rainfall map series for Southern Africa reflects the influence of physiography on MAP. It was recognised at the commencement of this project that techniques involving the prediction of rainfall surfaces from physiographic variables would most probably play a major role in this research.

Mandeville and Rodda (1970) state that those techniques of areal integration of rainfall which rely on the arithmetic mean and the Thiessen polygon methods or on the subjective drawing of isohyets tend to be used without sufficient appreciation of their limitations. The Thiessen polygon methods employ a fundamental assumption of the juxtaposition of similar amounts of rainfall. Their use is therefore confined generally to flatter areas where rainfall gradients are not marked (Mandeville and Rodda, 1970). According to Hall and Barclay (1975) the technique of fitting multiquadric polynomial surfaces, as employed by Shaw and Lynn (1972), does not allow for orographic effects and care is required in selecting the stations, as some combinations produce instability of the solution. An inspection of the existing MAP maps of Southern Africa revealed that the flatter areas, for example, the Karoo and parts of the Orange Free State and Transvaal would probably be more suited to have rainfall mapped by multiquadric surface or direct interpolation techniques. The mapping of precipita-

tion statistics in the western as well as southern Cape and eastern seaboard, the Drakensberg and the eastern Transvaal escarpment would probably be undertaken more suitably using trend surface techniques which relate rainfall, inter alia to physiography. In the absence of a formal report on the techniques employed to produce the existing 1:250 000 maps of MAP in South Africa and on inspection of these isohyets, it must be presumed that a major decision was taken with regard to the influence of altitude on MAP in the production of these maps.

A review of the literature on the influence of physiography on precipitation is therefore considered appropriate at this stage, in order to explain why so much importance was attached to the creation of a physiographic data set. These physiographic data are essential for any form of trend surface analysis, which was one of the major techniques employed in this project.

2.1 The importance of physiographic information in mapping precipitation

Research into the quantification of the physiographic factors which affect rainfall has been well documented for Southern Africa as well as many other parts of the world.

2.1.1 In other parts of the world

Spren (1947) reported that altitude accounted for 30 per cent of the variation in mean seasonal precipitation in western Colorado. Schermerhorn (1967) found, for western Oregon and the state of Washington, that station altitude on its own did not account for much of the variation in annual precipitation. He found it necessary to include indices, inter alia, of barrier elevation to explain most of the variation of MAP. Hutchinson (1968; 1969) used a regression equation which involved altitude, distance from

the coast, exposure and aspect to map MAP in Otago and Canterbury in New Zealand. A word of caution in using topographic variables in the estimation of annual precipitation, when such precipitation may be separated into that resulting from winter frontal systems and summer-time convective storms, was sounded by Duckstein et al. (1972). They maintain that such analyses using annual precipitation data may also introduce considerable bias for determining topographic effects on precipitation. Duckstein et al. (1972) did find, however, that for summer convective storms in a region within Arizona the mean number of events per season increased considerably with altitude, that the mean amount of rainfall increased moderately with altitude and that the mean total seasonal rainfall increased as a quadratic polynomial of altitude. The superposition of local topographical influences on regional rainfall patterns was noted by Lee (1973). Edwards (1973) stated that trend surface analysis based on physiography was one of the most promising fully automated techniques for providing rainfall surfaces to unevenly spaced data points. Research by Torrance (1973) in the eastern border mountains of Zimbabwe illustrated the increase of precipitation with altitude. Huff et al (1975) reported an average 15 per cent more summer precipitation on the slopes of a low range of hills in southern Illinois than on the flat surrounding farmlands. McKay (1976) reported that, in hydrology, most important isolines are consistent with topography and that isohyetal maps are no exception in this regard. Colton (1976) used a mathematical model to simulate the orographic precipitation process in order to determine the spatial distribution of precipitation in the mountainous areas of the central Sierra Nevada of the U.S.A. Trend surface analysis with altitude as a variable was used successfully by Solomon et al. (1968) to map MAP in Newfoundland, Canada. Dunne and Leopold (1978) presented data from Vermont, U.S.A. in which the MAP rose by 73 mm per 100 m rise in altitude. March et al (1979)

showed an increase in MAP of 86 mm per 100 m increase in altitude at Coweeta, North Carolina, U.S.A. According to Dingman (1981), altitude has a major influence on the hydrology of New Hampshire and Vermont, U.S.A., due mainly to the marked increase in MAP with altitude.

2.1.2 In South Africa

Whitmore (1968) showed that a regression equation which used altitude, aspect, continentality and longitude could explain most of the variation in MAP in an area of the southern Cape. Schulze (1976; 1979) used trend surface analysis with physiographic variables of altitude, distance from a physiographic barrier and alignment of topography to map MAP and other precipitation statistics in the Drakensberg. Buys *et al.* (1979) used altitude contours as a guide in drawing maps of expectancy of monthly rainfall for Southern Africa. This was an indirect form of trend surface analysis. Alexander (1980) reported that the rainfall on the eastern escarpment of Southern Africa is influenced by the local topography and is therefore predictable in its location and extent. Physiography was used successfully by Hughes (1982) in a trend surface analysis to map the MAP of the southern Cape coastal mountains in the vicinity of the Wilderness. De Villiers and Bond (1982) employed physiographic variables in a trend surface analysis which included continentality and altitude to map MAP in the Durban/Pinetown area.

Conclusions by London and Emmitt (1986) who reported the evaluation of more than 3 500 radar storm cells in the eastern Transvaal Drakensberg have provided evidence for the following statements related to the effect of physiography on rainfall:

- (a) The escarpment is in general a preferred region for the initiation of convective storms.

- (b) Within the broad escarpment area there are "hot spots" or locations with distinctively higher radar echo activity.
- (c) All major "hot spots" are on the south east faces of ridges, i.e. the faces directed towards the predominant east-southeast surface flow of moisture laden air masses.
- (d) The ruggedness of the topography appears to have some influence on the lifetime and movement of storms.

Ongoing research by London and Emmitt (1986) is addressing the following questions, viz.

- (a) whether recognizable preferred topographical locations for storm initiation exist, and
- (b) whether the topography can affect the evolution of these storms in ways that are similar to the expected response to seeding.

2.2 Need for establishment of a physiographic data base

In the light of the above review it was evident that trend surface analysis, which regressed MAP inter alia on physiographic variables, would be one of the techniques to be used in this project. A major exercise in gathering physiographic data was therefore initiated. This exercise involved the following facets:

- (a) The exact location and altitude of each rainfall station in the data base had to be checked. This was to be undertaken using topographic maps of scale 1:50 000 and the location and altitude information would be computer coded.
- (b) Point altitude values on a 1 minute of a degree grid had to be gathered, checked and computerised for Southern Africa.

- (c) Computer programs had to be developed to calculate the distance from a grid point to a pre-determined mountain barrier or to the coastline.
- (d) A computer program had to be developed to allocate each point in the entire grid system to a preselected rainfall region.
- (e) Computer programs had to be developed to generate physiographic variables such as exposure, surface roughness and aspect for the altitude grid covering Southern Africa.

Each of these facets will be elaborated on in more detail in the following sections.

The large mass of data involved in each of these tasks necessitated a high degree of computerised automation in this project. This automation is consistent with one of the objectives of the study, namely to automate the techniques. Nevertheless, a great deal of manual data gathering and checking was also necessary, especially in establishing the location of rainfall stations and extracting point altitude data off 1:50 000 maps..

2.3 Establishing the location of rainfall stations

It was important to establish the exact location of each rainfall station since its spatial location would "fix" its altitudinal location which was often the important variable in the regression analysis. The latitude, longitude and altitude of all the rainfall stations were therefore checked on maps of scale 1:50 000 with a contour interval of 20 m. Most of the positions and altitudes of rainfall stations had been estimated by the South African Weather Bureau (SAWB) from 1:250 000 maps which have contour intervals of 500 feet. This information is reported in SAWB Yearbooks and SAWB computer files. The large number of stations involved made this a very time consuming

undertaking. In many instances intelligent guesswork was necessary to "place" the raingauge before its altitude could be estimated. The information gathered in this exercise had to be subjected to numerous computerised and hand checks to ensure that it was as error free as possible.

The stations used from the SAWB were those which appear on their daily rainfall data base. A complete list of all the stations together with latitude, longitude (to nearest minute), altitude, SAWB reference number, name, starting year of record, ending year of record, and number of complete years of data was compiled. New latitude and longitude (to a resolution of 0,1 minute) and "new" SAWB numbers were included if the station was relocated. These data are contained on the file INDEX87 which is kept in the common area of the Computing Centre for Water Research (CCWR) computer file system and may be accessed by any user of the CCWR. The original SAWB numbers have, however, also been retained for all SAWB stations as the revised station latitudes and longitudes (to 1 minute of a degree resolution) may not correspond with the SAWB numbering system. The explanation for this is that when the stations were originally "located" they may have been incorrectly placed by the SAWB due to the map resolution and other problems at the time of allocation of the locations. For the sake of continuity with the existing data files at other institutions and existing SAWB publications it was therefore decided to retain the original SAWB numbers but also to record the exact location to 0,1 of a minute and the revised SAWB number in another field on INDEX87.

Some of the rainfall stations' locations were difficult to identify exactly on the 1:50 000 topographic maps. The amount of time spent deciding the location of such stations was dependent on:

- (a) the length of record at the station,

- (b) the date on which the station opened and closed,
- (c) the number of other stations in the vicinity, and
- (d) whether the station's exact location and hence altitude would be likely to improve the precipitation estimation in that area.

Where the answers to all these questions were positive and station location and altitude proved to be a problem, then assistance was sought from the SAWB. This analysis yielded 106 stations which required further investigation by the SAWB.

The following points summarise the types of problems which were encountered:

- (a) The name given to the rainfall station could not be found in the vicinity of its co-ordinates on the 1:50 000 map and therefore no "positive" identification could be made,
- (b) several identical place names were given but the places were not at the co-ordinates given by the latitude and longitude of the station according to the supplying organisations' records, or
- (c) no name was given to the rainfall station by the supplying organisation, in which cases no "positive" identification of the rainfall station could be made.

The general rule which was followed in locating the point position of the rainfall stations was that they were most likely to be near a road, farm house or village centre. Gauges were unlikely to be remote from places of convenient access. Individual farmsteads are shown on the 1:50 000 maps and this assisted place name identification considerably.

The latitude and longitude of each station was converted into a SAWB number according to their sector and co-

ordinate numbering system. This served as a check that the latitude and longitude and SAWB number had been entered correctly. A further check on station altitude was made by comparing them with the grid points of altitude in the vicinity (Section 2.5). This check served to highlight gross errors in both the station and grid point altitudes. The final check on the station altitudes was made in a residual analysis of observed and estimated rainfall at the stations (Chapter 5). Frequently, the altitude was one of the variables most highly correlated with rainfall and as such an incorrect altitude would result in a high residual. In all such cases of high residuals altitudes were checked again manually on the 1:50 000 maps.

While the above section has referred to the stations on the SAWB files, the same location altitude checking procedure was followed for all the stations from other organisations used in this study (Chapter 3).

2.4 The grid system of altitudes

The primary data in the grid system consisted of latitude, longitude and altitude at each 1 minute by 1 minute of a degree point in the grid. This grid system formed the framework for the entire physiographic and regional information base, and subsequently for the trend surface analyses of rainfall.

The grid covers all of South Africa plus the TBVC States, non independent National States and Lesotho. Approximately 75 per cent of these data were extracted by the staff of this project. Other substantial contributions have been from the University of the Witwatersrand's Hydrological Research Unit (HRU), the Department of Land Surveying and Mapping at the University of Natal (LSM), the Department of Agriculture and Water Supply's Soil and Irrigation Research Institute, Western Cape Region (SIRI).

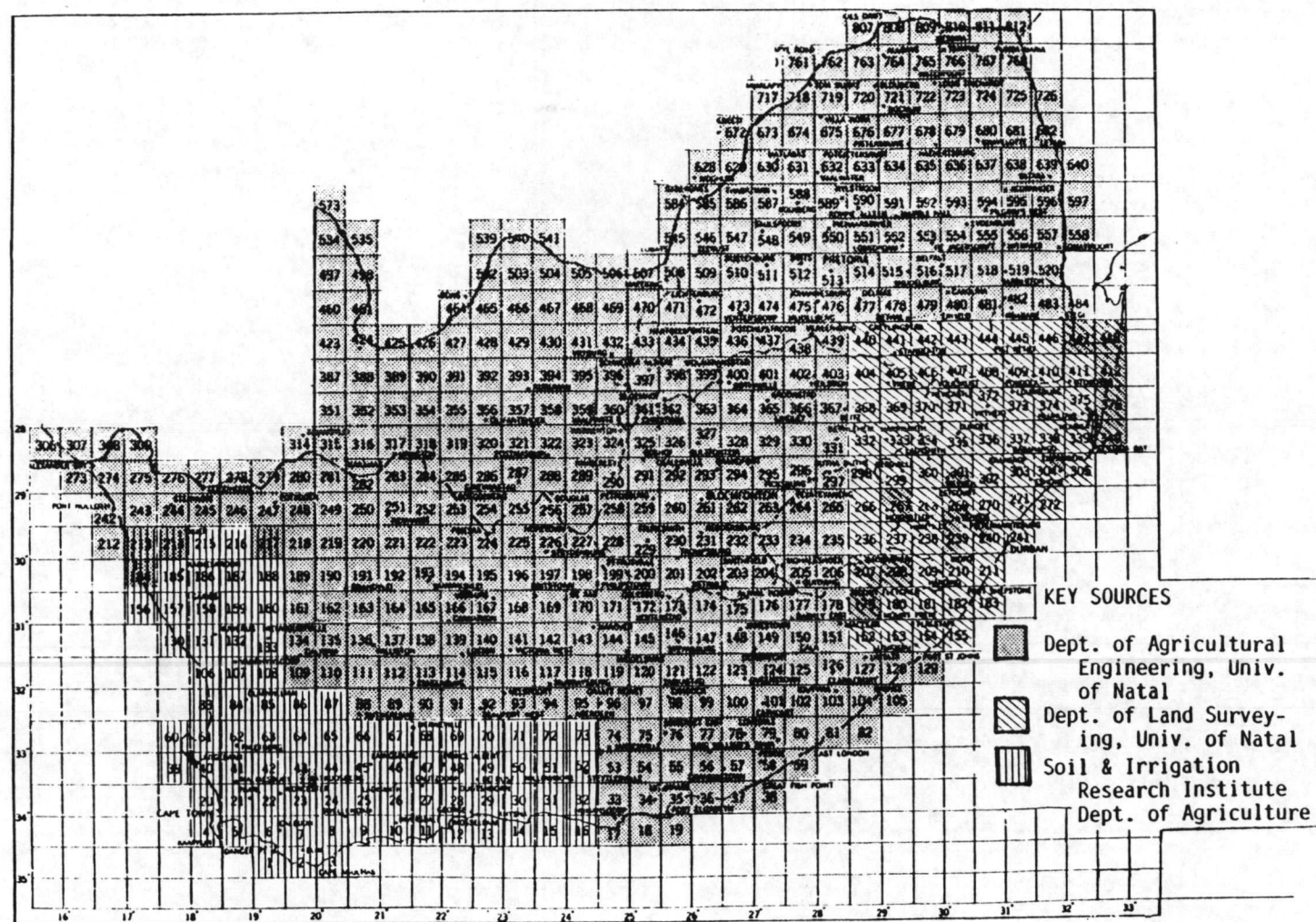


Figure 2.1 Altitude grid data over Southern Africa (sources of information)

All these data were combined into a single file named SAGRID. Figure 2.1 shows the areas covered by altitude data from the various sources mentioned above.

The original HRU data was extracted manually from 1:250 000 topo-cadastral maps with a contour interval of 500 feet. The latest edition of maps of this scale have 50 m contour intervals.

The SIRI, LSM and Department of Agricultural Engineering data were extracted from 1:50 000 topographic maps with 20 m contour intervals which are easier to follow in the mountainous terrain. These data were therefore taken to supersede the HRU data set for the Western Cape, Natal Drakensberg and eastern Transvaal Drakensberg areas. However, the altitudes were still compared with those from the HRU data set in order to check for gross errors.

Figure 2.2 shows how the altitudes from maps of scale 1:50 000 differed from those of the HRU. A dot depicts each grid point in which the HRU and altitudes from more detailed maps differed by more than 100 m. It is significant to note that the most important water resources areas of Southern Africa, where the rainfall surfaces are complex, are covered by the grid system extracted from the more detailed and thus exact maps with 20 m contour intervals. The original (HRU) altitude values were retained on file but displaced to another field in the record. On completion of the MAP maps it was realised that it would be preferable to use only altitude data extracted off 1:50 000 maps, since the use of altitude data which had 500 foot contour intervals led to unsatisfactory results in some areas. The altitude for all such areas (208 SAWB sectors) were then extracted from the 1:50 000 maps by the staff of this project.

The altitude data were derived in the following manner. A grid of 1 minute of a degree was overlain on the 1:250 000 or 1:50 000 map and the altitude at the grid intersection point was noted. No attempt was made to average altitudes for a 1 minute square or to find a "representative" altitude in the area. It is entirely possible therefore that

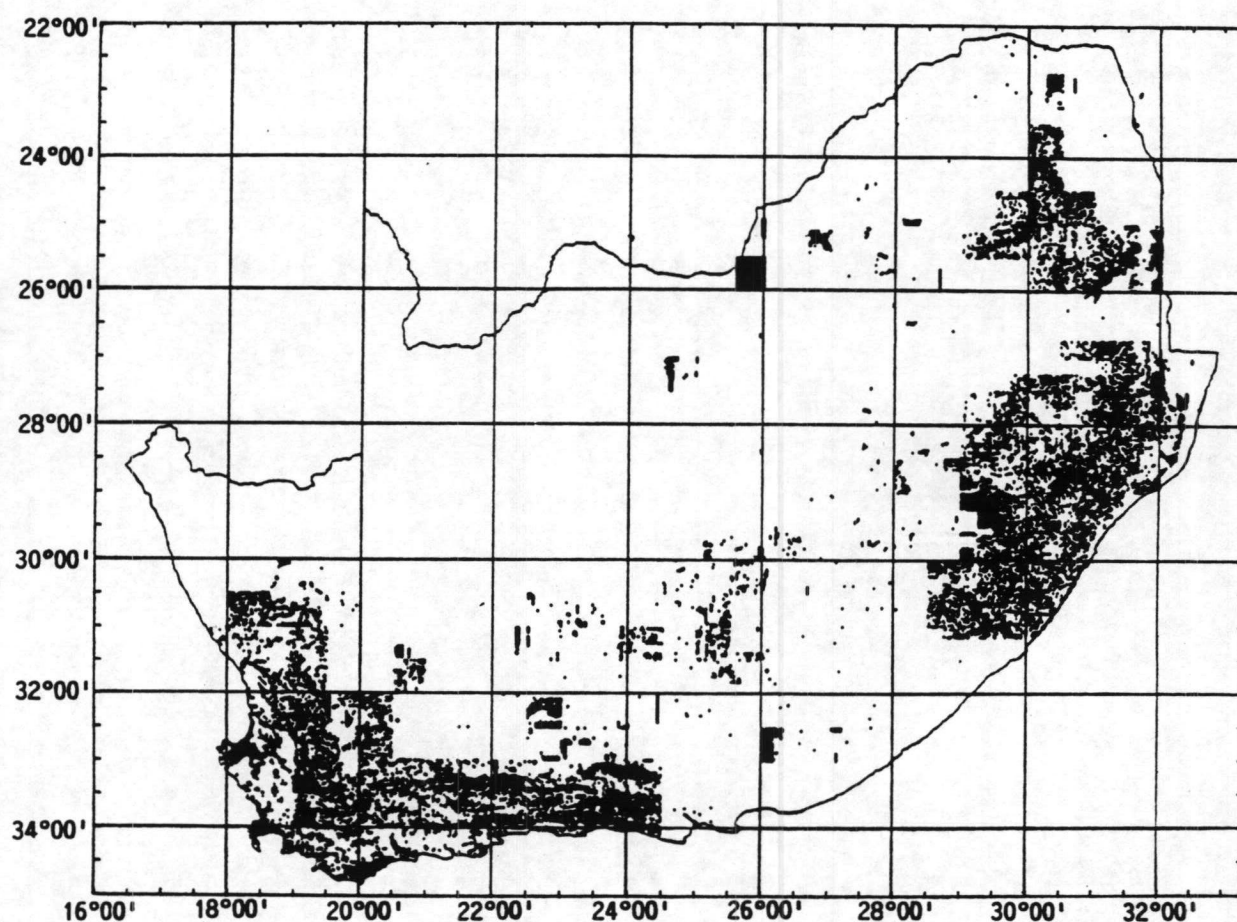


Figure 2.2 Points at which the altitude from the HRU data set differed by more than 100 m from altitudes extracted off the 1:50 000 maps

small topographical features will not be reflected by this digital image of altitudes which in many areas is adequate but in other areas may not be detailed enough to depict all the features of a complex topography.

The choice of resolution for the grid (or grid size) is important, according to McKay (1976). If the grid is too coarse the resultant trend surface lacks resolution, as has been demonstrated by Pawling (1973) and Schulze (1977). If, on the other hand, the grid resolution is too fine it results in an unmanageable amount of data which have to be processed. The rainfall station network density should also be considered in determining the grid's resolution. It should be noted that one cannot create new rainfall station data by introducing a very fine grid. After due consideration of the aforementioned factors, a grid size of 1 minute by 1 minute of a degree was therefore chosen for the entire study area of Southern Africa.

Thus the grid size is adequate in flat areas. However, in areas of more complex terrain problems were experienced, in this project. These problems stem from the fact that the grid altitude points are the estimated altitudes at the exact intersection of the grid points. The rainfall station altitudes on the other hand were recorded at the station locations which were estimated to the nearest 0,1 of a minute. The rainfall station locations were then converted to the nearest minute of a degree for all subsequent processing, whilst still retaining the estimated altitude of the original location. This introduced a problem when the altitude of a grid point and a station point with the "identical" latitude and longitude were compared. Figure 2.3 illustrates this problem by showing the rainfall station (RS) at latitude $27^{\circ}32'$, longitude $30^{\circ}15'$ (to the nearest minute) to have an altitude of 930 m, while the grid point latitude $27^{\circ}32'$ and longitude $30^{\circ}15'$ has an altitude of 1 010 m. In order to circumvent the problems

posed by this situation in the application of the trend surface, it was decided to replace the altitude of the grid point with that of the station point in all such cases, which were restricted generally to mountainous areas.

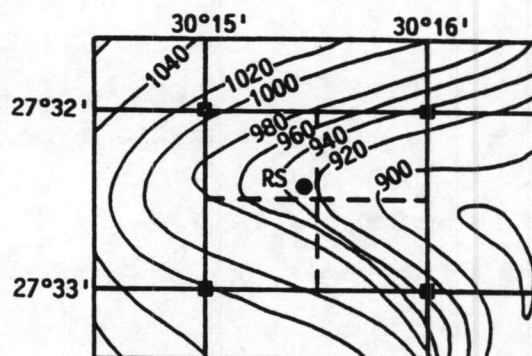


Figure 2.3 Illustration of the grid point altitude/rainfall station altitude discrepancy

Although nearly 70 per cent of the original grid point altitude data were given to this project from other sources, the large task of loading and checking these data on the University of Natal computer system still remained. In the case of the HRU data, these were contained on computer cards, one per one grid point. These data have undergone comprehensive testing to establish their completeness and accuracy.

2.5 Checking the altitude of grid points and rainfall stations

The two extensive tasks of producing and checking the altitude data file for grid points and rainfall stations took up a large proportion of the time of this project. In view of the importance of these data both to this project and for numerous other applications, it is pertinent to review the altitude data gathering and checking procedures. This review will give the reader an idea of the quality of this large data set and will possibly be of interest to scientists who wish to assess the appropriateness of these data

for other applications.

The large amount of data involved implied that besides an extremely time consuming method of checking, such as an independent reassessment of the altitudes (as was undertaken for the south western Cape and Natal) followed by a direct comparison of the two surfaces, there was no other way of checking grid altitudes, other than searching for contradictions in and exceptions to the data.

The grid point and rainfall station altitudes were produced entirely separately. Therefore a comparison of altitudes at each station point provided a useful independent check at 9 409 points. Differences of more than 150 m were listed for further investigation from 1:250 000 maps. At the same time the following checks were being conducted on the altitude grid:

- (a) Initially the 900 points in each SAWB sector were printed out with a single classification of 0's at all points which had altitudinal data. This was to check whether points had been omitted and that no points were duplicated.
- (b) A classified numerical image of altitudes, an example of which is shown in Figure 2.4, was produced for blocks which covered each of the 70 topographical maps at 1:250 000 in the study area. The digits on the numerical image changed at the same altitude class boundaries as the colours on the 1:250 000 map series. The checking of topographical features such as hills, valleys, spurs and plains was facilitated in this manner.
- (c) The altitude data from the HRU, LSM and SIRI data sets were compared as a further independent check. Large differences in altitude were investigated and corrected.

- (d) Towards the end of the project the plotting facilities at the CCWR were used to plot contour maps from the extracted altitude points. The originals and the computer drawn contour maps were then compared, feature by feature.
- (e) Altitude was generally one of the variables which was most highly correlated with MAP. Hence, in the regression analysis the outliers of MAP were often caused by incorrect altitudes at rainfall stations. Such cases were investigated and corrected.

This undertaking was extremely time consuming and required diligence and tenacity, since in all cases each problem had to be treated individually and the appropriate action taken. Several thousand corrections were effected during these checks.

2.6 Exposure, roughness, aspect, rainbearing wind direction and continentality indices

Apart from providing the altitude data this data set formed the foundation of a suite of techniques which produced other physiographic variables at each grid point. These variables were exposure, roughness and aspect indices, continentality and distance to a mountain barrier.

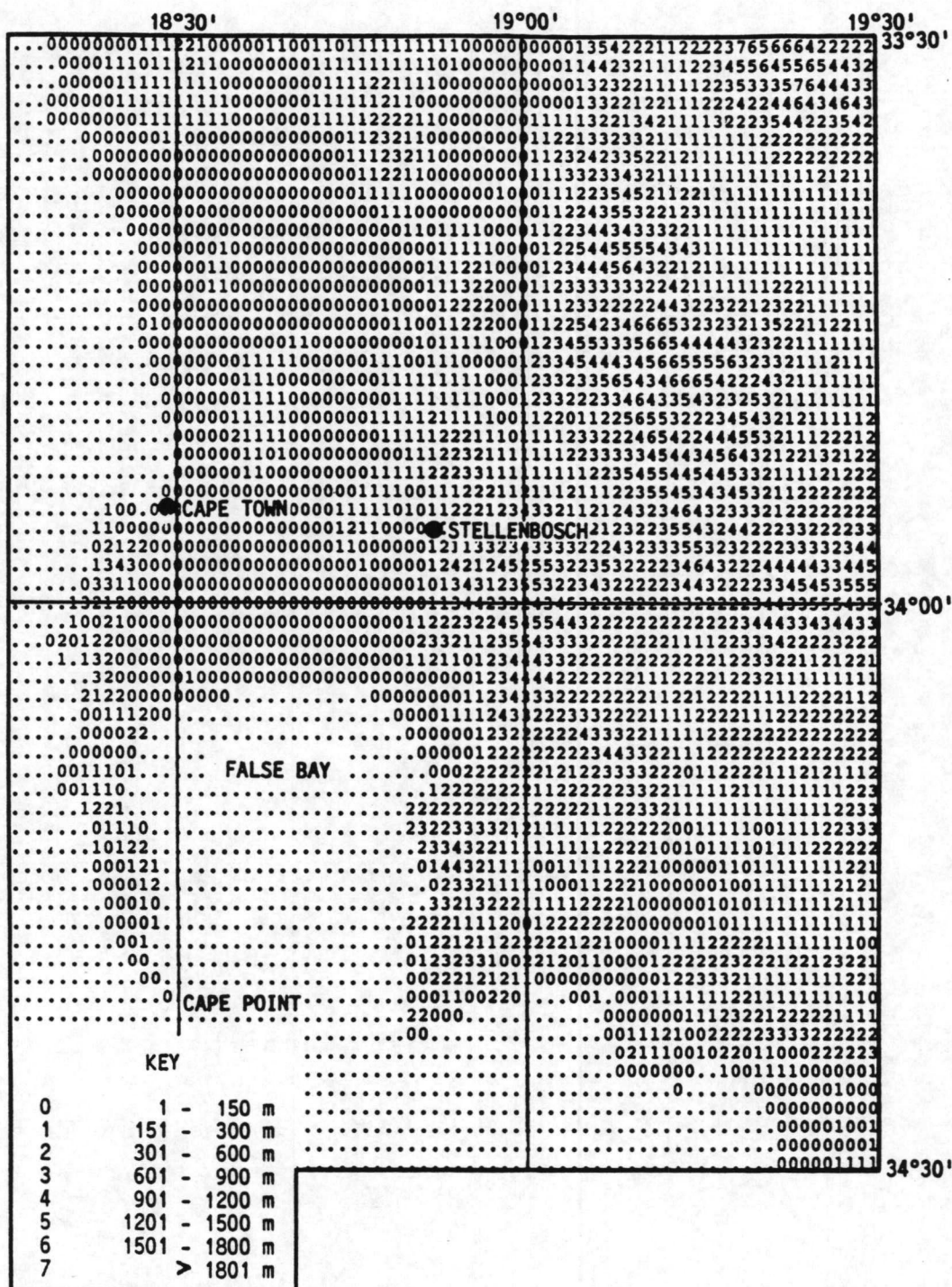


Figure 2.4 Classified image of altitudes in the south western Cape

2.6.1 Exposure index

The exposure index was selected to reflect a trend observed by Hutchinson (1968) and Hughes (1982) in which hills that are exposed relative to the surrounding countryside receive more rainfall than would a nearby area at the same altitude but which is not exposed. An exposure index suggested by Seed (1987) was calculated using a program which placed a 5 minute by 5 minute mask over the area and compared the altitude of the surrounding points to that at the centre point of the mask. The exposure index is a count of the number of points in the 5 minute by 5 minute mask which are lower than the centre point of the mask. This index ranges from 0 to 24. The centre point of the 5 minute by 5 minute mask is placed in turn on all the grid points in the image.

2.6.2 Roughness index

The roughness index reflected the steepness and complexity of the topography and in this way complemented the aspect index. In addition the roughness of the terrain causes updraughts and turbulence which may influence the rainfall. The roughness index was formed by calculating the standard deviation of altitudes of a 5 minute by 5 minute mask. This standard deviation was used as the index of roughness for the centre point of the mask. A roughness index was calculated for every station and every point on the grid of point altitudes.

2.6.3 Continentality index

The effect of continentality on MAP has been demonstrated inter alia by Whitmore (1968), Hutchinson (1968:1969) and De Villiers and Bond (1982). It was therefore decided to

develop an index of continentality for inclusion in the trend surface analysis.

The determination of an index of continentality involved digitising the coastline of South Africa. The coastline was thus represented by a series of 131 straight lines. The limited number of straight lines defining the coastline was a compromise between accuracy and computation time of the routine which calculated continentality. A very fine resolution (that is a large number of lines) defining the coastline would have led to an excessive amount of computing for negligible gain in accuracy, of the distance to the sea. A computer program was developed to calculate the perpendicular distance from a point to a straight line (defined by the latitudes and longitudes of its end points). Furthermore this program had to distinguish whether the perpendicular from the point to the line intersected between the end points or on the extension of the line. The continentality calculation was conducted for every point within approximately 300 km of the coast. It was assumed that the effect of the sea would be negligible further "inland" than 300 km. The same technique used in the determination of continentality could also be used to determine distance to a mountain range, distance from a ridge defining a rainshadow or, for that matter, from any predetermined line or point.

2.6.4 Distance from a physiographic barrier

The dependence of rainfall on distance from a physiographic barrier has been reported by Peck and Brown (1962) in the U.S.A., by Hutchinson (1968) in New Zealand, by Wicht et al. (1969) in the south western Cape and by Schulze (1979) in the Drakensberg. The sharp rainfall discontinuity at the edge of the Drakensberg escarpment was shown by Schulze (1979) and is illustrated in Figure 2.5.

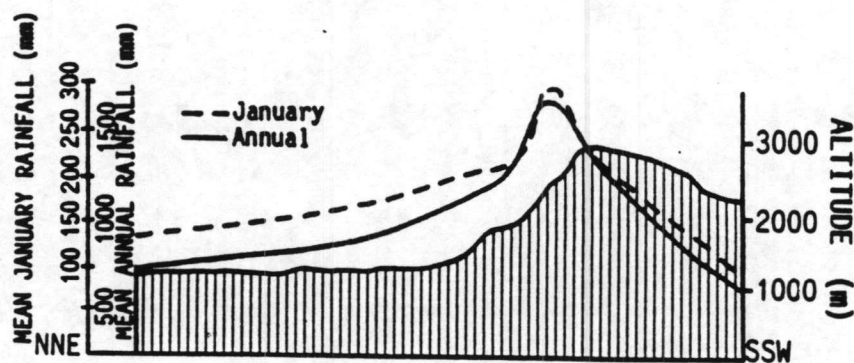


Figure 2.5 Rainfall/altitude relationship along section from Bergville in the NNE to Mothelsessane in the SSW (after Schulze, 1979)

It should be noted from Figure 2.5 that the rainfall peaks before maximum altitudes are attained and then decreases rapidly in the lee of the Drakensberg - a phenomenon not uncommon in mountainous zones.

2.6.5 Aspect index

Hillslopes which face rainbearing weather systems receive more rainfall than leeward slopes. The term aspect, which is usually used in the context of solar radiation loading, has however been used in this study to describe the direction of the hillslope in relation to the rainbearing weather mechanisms.

Aspect is an important factor in the rainfall/topography interaction (for example Hutchinson, 1968; Whitmore, 1972; Hughes, 1982). An index of aspect was thus sought. The idea which subsequently became the fully automated routine described below, developed out of research by England (1971).

At the outset it should be remembered that the grid points of altitude are spaced 1 minute of a degree apart (approximately 1,6 km); hence small topographic featural changes will not be reflected in the aspect index. However, large trends in aspect will be discernible as may be seen in Figure 2.6, which shows the classified digital image of aspect for the south western Cape area.

Aspect was determined by a procedure which involved two major stages. First, a plane was fitted by regression analysis through the 9 points of a 3 minute by 3 minute mask of altitudes. This plane yielded a function of the form:

$$\hat{z} = b_1 + b_2 x_1 + b_3 x_2$$

where

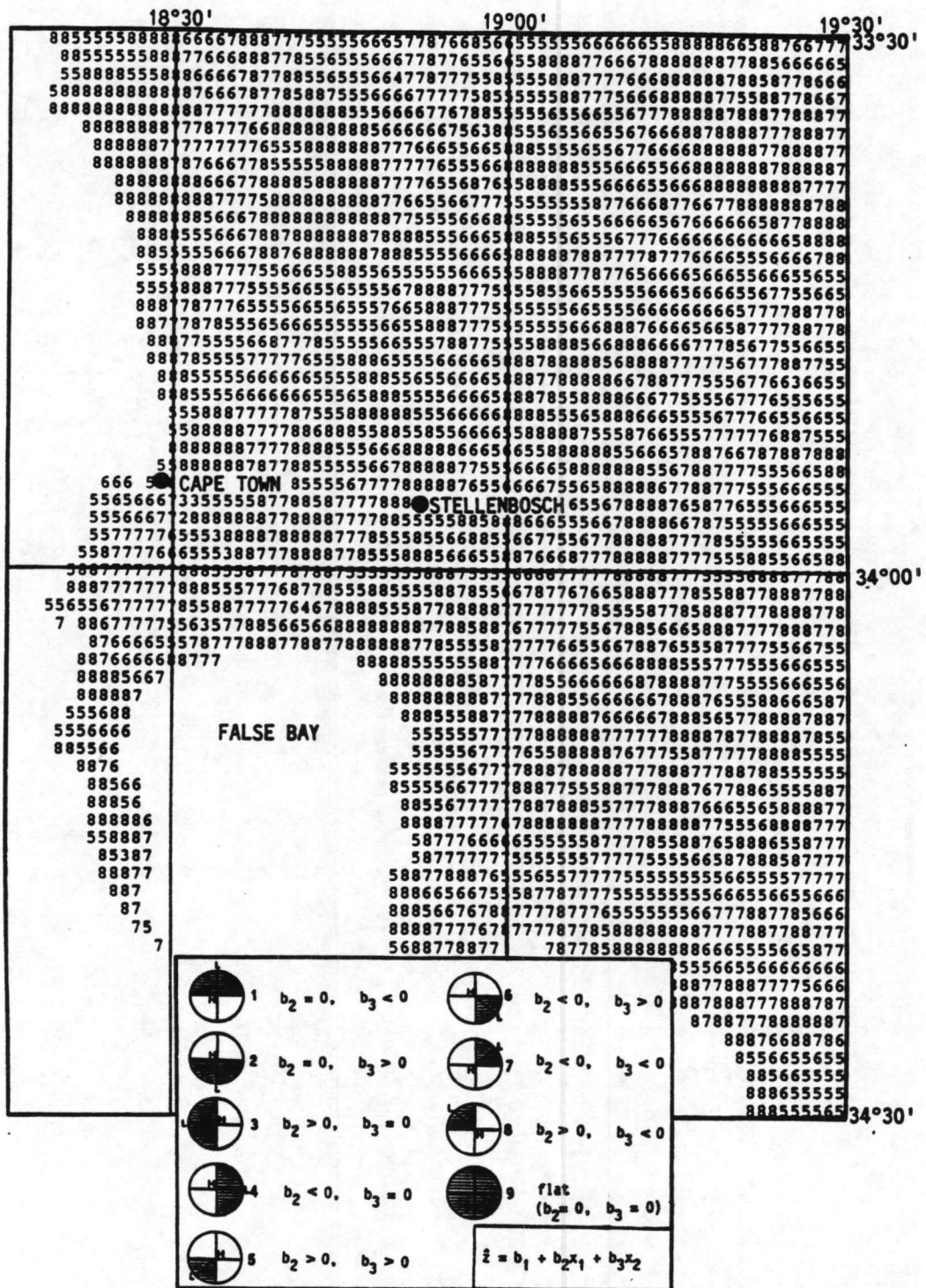
$\hat{z}$, x_1 and x_2 = altitude, longitude and latitude respectively

b_1 , b_2 and b_3 = regression coefficients.

In matrix terms the determination of the coefficients (B) requires the solution of the following ;

$$(X^T X)^{-1} \cdot X^T \cdot Z = B$$

The origin of the mask is taken at the south west corner and the co-ordinates of the nine points are then calculated in terms of this origin. The co-ordinates thus are 0 or 1 or 2. The calculation of $(X^T X)^{-1}$ was performed once and thereafter the value of the resultant matrix inserted into the expression as a constant. This led to a saving in time when one considers that this routine has to be performed for each point in the gridpoint altitudes for the entire study area of Southern Africa (451 700 points). Aspect was then classified into one of nine directions as depicted schematically in Figure 2.7, depending on the value of b_2 and b_3 . The shaded segment displays the aspect if one imagines the segment dropping towards the circumference of the circle as shown by the high and low altitude symbols L respectively.



NOTE: The numerical classification system for aspect is explained overpage.

Figure 2.6 Classified digital image of the aspect index in the south western Cape

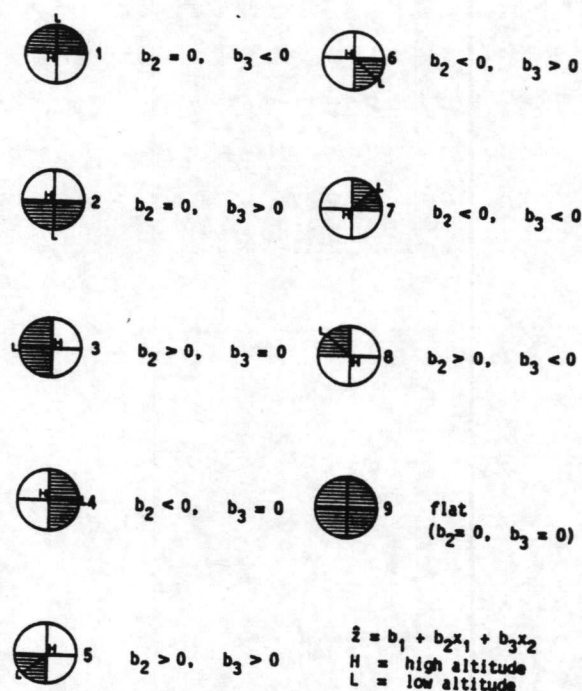


Figure 2.7 Schematic diagram of the numerical classification system for aspect

The aspect index was an essential element in the determination of the rainbearing wind direction index described below.

2.6.6 Rainbearing wind direction index

This index was developed to accommodate differences in MAP caused by the direction of the rainbearing wind in relation to the aspect of the slope. It was only necessary to provide for two wind directions, since the regression analysis accommodated rainbearing wind systems from the opposite directions by manipulating the sign and magnitude of the coefficients of the rainbearing wind direction index in the regression of MAP on the physiographic variables. The chosen wind directions were south and east. A description

of the rainbearing wind direction index is assisted by reference to Figure 2.8. The inner circle contains the rainbearing wind direction index for southerly winds and the outer circle the index for the easterly rainbearing winds. South facing slopes thus receive the highest index for a southerly rainbearing wind and east facing slopes for an easterly rainbearing wind.

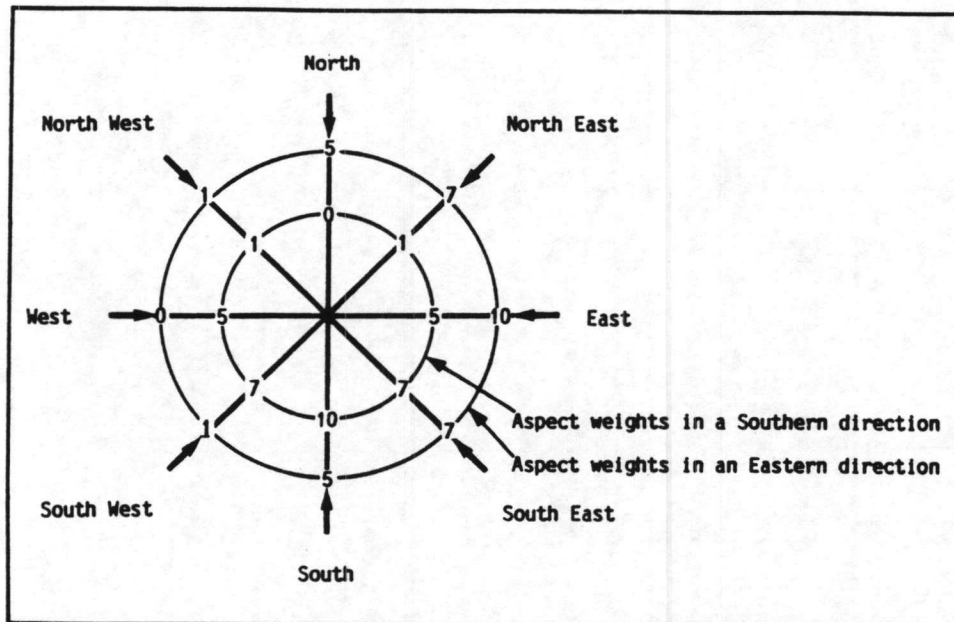


Figure 2.8 Illustration of the development of a rainbearing wind direction index

Should the rainbearing winds be from the south for example then the regression equation would respond by giving positive coefficients to the rainbearing wind direction index variable. The effect of this would be to increase the MAP as the aspect moves more into line with the winds. However, if the rainbearing winds are from the opposite direction i.e. north then the coefficients of the rainbearing wind direction index variable in the regression would be negative, thus reducing the MAP as the rainbearing wind direction index increases.

The rainbearing wind index for each station or grid point is dependent on the aspect index which was described in Section 2.6.5, and the direction of the wind as shown in Table 2.1.

Table 2.1 Rainbearing wind direction indices

Southerly rainbearing wind direction								
Aspect	S	SW	SE	E	W	NE	NW	N
Rainbearing wind direction index	10	7	7	5	5	1	1	0
Easterly rainbearing wind direction								
Aspect	S	SW	SE	E	W	NE	NW	N
Rainbearing wind direction index	5	1	7	10	0	7	1	5

The physiographic data sets and indices have been described in terms of altitude, exposure, roughness, continentality, distance from a physiographic barrier and aspect in this chapter. The next chapter deals with the rainfall data file which served as the major input into techniques for presenting MAP cartographically.

CHAPTER 3

THE RAINFALL DATA FILE

The nature of this project dictates that it should utilize the most comprehensive precipitation data set available in Southern Africa. The logistical problems involved in collating the daily rainfall data from a number of sources into one comprehensive data file are considerable. Computer networking between the mainframe computers which host these data was in its infancy in Southern Africa while this project was underway. One may therefore appreciate the reluctance of researchers and consultants to gather a comprehensive rainfall data base for any "once off" or time constrained study. Consequently, at the time of the inception of this project, such a "complete" collation of these various data had not been accomplished. This project provided the stimulus to assemble all the data sets mentioned below (Section 3.1), into a single, readily accessible and flexible data system in which selected data products could be extracted easily and speedily.

The large size of this data set, combined with the project's overall aim to automate the whole process, necessitated the development of many computer programs to perform functions which would, in smaller localised studies, have been performed manually.

3.1 Sources of rainfall data

The project has depended on many organisations and individuals for its data. Daily rainfall data in the Republic of South Africa, Lesotho, and Swaziland are assembled by a number of organizations. The South African

Weather Bureau and the Department of Agriculture and Water Supply are the principal collectors of daily rainfall data. Other sources are the South African Forestry Research Institute of the Department of Environment Affairs, provincial parks boards, organized agriculture and in particular the South African Sugar Association, municipalities, and mines. In addition, a large number of individuals record daily rainfall but do not submit these data to any of the aforementioned organizations. Daily rainfall has been recorded at approximately 9 500 stations in Southern Africa. A breakdown of the sources of these data and locations of the rainfall stations is presented in Table 3.1.

Table 3.1 Breakdown of the sources and location of rainfall data

SOURCES	NATAL	TRANSVAAL	ORANGE FREE STATE	SWAZILAND	LESOTHO	CAPE	SOUTH WEST AFRICA
SAWB	537	2 411	1 324	62	38	2 929	291
AWS	130	95	36	2	0	233	0
FOR	5	0	0	0	0	0	0
SASA	126	3	0	11	0	2	0
PRIV	234	212	111	3	3	604	7
	1 032	2 721	1 471	78	41	3 768	298

Key to Sources

SAWB	South African Weather Bureau
AWS	Department Agriculture and Water Supply
FOR	Department Environment Affairs (Forestry) stations which are not on SAWB files
SASA	South African Sugar Association
PRIV	Private individuals, organised agriculture, mines and municipalities

The Department of Agricultural Engineering, University of Natal is very appreciative of the abovementioned organisations and individuals within these organisations for their enthusiastic co-operation in supplying rainfall data. In addition to the established sources of data, another large source of data was tapped through a wide ranging appeal for rainfall data, launched on radio, in the press and by letter to organised agriculture, other bodies and individuals. The response was excellent, data having been received from 1 174 new stations. Included in these are an additional 241 stations from the SIRI in the western Cape which were gathered recently from farmers by Myburgh and Hoon (1985). The area from which these latter stations were drawn is presented in Figure 3.1. Some of these records contain daily rainfall data of long duration (70 years or longer) and/or are from locations where hitherto the raingauge network has been sparse. When this search for new data was embarked upon, the philosophy of "casting the net wide" was adopted in the hope of obtaining some important stations. Arising out of the excellent response, the handling of which occupied a great deal of the project staff's time, the data for a number of new key rainfall stations were forthcoming. An example of the degree to which these new data augmented those from the established data collection agencies may be seen in Figure 3.2.

3.2 Quality of the rainfall data

In Southern Africa volunteers operate most of the daily rainfall stations, the data from which are collated by the principal contributors of data to this project, viz. the SAWB, the Department of Agriculture and Water Supply and the South African Sugar Association. It is possible that human error occurs in the measurements at these stations and there is no means of checking this, besides the techniques discussed in the following section. At some stations it was apparent when the recorder's annual leave was

taken and many months of data had to be omitted in the aggregation to monthly totals and then synthesised later. The SAWB cautions against errors which may occur in transposing their data into computer compatible form. The SAWB claim 95 - 98 per cent "accuracy" in the data acquired from them. Subtle errors will unfortunately not be detected by the error checking procedures discussed below. However, gross errors of data capture would probably have been detected, and either corrected or "quality" coded for omission from subsequent statistical analyses.

The SAWB employs a system of quality codes for their daily rainfall data. These codes and the quality codes used to flag the monthly rainfall totals in this project are presented in Table 3.2. Only the uncoded, \$ and synthesized monthly data are used to calculate the monthly and annual statistics which are used finally in the mapping.

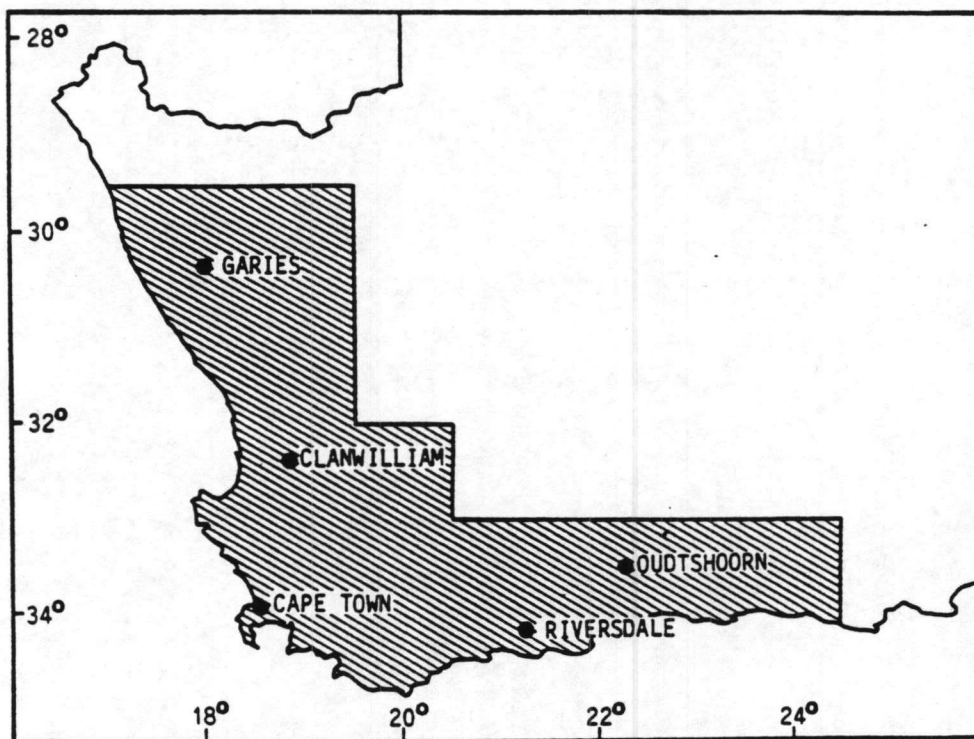
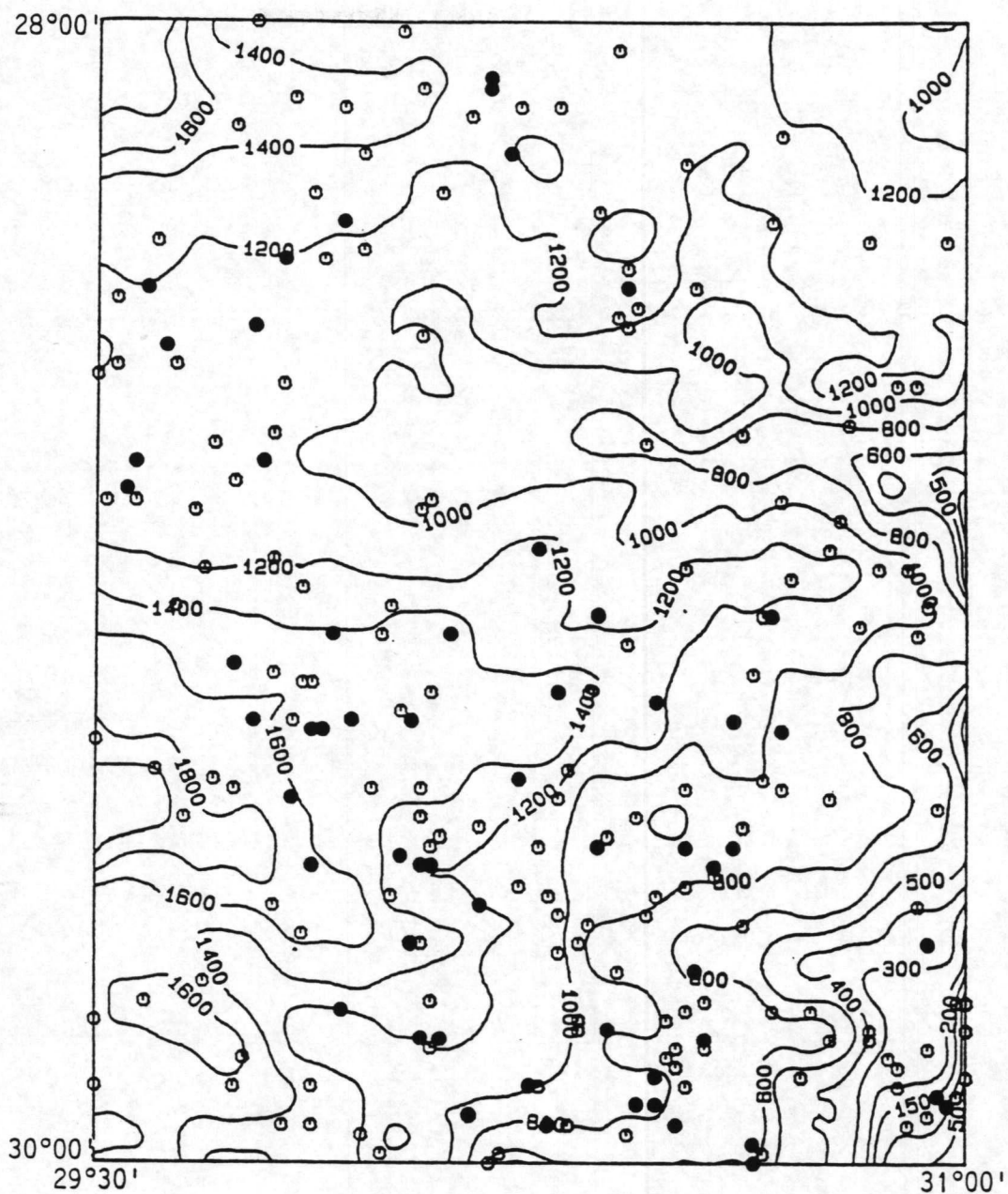


Figure 3.1 Source area for new rainfall stations from SIRI western Cape region



KEY

- 1200— altitude (m)
- (58) Newly found stations with > 15 years rainfall record
- (162) SAWB or Department of Agriculture and Water Supply stations > 15 years

Figure 3.2 Location of the "new" and "old" long term (>15 years) stations in an area of Natal/Orange Free State

Table 3.2 Quality codes used in the monthly rainfall data file and the SAWB daily equivalents

Monthly data code	Description of data quality	SAWB daily data code
Uncoded (blank)	reliable daily rainfall	uncoded
\$	accumulated daily rainfall data (total <u>reliable</u>)	-6666 80000+ 90000+ 70000+
*	accumulated daily rainfall data (total <u>unreliable</u>)	60000+ 50000+ 40000+
#	estimated daily rainfall data (not accumulated and not reliable)	30000+
M	missing daily rainfall data	-9999

There is naturally doubt as to the quality of some of the data which were obtained by this project from private individuals. However, when one considers the following points, viz.

- (a) that such data were only used when no data from established agencies were available in the area,

- (b) that these data were subjected to the same screening tests as all the other data,
- (c) that the respondents who supplied data obviously had a personal interest in their rainfall data, as was evident from the diligence of many of the responses and
- (d) that if the philosophy be accepted that a diligent observer at a point will obtain a more accurate estimate of the precipitation than interpolation or extrapolation techniques from surrounding stations would at that point, then the value of these "new" data is appreciated.

The daily data were supplied by the organizations mentioned above on various media, such as magnetic tapes, computer printouts, punch cards, or photocopies of individuals' records. These data were processed to yield monthly and annual statistics of rainfall. Monthly rainfall data have much relevance in Southern Africa since most of the large-scale hydrological planning on the subcontinent is carried out using models that have a monthly time base (for example Pitman, 1973, and Pitman et al., 1981). Ninety per cent of the daily rainfall records are obtainable on magnetic tape from various organizations. The daily data on these tapes are generally quality coded according to the classes shown in Table 3.2. The quality of the data which aggregate to form the monthly totals is of fundamental importance. Hence, it was decided to include all of the data together with an index of quality. The individual using the data is thus able to make a decision as to the acceptability of the data, since the user is considered to be a more discerning judge of the appropriate data quality level for a particular use than this project's staff.

No correction has been applied to the recorded rainfall as

a result of estimates of raincatch deficiencies or the contribution provided by fog. The magnitude of these deficiencies is dependent upon specific location and site factors. Fog contribution and raincatch from protected/shield gauges measured at only a handful of selected (usually research) sites and consequently the inclusion of such data would make those point values incomparable with those from standard gauges. Discrepancies in point rainfall estimation are reviewed elsewhere in some detail, inter alia by Rodda (1967), Schulze (1975), Ward (1975) and De Villiers (1980).

3.3 Checking the rainfall data

Checking such large volumes of rainfall data required programs which tested for values which appeared erroneous in magnitude and in their temporal context. Checks for spurious values which could have been errors were conducted at the daily, monthly, annual and MAP levels. These "erroneous" data were investigated and flagged if necessary. Other statistical indicators such as the coefficient of variability (CV) of monthly totals and occurrences of zero rainfall in months normally associated with high rainfall were also used as signals to alert data processing staff to inspect the data more closely.

The SAWB magnetic tapes which contain daily rainfall data up to May 1987 were used as a primary source of rainfall data in this project. Much of the rainfall data in this study will therefore contain the records from the severe drought of the early 1980's. The daily data were subjected to a computerised screening in which the daily values for the entire month were printed out if a daily value in that month exceeded 200 mm. This screening yielded approximately 3 500 such high values. There followed a careful and time consuming checking procedure in which each of these events of >200 mm per day was checked

in the following manner. The daily rainfall data of surrounding stations were investigated. Extensive use was made of extreme daily rainfall information published by Adamson (1981) in order to gain some idea of the magnitudes of extreme values which were possible in the vicinity of the station. In addition the concurrent daily rainfall at nearby stations was investigated. Based on these two indicators a decision was made whether to accept or reject a value >200 mm. As a result of this exercise it was decided to accept 1 300 of these values as valid data. For the purpose of this project the remaining values of >200 mm per day were considered to be in error and were flagged as "missing" data. The file which contains these suspect data has been retained on disc as it forms a most useful reference, whenever daily data are extracted from the SAWB tapes or when extreme value analyses are performed.

Several checks were implemented on the monthly data, viz:

- (a) all months with more than 1 000 mm rainfall were flagged and investigated manually,
- (b) all months which had a coefficient of variation >80 per cent were flagged and investigated manually.

The third check on the monthly rainfall involved a check for zero rainfall in high rainfall months. Stations which met all the following three criteria were flagged as suspect and after inspection were not used in any of the analyses, viz. synthesis, regression or mapping:

- (a) monthly means for each of the three wettest months exceeded 75 mm and,
- (b) minimum period of record > 10 years and,
- (c) zero rainfall recorded on five or more occasions in the three wettest months.

The result of this analysis produced 77 stations for which the summer rainfall totals are highly suspect. The rationale behind the above check was that an error such as zero in a "wet" month was more damaging to the monthly and annual statistics than a zero in the dry months. The stations which failed this check were flagged on the computer file and should be treated with caution.

A problem which was revealed during the abovementioned checks was that of accumulated data which straddled a month end. The daily and monthly data for Jonkershoek-Jonkersnek and Jonkershoek-Virgin Peaks which are shown in Table 3.3 illustrate this problem. In these cases daily values are accumulated for six weeks with no rainfall credited to the four weeks of the first month; rather all the rainfall is credited to the second month. The six week accumulated total sometimes exceeded 1 000 mm and hence it was detected by the filter and classified as an error. These problems are of no consequence for annual totals provided the error does not occur over the year end.

The relationship between MAP and altitude was used as a check on the MAP of each station in relation to those at similar altitudes in the vicinity. A classified digital image of grid point altitudes and the map at each station was printed in colour. In this form of presentation clear trends of MAP and especially MAP as related to altitude could be distinguished. A careful visual check in search of anomalous values was then made of the MAP at each station. A system of colours, digits and alphabetic characters was developed to allow for easy identification of stations from the list of almost 7 500 with > 5 years of record.

Table 3.3 Daily and monthly rainfall data for Jonkersnek and Virgin Peaks

022030	2040	1141	1219	1972	1984	11	10013	JONKERSHOEK-JONKERSNEK					3056.2		
	JAN	FEB	MARCH	APRIL	MAY	JUNE	JULY	AUG	SEPT	OCT	NOV	DEC	ANNUAL		
1972	.0	148.0\$	107.2\$	56.8\$	662.3\$	.0	340.8\$	434.0\$	285.3\$	109.7\$	.0	41.4\$	2185.5		
1973	.0	57.6\$	98.6\$	.0	78.5\$	328.3\$	.0	942.9\$	.0	309.3\$	300.5\$	.0	2115.7		
1974	.0	50.6\$	49.8\$	53.7\$	.0	.0	724.7\$	329.0\$	999.90	.0	281.0\$	170.5\$	2659.2		
1975	.0	148.0\$	91.5\$	92.7\$	510.4\$	447.3\$	550.6\$	460.5\$	115.6\$	274.8\$	.0	147.4\$	2838.8		
1976	.0	34.5\$	127.6\$	.0	207.8\$	957.0\$	.0	.0M	270.0\$	.0	612.8\$	448.5\$	2658.2+		
1977	92.4\$	78.3\$	135.6\$	525.0\$	400.0\$	915.9\$	803.9\$	945.0\$	285.6\$	148.4\$	44.0\$	189.5\$	4563.6		
1978	242.4\$	60.8\$	50.8\$	295.8\$	338.5\$	92.0\$	178.3\$	917.2\$	397.4\$	315.7\$	174.5\$	.0	3063.4		
1979	354.0\$	89.2\$	132.1\$	117.5\$	290.8\$	917.6\$	254.0\$	194.9\$	180.1\$	447.7\$	128.4\$	12.8\$	3119.1		
1980	151.6\$	113.5\$	27.9\$	365.4\$	487.9\$	450.0\$	98.2\$	477.3\$	164.9\$	181.5\$	262.0\$	248.4\$	3028.6		

Monthly rainfall data 1972-1980 (Jonkersnek)

[illegible]

Daily rainfall data 1973-1974 (Jonkersnek)

Key

QUALITY CODES:

- | | |
|----|---|
| M | Missing data (-99.9M) |
| \$ | Accumulated data (total reliable) |
| * | Accumulated data (total unreliable) |
| # | Estimated data contained in the record (total unreliable) |
| O | Monthly total exceeds 999.9 mm |
| & | (INDEX) More than 5 months of zero rainfall during rainy seasons |
| Q | (INDEX) MAP inconsistent with neighbouring stations and topography |
| P | Synthesized value using the technique of Zucchini and Hiemstra (1984) |
| X | Value not synthesized but substitution by original month's mean |

Monthly rainfall data 1972-1980 (Virgin Peaks)

Daily rainfall data 1973-1974 (Virgin Peaks)

Key

```

M Missing data (-99.9M)
$ Accumulated data (total reliable)
* Accumulated data (total unreliable)
# Estimated data contained in the record (total unreliable)
O Monthly total exceeds 999.9 mm
& (INDEX) More than 5 months of zero rainfall during rainy seasons
Q (INDEX) MAP inconsistent with neighbouring stations and topography
P Synthesized value using the technique of Zucchini and Hiemstra (1984)
X Value not synthesized but substitution by original month's mean

```

The rainfall station location check described in Section 2.3 revealed that on a number of occasions the responsibility for the maintenance of stations had changed from SAWB to the Department of Agriculture and Water Supply or vice versa. In such cases the station data often appeared on the primary sources from both these organisations. Manual manipulation was necessary to compare the overlapping data in order to establish that it did in fact belong to the same station. Thereafter the data from the two stations were merged and the station number assigned to the most recent of the data agencies.

The final check on the MAP came in the form of the residual analysis which followed the regression of MAP on physiographic variables, stations which exhibited high residuals were examined manually (Chapter 5).

3.4 Managing the rainfall data

A system of computer programs and files was developed in order to manage these rainfall data. The system consisted of two main data files, namely an index file (INDEX) and the monthly data random access file (DATABASE). The latter file contains all the monthly and annual data as well as selected attributes of each station and a statistical summary of each station's data. The file DATABASE is over half a million, 80 character lines, in length. An example of the format of one station's data on this file is presented in Table 3.4.

The INDEX file contains only selected attributes of each station. These attributes are the station name, geographic position expressed as latitude and longitude in minutes of a degree, altitude in metres, number of years of complete data, the first and last years of data, and the record number in the DATABASE. An example of the format of this index file is presented in Table 3.5.

Table 3.4 Format of the monthly rainfall data for one station

QUALITY CODES:

M Missing data (-99.9M)
 \$ Accumulated data (total reliable)
 * Accumulated data (total unreliable)
 # Estimated data contained in the record (total unreliable)
 O Monthly total exceeds 999.9 mm
 & (INDEX) More than 5 months of zero rainfall during rainy seasons
 Q (INDEX) MAP inconsistent with neighbouring stations and topography
 P Synthesized value using the technique of Zucchini and Hiemstra (1984)
 X Value not synthesized but substitution by original month's mean

PUB0861	2030	1145	297	1958	1976	16	338396	WEMMERSHOEKDAM										823.9
19	LINES OF DATA			PUB0861				YEARS OF GOOD DATA: 16										
1958	14.7	26.2	13.2	19.1	94.7	55.1	12.2	129.5	17.0	22.1	40.4		.0	444.2				
1959	21.1	6.9	18.5	81.0	297.7	28.2	14.7	88.6	29.7	26.4	.3	14.5	627.6					
1960	7.1	7.4	39.4	42.9	107.4	150.9	30.7	43.2	51.6	27.7	2.8	13.0	524.1					
1961	46.7	27.4	22.6	35.1	83.1	154.4	63.8	101.9	129.8	20.6	.0	16.0	701.4					
1962	10.7	24.6	34.8	80.5	32.8	400.6	102.9	247.4	77.2	136.1	17.8	5.6	1171.0					
1963	26.2	.0	3.0	21.1	24.1	68.8	191.5	216.7	42.9	9.4	49.3	45.5	698.5					
1964	.5	101.9	5.8	17.8	59.2	197.6	112.0	165.1	23.4	77.7	55.9	11.2	828.1					
1965	29.5	62.2	92.5	95.3	114.6	91.2	64.0	166.1	46.2	35.1	11.9	-99.9M	808.6+					
1966	2.0	7.4	137.9	50.5	48.0	134.9	193.5	71.9	65.8	1.8	11.9	19.8	745.4					
1967	22.6	.8	3.3	133.4	103.1	257.6	110.7	78.2	98.3	163.1	21.1	48.3	1040.5					
1968	33.5	15.2	23.6	64.0	8.1	93.0	90.7	159.3	133.4	100.6	10.2	4.1	735.7					
1969	5.1	10.2	5.1	4.6	186.7	239.0	164.3	152.4	90.4	44.2	29.5	40.9	972.4					
1970	5.1	1.5	36.6	3.3	87.4	127.3	195.3	191.3	19.1	22.9	7.4	53.3	750.5					
1971	31.5	16.0	17.5	83.6	155.2	91.9	67.8	111.5	69.3	21.6	.0	46.5	712.4					
1972	2.6	1.0	35.3	12.4	49.8	32.0	39.7	80.8	98.0	52.8	10.7	92.7	507.8					
1973	6.4	1.3	21.6	1.1	137.2	255.8	117.3	544.8	89.2	111.5	24.4	12.0	1322.6					
1974	26.7	23.4	6.7	90.7	260.4	131.8	218.9	112.3	18.5	85.1	-99.9M	5.6	980.1+					
1975	.0	1.3	23.4	26.9	115.1	406.1	325.4	62.0	71.6	17.3	202.2	148.8	1400.1					
1976	33.3	49.3	23.1	133.6	286.8	365.0	-99.9M	-99.9M	61.0	-99.9M	-99.9M	24.9	977.0+					
MEAN	17.1	20.2	29.7	52.5	118.5	172.7	117.5	151.3	64.9	54.2	29.2	33.5	823.9					
MEDN	14.7	10.2	22.6	42.9	103.1	134.9	106.8	120.9	65.8	31.4	11.9	17.9	740.5					
STD	14.0	26.1	33.1	42.4	85.6	118.6	83.1	112.9	35.8	47.1	47.7	37.3	283.5					
C.V.	82.0	129.2	111.6	80.9	72.2	68.7	70.7	74.6	55.2	86.9	163.6	111.3	34.4					
SKEW	.4	1.9	2.2	.5	.9	.8	.8	2.5	.3	1.0	3.0	1.9	.7					
YRS	19	19	19	19	19	19	18	18	19	18	17	18	16					
				PERIOD OF RECORD				1958 MONTH 1 TO 1976 MONTH 12										

Table 3.5 Format of file INDEX

Station number	Latitude (mins.)	Longitude (mins.)	Altitude (metres)	Start year	End year	Number of years of record	Record no. in DATABASE	Station name	Latitude (degrees)	Longitude (degrees)
001517	2078	1158	14	1905	1984	73	1	DANGER POINT (VRT)	3437.8	1918.0
001605	2075	1161	17	1926	1984	47	84	GANSBAAI	3435.0	1921.0
001726	2076	1165	20	1900	1984	32	142	UILENKRAAL (BOS)	3436.0	1925.0
001813	2073	1168	152	1950	1957	6	297581	LOEFONTEIN	3433.0	1928.0
002069	2079	1172	15	1944	1982	37	10793	PETERS GATE	3439.3	1932.0
002256	2086	1179	12	1900	1976	42	10838	QUOIN POINT (BOS)	3446.0	1939.1
002456	2076	1186	60	1937	1984	23	10888	ELIM (POL)	3435.5	1945.5
002639	2079	1192	15	1916	1984	37	10923	VOELVLEI	3438.5	1952.1

The INDEX and DATABASE files and the search and extraction programs described in this section have been placed on the CCWR mainframe and are available to all users of the CCWR. The description which follows is therefore phrased in terms of "the user". However, the research staff on this project used the system extensively and in the same way. Ideally, researchers should be able to access the data base and extract data themselves. The specifications for data retrieval are contained in a clearly stated menu and the extraction operation is conducted with ease and speed. To achieve this, it was necessary to examine the manner in which most data are requested.

- (a) Invariably, the study area is defined. The most important attribute of the search and extraction technique therefore was that it should accommodate requests for data from a specific area.

- (b) Generally, the minimum length of record at the acceptable stations was also specified, along with
- (c) the type of data required such as monthly or annual totals, or a statistical summary of these values.
- (d) Finally, the output medium such as disc file or printout was specified. A schematic block diagram of the logic flow path followed by the data search and extraction program is shown in Figure 3.3.

The search and extraction techniques have been designed to fulfil the abovementioned criteria. Additional features include extraction of key attributes of rainfall stations and depiction of the spatial distribution of such stations. Selected statistics of monthly and annual rainfall also have been included in the data set and require extra disc space. However, the increase in speed (since they are not calculated each time they are requested) compensates for the extra disc space.

Hydrologic study areas are generally defined by drainage basin boundaries or regions of similar rainfall characteristics. The rainfall statistics mapping project required some form of regionalization of the rainfall data. The Welding and Havenga (1974) rainfall regions for South Africa were selected initially. These regions are discussed in Chapter 7. The boundaries of the Welding and Havenga (1974) regions were digitized and added to the files; stations may therefore be accessed by region. These regional boundaries may change with future research. It was therefore essential that the allocation of stations to regions be conducted automatically by a computer program. The routines that complete the regionalization procedure are time consuming; therefore, this is the last

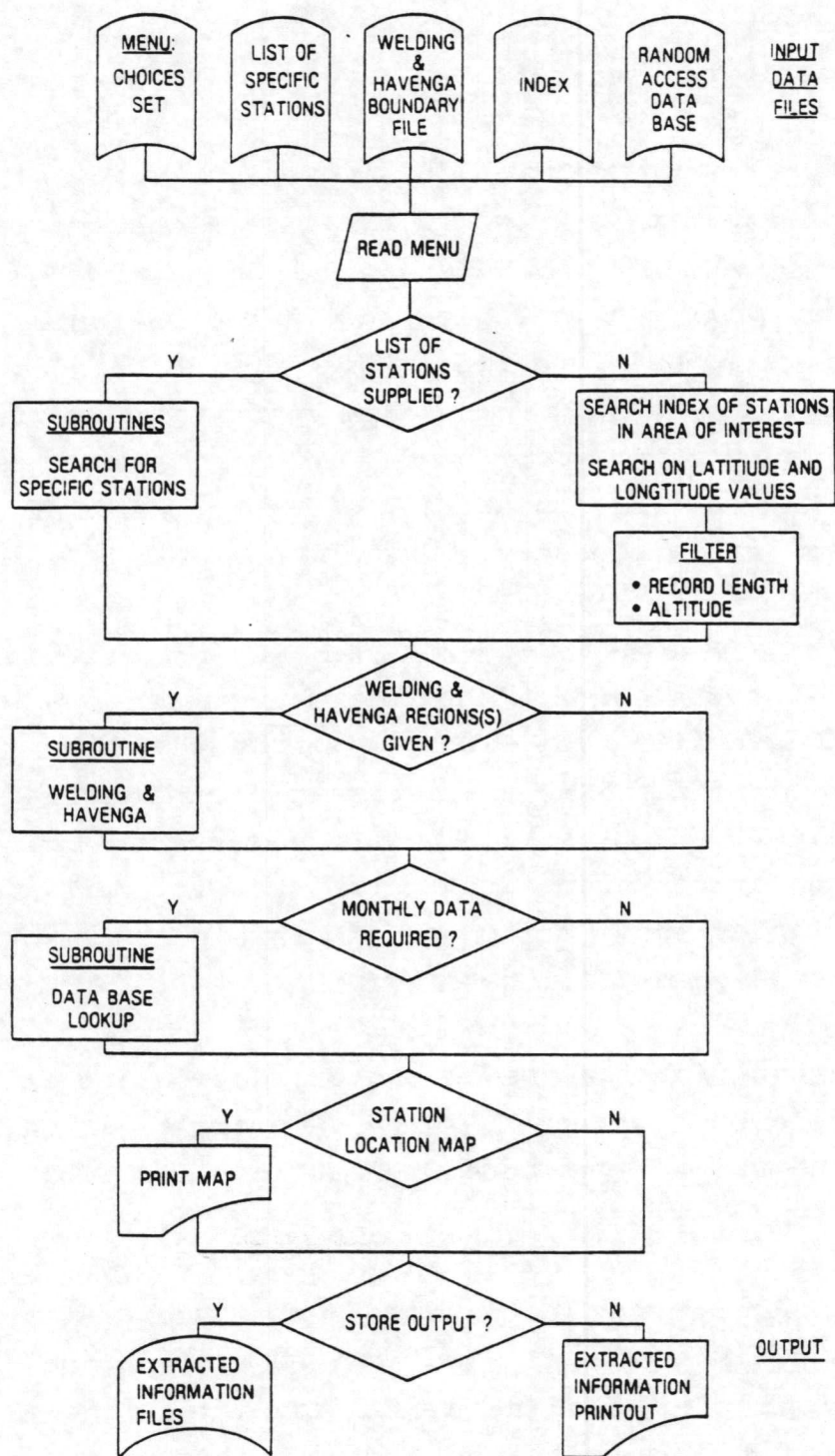


Figure 3.3 Logic path of the data search and extraction program

of the search criteria, since the least number of stations remains in the selected INDEX file by this stage.

An additional feature of the regionalization routine is that in order to extract rainfall stations for a particular drainage basin, the Welding and Havenga (1974) regional boundary file can be copied and the co-ordinates of any one of the regional boundaries replaced with the drainage basin boundary co-ordinates. This particular region is then specified in the search and retrieve request presented in Table 3.6 in order to extract stations that fall within the drainage basin boundary.

All searches are first executed on the INDEX file, thereafter the DATABASE record numbers are read from the selected subfile from INDEX. The speed of data extraction from the DATABASE is enhanced considerably by this mechanism. Data selection and extraction based on these file structures and search techniques is approximately five times faster than performing the same processes by means of computer system editing software.

The output from the search and extraction program may be in disc file or a printout form. The types of data that may be requested are presented in Section 7 of Table 3.6.

The user may request a printout depicting the spatial distribution of the selected stations. Such a feature is often useful in regional studies. The operation of the spatial depiction routine is illustrated schematically in Figure 3.4. The minimum latitude and longitude of the study area block is subtracted from the latitude and longitude of each station. The converted co-ordinates then form the row and column position of the station symbol in the array DEPICT. A symbol classification routine is executed on the points before they enter the array DEPICT.

3 - 19

5. WELDING AND HAVENGA REGIONS

- . maximum of 5
- . arrange in ascending alphanumeric order
- . character format

*** '6A1b' '6B1a' '20A2b' ' '

6. FORM OF SEARCH AND EXTRACTION

- search and report INDEX only (YES/NO)

*** 'NO'

- search INDEX then extract and report rainfall data (YES/NO)

- *** 'YES'

7. RAINFALL DATA TYPE

- (a) annual rainfall totals only

- (b) annual rainfall statistics only

- (c) monthly rainfall only

- (d) monthly and annual statistics only

- (e) all rainfall data and statistics

(a)	(b)	(c)	(d)	(e)
-----	-----	-----	-----	-----

- *** 'NO' 'NO' 'NO' 'NO' 'YES'

8. OUTPUT MEDIUM

6 = printer F = file called EXTRACT

*** '6'

9. SPATIAL DEPICTION OF STATIONS

- spatial depiction (YES/NO)**

*** 'YES'

- grid symbol station symbol boundary symbol

• • •

Stations in the selected file may be arranged in any sequence. The process is repeated until all the stations have been allocated a position in the array DEPICT and the array DEPICT is then printed. The classification routine enables a distinction to be made between station points, boundary points, and grid points. The symbols for these points are specified in the input request file (MENU). It is also possible to link this spatial depiction routine to a classified altitudinal image. This allows the user to view the stations in their topographic as well as spatial context. At present the topographic routine is programmed to depict at a resolution of 5 minutes of a degree; however, a 1 minute of a degree topographic depiction is possible for limited areas, given the computer time and disc space.

An analysis of each element of this monthly rainfall data management system shows that there is nothing particularly difficult about the system. The significant point is that these data sets and the search and extraction system have been completed. Hydrologists and engineers may now proceed with the task of research and design using monthly rainfall data without the burden of data collation, reformation and extraction.

A particular advantage of this data extraction system is that the specifications for extracting data are formulated in the manner in which a request is usually phrased by a user. This advantage, combined with the fact that station position oriented searching routines form a key element of the extraction process, has made the system user friendly.

The following is a very useful consequence of this search and extraction system. The daily data from which the monthly data set was compiled, are too voluminous to reside on disc files. Therefore, there is often a need to review the station locations and the quantity and quality of data

at the monthly level from stations within the study area before initiating the time consuming and costly process of extracting daily data.

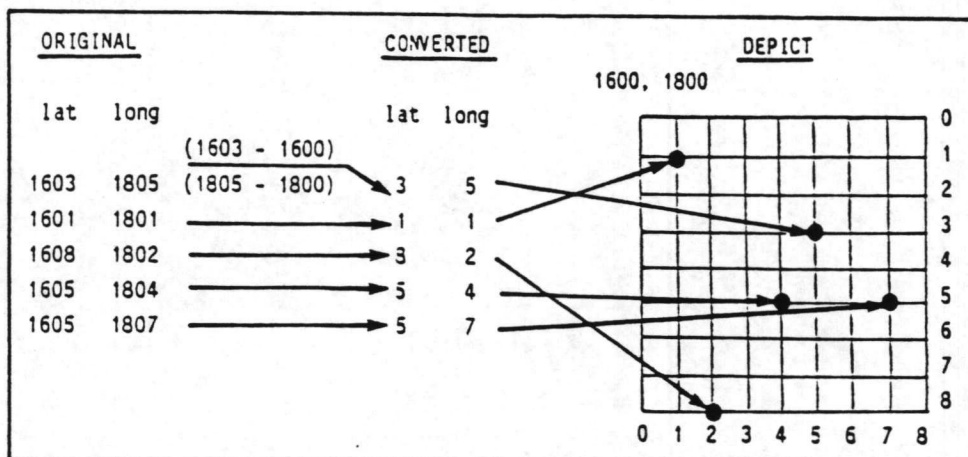


Figure 3.4 Schematic illustration of the spatial depiction routine

The problem of updating data sets is often as great as creating the data sets originally. Ninety per cent of the daily data which form the basis of this system are available on magnetic tape. The monthly data set has been updated once already and a two yearly update of the data base is planned.

The data base management system was devised in order to fulfil the requirements of the rainfall statistics mapping project. However, the system's ability to extract data products of monthly and annual rainfall statistics, as well as to depict the location of such rainfall stations, has resulted in a number of requests for these data products from researchers throughout Southern Africa.

CHAPTER 4

SYNTHESISING MONTHLY RAINFALL

The reliability of the estimates of the mean and variance of a population, increase with the length of the sequence. In this project it was desirable to maximise, within limits, the reliability of the estimates of the rainfall statistics being mapped. The annual rainfall sequences were too short, in many cases, to achieve the desired reliability. It was therefore necessary to improve the reliability of the estimated statistics by some synthetic means.

Matalas and Jacobs (1964) showed that if the product-moment correlation coefficient between rainfall records at two stations exceeds $1/\sqrt{n_1-2}$, where n_1 is the length of the concurrent period, then the reliability of the estimate of the mean of the record, extended by linear regression, is greater than that based only on the observations of the short sequence. Similar conclusions were revealed by Cochran (1953) as cited by Matalas and Jacobs (1964). Rosenblatt (1959) also cited by Matalas and Jacobs (1964), however showed that a rainfall sequence, which was lengthened by linear regression yielded a biased estimate of the variance. Fiering (1963) therefore used multiple linear regression in order to extend a short sequence by employing more than one long sequence and obtained similar results to the abovementioned with respect to the mean and variance.

In extending the rainfall record it was important to preserve both the mean and the variance for two reasons, viz:

- (a) firstly, in order to map reliable estimates of the statistics and
- (b) secondly to ensure that the synthetic rainfall could be used with confidence in rainfall/runoff models for water resources assessment.

Matalas and Jacobs (1964) added a random or noise component to the regression estimate. The addition of this noise was shown to have no effect on the reliability of the estimate of the mean. The reliability of the estimate of the variance was, however, improved and no bias was introduced by the addition of noise. Since they are based on random noise, the rainfall records synthesised by this technique are not suitable for use in rainfall/runoff models when such models are used to compare the observed and estimated runoff.

It is not only essential to have a good estimate of the population statistic, for example of the MAP, at the rainfall station but it is also desirable to have a comprehensive spatial and representative altitudinal distribution of stations with an acceptable record length. It was anticipated that increasing the length of the rainfall record at each station would improve the spatial distribution of rainfall stations which meet any record length criterion for inclusion into the data set. This point is aptly illustrated in Figure 4.1, which depicts the location of stations which would qualify for inclusion, should 10, 20 or 40 years' minimum duration limits be set. In addition to the improved two dimensional spatial distribution shown in Figure 4.1 the altitudinal representation should also be considered. Figure 4.2 depicts the cumulative frequency distribution of the rainfall station altitudes and the grid point altitudes for the area shown in Figure 4.1. The dearth of stations in the higher altitudes of, for example the south western Cape region, is noticeable and is of

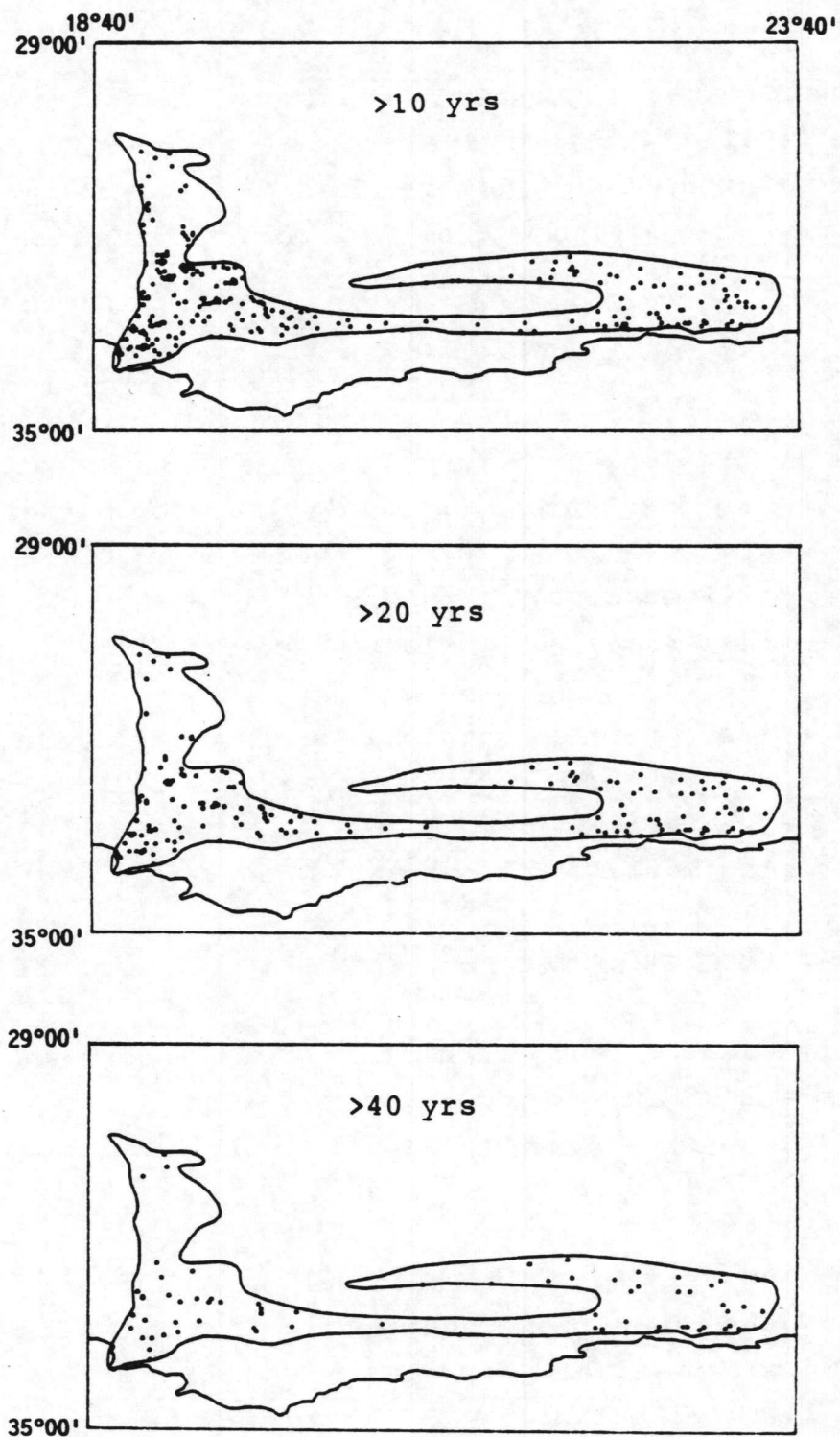


Figure 4.1 Spatial distribution of rainfall stations in the south western Cape with record lengths of more than 10, 20 and 40 years

concern when regression techniques to estimate MAP using inter alia altitude are applied, particularly where the regression equations are used to estimate precipitation beyond the altitude range of rainfall stations for which the equations were developed in a region.

The rainfall record may be lengthened beyond and prior to the period of active recording at the station and gaps in the record filled in to produce the same effect on the distribution of stations which meet a record length criterion as would the lowering of the record length criterion itself. A number of interesting issues arising out of this are addressed in the following sections, inter alia:

- (a) the period and desired length of the rainfall record,
- (b) the method of synthesising rainfall to fill gaps in the rainfall record and
- (c) the method of lengthening the rainfall record.

4.1 Period and desired length of the rainfall record

The choice of the period of rainfall record and its length is not a straight forward matter. There were three aspects to this question which required detailed consideration, viz:

- (a) the length of the base period,
- (b) the start and end years of the base period, or alternatively,
- (c) the acceptability of having no fixed period as long as the span of the record was long enough.

Whilst reviewing these questions it is important to reflect on the following extracts from Dyer and Tyson (1977), the World Meteorological Organisation (1974) and Dunne and Leopold (1978). Dyer and Tyson (1977) state that, "No

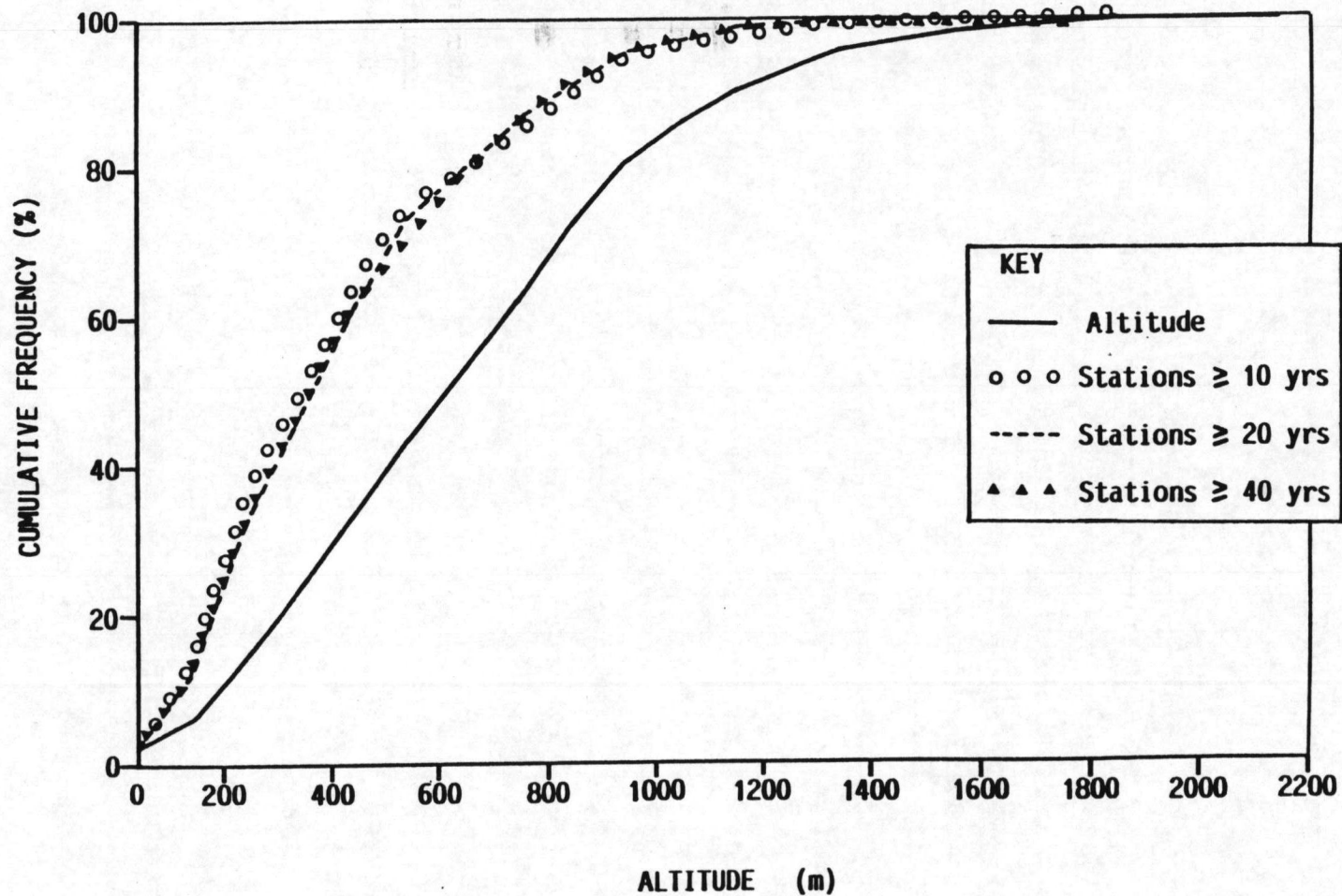


Figure 4.2 Cumulative frequency distributions of rainfall station altitude and grid point altitude for rainfall stations in the south western Cape with record lengths of more than 10, 20, and 40 years

evidence can be found to show that South African rainfall has increased or decreased progressively throughout the period of meteorological record. Instead quasi-periodic fluctuations appear to have persisted throughout the period" (p 145). This may be interpreted as implying that it does not matter when one takes one's period of record so long as it is of sufficient duration to have spanned enough of the "quasi-periodic fluctuations" that their influence does not affect critical statistics. The World Meteorological Organisation (1974) states that the choice of the optimum base period for MAP estimates is "a period long enough to represent a good sample of record through time. Too short a period may be unduly influenced by a particularly dry or rainy period. Too long a period will require too much synthesis of record" (p 5.3) and will most probably result in a poor network of rainfall stations.

Dunne and Leopold (1978) make a significant contribution to the discussion on the base period: "By international agreement, the average climate of the world used to be defined as the average value of precipitation, temperature and other variables for the period 1931 to 1960. As the period of observation was extended forward and backward in time, it became apparent that the period 1931 to 1960 had one of the most abnormal climatic regimes in the last several hundred years. The average climate of a place is now generally updated each decade to another 30 year period, 1941 to 1970, 1951 to 1980, and so on. Many published records and atlas maps, however, are based upon data from 1931 to 1960 and they should be checked against older and more recent information wherever possible" (p 71).

On the existing 1:250 000 isohyetal maps of Southern Africa it is stated that the MAP is for the period 1921-1960. A study of the rainfall record length and period for stations in several regions revealed that the records at these sta-

tions must have been "normalised" in order to allow the station to be used for the isohyetal mapping exercise, bearing in mind the 40 year base period from 1921 to 1960. Such a normalisation technique generally involved the determination of the ratio between the annual rainfall at the station to be normalised and a long term station (covering the base period). This ratio was then used to estimate annual totals for the "missing" years at the station with the short record. Such a technique, which is a rudimentary form of linear regression, has several shortcomings, notably that the variance of the synthetic annual values is considerably reduced as stated by Matalas and Jacobs (1964) and Zucchini and Hiemstra (1984). The variance of the data thus impairs the use of these data in regression analyses where the variance of the annual totals is used in the weighting of the input variable MAP. In addition the usefulness of one of the most important by-products of this study, namely the synthetic monthly rainfall data, would have been reduced, had the normalising technique been applied, since the variance of the monthly rainfall record is vital if used, for example in any rainfall/runoff modelling for the purposes of assessing water resources.

The several preliminary studies in this project had revealed the paucity of rainfall stations with 30 or more years of record and thus made it imperative to review the concept of a fixed long term base period. The objective of this project was to estimate various precipitation statistics such as means and variabilities at all points in the study area. It was decided that the best possible estimate of the population statistic at a rainfall station was to be gained by using the data at that station, irrespective of the start and end years of the rainfall record period, provided of course that the record was sufficiently long.

It therefore remained to establish a minimum record length

of annual totals in order to calculate a reasonable estimate of the long term MAP. Furthermore it was necessary to determine whether this minimum record length varied from region to region. Knisel et al. (1979) suggest that the use of short term records can bias estimates of MAP significantly and that this bias is more pronounced in semi-arid regions. Large parts of South Africa are semi-arid. Numerous studies on South Africa's annual rainfall have shown the apparent persistence of high and low periods (for example Dyer, 1976; Hall, 1976; Dyer and Tyson, 1977; Dyer and Gosnell, 1978). Rawson (1908), cited in Dyer and Tyson (1977) and discussed by Louw (1982), reported a 19 year oscillation in the position of the Atlantic and Indian Ocean anti-cyclones for the period 1841 to 1906. Nevill (1908) found an 18 to 20 year oscillation to be present in Natal for the period 1850 to 1908. Hall (1976) analysed the growth rings of a tree felled near Pietermaritzburg in 1910 and suggests that a quasi 20 year oscillation was evident from the width of growth rings from 1910 back to 1750. Dyer and Gosnell (1978) found a significant long term oscillation with a mean wavelength of 19,2 years in the sugarcane belt of Natal. Dyer (1976) used the quasi 20 year oscillations in order to estimate expected future rainfall in selected parts of South Africa. Whitmore (1962b) reported "a significant degree of persistence" in the annual rainfall of South Africa. Louw (1982) found oscillations of 18 years and shorter in the South African annual rainfall record. Research conducted by Bromley (1980) in Zimbabwe on annual rainfall showed a 19 to 20 year periodicity in Harare, an 18 year periodicity in Gweru and a 27 year periodicity in Bulawayo.

The fact that the existing 1:250 000 scale isohyetal maps employed a 40 year base period combined with the apparent tendency for a periodicity around 20 years in the annual rainfall led to the decision, that initially, a minimum record length of 40 years be considered. However, it be-

came apparent during the study that parts of the country are not represented realistically with rainfall data, if only the stations with 40 years or more of record are considered. It was evident that some relaxation of this record length criterion was necessary in order to achieve an acceptable spatial distribution of stations. However, such a reduction was not to lead to a marked error in the estimate of the mean. It was felt that the resultant inclusion of vital stations would certainly lead to more realistic isohyetal estimates than if those stations had been omitted.

A first step in the proposed relaxation of record length was to determine the possible error, in the estimate of the long term mean, which would be introduced if shorter term means were to be used. The rainfall stations were grouped according to the 25 major rainfall regions of Southern Africa proposed by Welding and Havenga (1974) and shown in Figure 4.3. Thereafter all the stations in each region which had more than 40 years of record were selected. Moving "windows" of 5, 10, 15, 20, 25, 30, 35 and 40 years were passed through the data and the mean annual rainfall was calculated for each of these windows. The difference between this short term (window) mean and the long term (>40 year) mean was then expressed as a percentage of the long term mean. The 80 and 90 percentile limits of these errors were established for each of the windows, i.e. the percentage error, both positive and negative that may be expected to be exceeded 20 and 10 per cent of years, respectively, if a shorter record was to be used to estimate the long term mean. A table similar to that shown in Table 4.1 was compiled for each of the 25 major Welding and Havenga (1974) regions. These tables, which are presented in Appendix A, were consulted closely when decisions were taken as to include or exclude stations with short records from the data set for each region in the

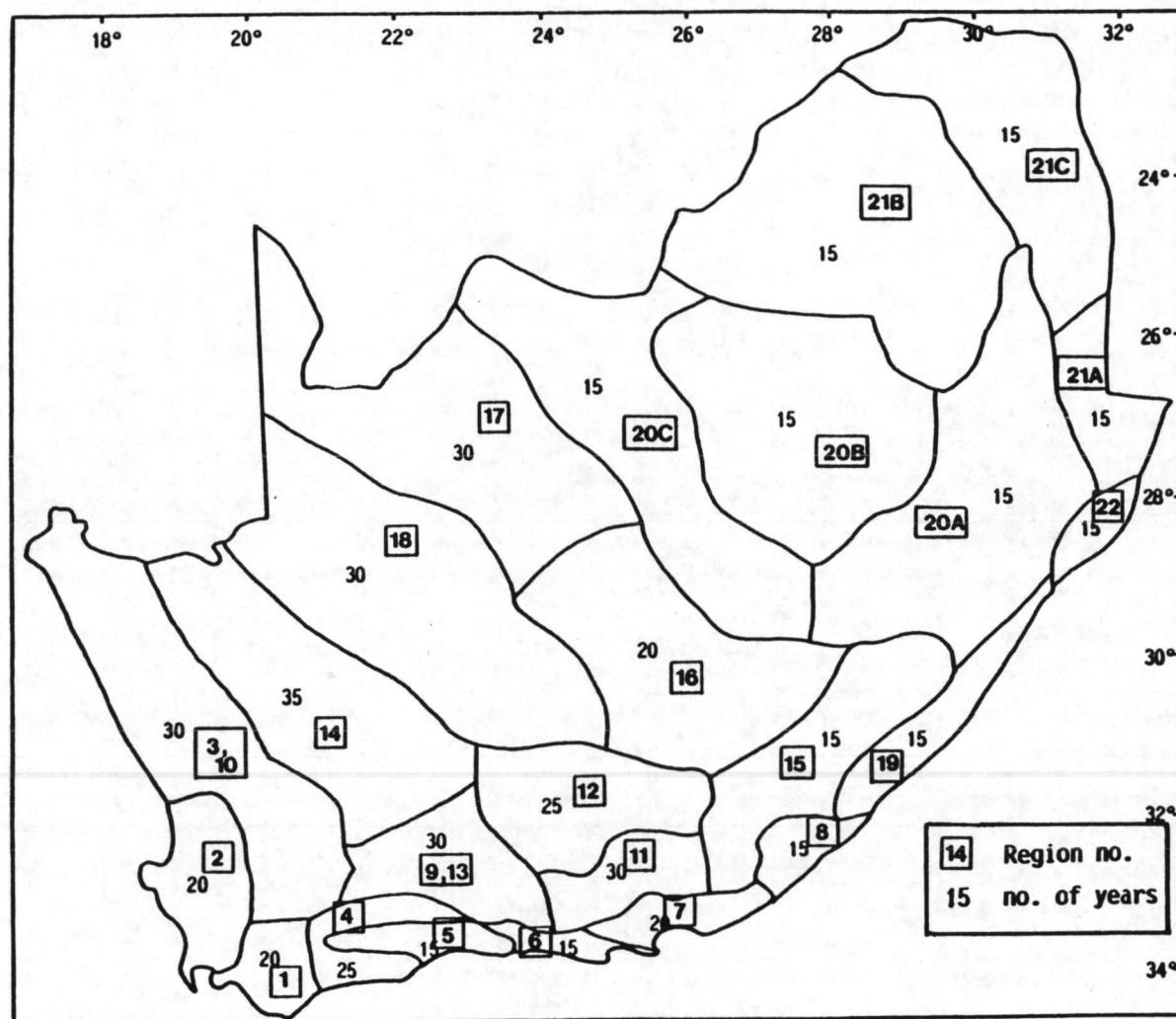


Figure 4.3 Record lengths required to ensure that the mean of annual rainfall estimates to within 10 per cent of the long term mean, 90 per cent of the time

Table 4.1 Upper limits of the percentage error of the short term mean: An example of the error analysis for the Welding and Havenga region 20B

WELDING & HAVENGA REGIONS				20B	POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 412 STATIONS											
				JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN	AT	90% CONFD.		36.5	39.5	42.5	59.5	74.5	100.5	95.5	100.5	88.5	42.5	41.5	36.5	17.5
		80% CONFD.		27.5	30.5	32.5	45.5	59.5	79.5	79.5	76.5	71.5	33.5	31.5	27.5	13.5
		SAMPLE SIZE		24262	24247	24234	24096	24026	23974	23938	23972	24023	24097	24091	24115	23812
10 YEAR MEAN	AT	90% CONFD.		24.5	25.5	28.5	46.5	51.5	80.5	75.5	80.5	58.5	29.5	27.5	24.5	12.5
		80% CONFD.		18.5	19.5	22.5	35.5	39.5	64.5	58.5	61.5	46.5	23.5	20.5	18.5	9.5
		SAMPLE SIZE		22202	22187	22174	22036	21966	21914	21878	21912	21963	22037	22031	22055	21752
15 YEAR MEAN	AT	90% CONFD.		19.5	19.5	21.5	39.5	39.5	63.5	59.5	63.5	38.5	23.5	20.5	19.5	10.5
		80% CONFD.		14.5	15.5	17.5	31.5	30.5	49.5	45.5	50.5	29.5	18.5	15.5	14.5	7.5
		SAMPLE SIZE		20142	20127	20114	19976	19906	19854	19818	19852	19903	19977	19971	19995	19692
20 YEAR MEAN	AT	90% CONFD.		16.5	16.5	17.5	35.5	34.5	51.5	50.5	51.5	31.5	19.5	16.5	15.5	8.5
		80% CONFD.		12.5	12.5	13.5	27.5	27.5	38.5	38.5	42.5	24.5	14.5	12.5	11.5	6.5
		SAMPLE SIZE		18082	18067	18054	17916	17846	17794	17758	17792	17843	17917	17911	17935	17632
25 YEAR MEAN	AT	90% CONFD.		13.5	14.5	14.5	30.5	30.5	47.5	40.5	45.5	28.5	16.5	14.5	13.5	7.5
		80% CONFD.		10.5	10.5	11.5	24.5	23.5	34.5	31.5	36.5	21.5	12.5	10.5	9.5	5.5
		SAMPLE SIZE		16022	16007	15994	15856	15786	15734	15698	15732	15783	15857	15851	15875	15572
30 YEAR MEAN	AT	90% CONFD.		11.5	11.5	13.5	27.5	27.5	41.5	35.5	41.5	23.5	13.5	12.5	11.5	6.5
		80% CONFD.		9.5	8.5	10.5	21.5	21.5	31.5	28.5	33.5	17.5	10.5	9.5	8.5	4.5
		SAMPLE SIZE		13962	13947	13934	13796	13726	13674	13638	13672	13723	13797	13791	13815	13512
35 YEAR MEAN	AT	90% CONFD.		10.5	10.5	11.5	24.5	24.5	36.5	32.5	36.5	20.5	11.5	10.5	9.5	5.5
		80% CONFD.		7.5	7.5	8.5	19.5	19.5	27.5	25.5	29.5	14.5	8.5	7.5	7.5	3.5
		SAMPLE SIZE		11902	11887	11874	11736	11666	11614	11578	11612	11663	11737	11731	11755	11452
40 YEAR MEAN	AT	90% CONFD.		9.5	8.5	10.5	21.5	22.5	31.5	28.5	33.5	19.5	10.5	8.5	8.5	4.5
		80% CONFD.		7.5	6.5	7.5	17.5	17.5	23.5	21.5	27.5	14.5	7.5	6.5	6.5	3.5
		SAMPLE SIZE		9842	9827	9814	9676	9606	9554	9518	9552	9603	9677	9671	9695	9392

regression analysis, of MAP on physiographic variables, which preceded isohyetal mapping.

It is evident from Figure 4.3 that the length of record required varies from region to region. As may be expected, the length of record required in the more arid regions of Southern Africa is longer than that required in the wetter regions. However, it should be noted that the expression of the possible error in per cent, may be somewhat misleading. For example when consulting Table 4.1 the percentage possible error at the 90 per cent confidence level of 13 per cent for the 10 year mean may seem high. However, this possible error should be seen in the light of the improvement in the estimate of the mean that can be gained by synthesising in order to attain a 40 year record. The improvement would only be 9 per cent. The corresponding reduction in the spatial and altitudinal coverage of stations would, on the other hand, be drastic, in most cases. The implications of these two facts were highly influential in the decision to reduce the required length of record required for the station means, where necessary.

The use of percentages to express the "error in the estimate" has a disadvantage in that a 10 per cent error in MAP in the Karoo may only be equivalent to 20 mm whilst a 10 per cent error in the Drakensberg may be 150 mm. The only "solution" to this dilemma is to bring the abovementioned points to the attention of the reader so that the results of these analyses are interpreted correctly. In some of the drier months the percentage errors show anomalous trends with respect to the record lengths. This "apparent error" may be attributed to the high variability of rainfall in the dry months.

After completion of the abovementioned studies and having gained experience in mapping the statistics of rainfall it became apparent that the question of the minimum length of

record was one which cannot be answered unequivocally. Generally the objective was to use an estimate which, at the 90 per cent confidence limits, yielded a probable error of less than 10 per cent.

4.2 Techniques for synthesising monthly rainfall

It has been shown in the previous section that it is desirable to obtain as many stations as possible with sufficiently long records for mapping MAP, and in order to do this it is acceptable to synthesise rainfall data. There are several different techniques which are suitable for synthesising monthly rainfall data. Considerable time was dedicated to investigating these techniques.

4.2.1 Background and concepts of synthesising

On inspection of the monthly data sets which made up the annual totals it became evident that many years were coded as containing missing data, and therefore not suitable for use in computing the relevant statistics, when only one or perhaps two months data were missing or incomplete. A single missing day is enough to establish a missing data code on the month and hence the year. A missing data flag on the year meant that the whole year of data was then discarded for the sake of perhaps one missing day. An analysis was conducted on the DATABASE to establish the extent of this problem. In the analysis the missing months and years were expressed as a percentage of the total number of months or years of complete data. The results of this analysis are summarised in Table 4.2.

It was evident from these results that the data synthesis would have to take place at monthly level rather than annual, implying a 12-fold increase in the computer time and also additional program development to collate the monthly files back into a single annual file once the

Table 4.2 Percentage of missing months and years in the
rainfall data set

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
% MISSING	6.47	7.12	6.53	6.95	6.54	6.74	6.64	6.59	6.88	6.51	7.38	7.27	23.79

analysis had been completed.

A further measure which involved the replacement of a selected but limited number of missing months by the mean for that month in order to complete the year, further increased the number of complete years without impairing the statistics of MAP or standard deviation of annual totals appreciably. This technique and the results emanating from its application are discussed in Section 4.7.

The large number of stations to be processed in this study necessitated a high degree of computational efficiency in terms of speed, automation and pre-programmed decision making. Furthermore the methods were required to be robust with respect to possible non-optimal parameter selection as a result of this automation.

It should be emphasised that all the techniques which were considered (Sections 4.2 and 4.3), require rainfall record generally of at least 15 complete years to exist at a station whose record was to be augmented. The synthesised values were also required to be repeatable and be well correlated with observed data at stations in the vicinity.

Linear regression models using one or more control stations have been used extensively to synthesise missing data (Zucchini and Sparks, 1984a). These models generally ful-

fill the requirements discussed above. Optimal interpolation methods are also linear models but differ from least squares techniques in that the correlations between target and control stations are used to calculate the relative weightings of stations used to synthesise missing data. These techniques are perhaps better known as methods for spatial interpolation but have features which make them suitable for data synthesis.

These three "families" of methods, viz.;

- (a) methods requiring a single control station,
- (b) methods requiring more than one control station, and
- (c) optimal interpolation techniques

will be discussed and the two methods which performed the best in tests will be compared in detail.

4.2.2 Selection of target:control station pairs

Throughout the description and assessment of the synthesising techniques presented by Zucchini and Hiemstra (1984) the terms "target" and "control" station are used to indicate which stations provide the dependent and the independent variables respectively. A short description of how the target and control station pairs were selected is appropriate at this stage. This selection process was automated out of necessity and in accordance with the objectives of the project.

Each station in the rainfall data set was investigated with a view to creating one or more pairs of target and control stations for each target station. The selection process dictated that the control station should meet the following criteria with respect to the target station, namely, that it should:

- (a) have at least 15 years' concurrent record,
- (b) be one of the five nearest stations which meets the above criterion, and
- (c) have a correlation coefficient of more than 0,5 with the target station at the level of monthly data.

The control stations which met the above criteria with respect to the target station were then subjected to a scoring system in order to select the final target:control station pair to be used in the synthesising procedure. The scoring system was designed to prevent decisions as to the "best" target:control pair being taken purely on the basis of correlation coefficient. Such a system could have led to possible anomalous pairings, as the following example illustrates:

Consider the case of two possible control stations with correlation coefficients of 0,82 and 0,83 with the target station, the distances between the target station and each of the control stations being 5 and 50 km respectively. On the basis of correlation coefficient alone, the latter control station would have been paired with the target station. However, this is not necessarily the best pairing when one considers the distance criterion as well. Similarly the total length of concurrent record may have been 30 years and 15 years respectively. To reiterate, the case is made not simply to take correlation coefficient as the final criterion but to look at the temporal and spatial relationships between the stations.

It became evident that a scoring system to weight the above criteria was necessary in order to produce, eventually, one number upon which the decision regarding target:control station pairs could be based. The scoring system described below was tested in order to establish which weighting produced the "best" station pair selection.

The distances and record lengths were standardised to a maximum value of unity as follows

- (a) interstation distance index = $(100 - d)/100$, and
d = interstation distance in km, assuming a maximum scan for stations of 100 km radius, and
- (b) length of record index = length of concurrent record/50, with an upper limit of unity.

Initially the scoring system weighted the correlation coefficient, interstation distance index and the length of concurrent record index in the ratio 7:2:1 respectively. The best five target:control station pairs were selected using this ratio. Thereafter some of the observed data at the target station were "hidden" i.e. did not take part in establishing the interstation model parameters and the control station data were used with a technique described in Section 4.2.3 as the Zucchini and Hiemstra (1984) model 5, to synthesise this hidden target station data, using each control station in turn. The control stations were then ranked according to the quality of the synthesising which they produced at the target station. It remained then to establish which of the weighting ratios produced the most reliable estimate of the abovementioned rank order.

Permutations of weighting involving assigned weights between 1 and 10 to each of the three scoring criteria were used to establish which weighting system predicted that best control station to partner the target station. After extensive tests on 137 stations the weighting 6:3:1 was selected.

4.2.3 Synthesising models reported by Zucchini and Hiemstra (1984)

Zucchini and Hiemstra (1984) consider the following criteria to be important in the choice of a model to synthesise monthly streamflow.

- (a) The method should be "transparent" in the sense that the influence of each parameter is evident.
- (b) The results of the estimation should be repeatable.
- (c) The method should not introduce systematic bias into the statistics of the extended series.
- (d) The method should allow some movement in the statistics in order to account for short term changes in the temporal distribution of rainfall.

The six models reported by Zucchini and Hiemstra (1984) were used by them to estimate streamflow. However, the techniques may also be used for synthesising monthly rainfall (Zucchini and Sparks, 1984b). Matalas and Jacobs (1964) used Model 3 for synthesising rainfall.

All the selected target/control station pairs which contained more than 30 years' concurrent data were used to assess the performance of the models reported by Zucchini and Hiemstra¹ (1984). A "window" of 15 years' concurrent record was moved through the target and control data. For each setting of the window the model parameters were established and then the synthesis was conducted. The estimated and observed monthly totals were then compared. The 15 year window then moved on by one year and the "hidden" data "wraps" around to include the data "behind" the window and the assessment process was repeated. This system is illustrated schematically in Figure 4.4.

¹ These models will hereafter be referred to as the Zucchini and Hiemstra models for the sake of brevity.

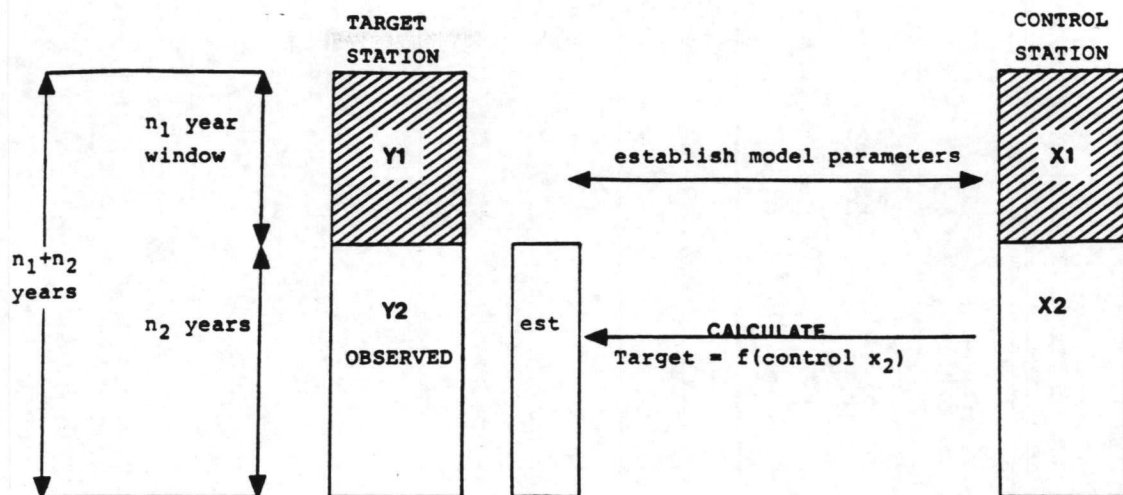


Figure 4.4 Schematic presentation of the procedure to assess the Zucchini and Hiemstra (1984) models

Zucchini and Hiemstra (1984) cite Hirsch (1982) and show that all the models which he considered, which are in fact all the models which they present, lead to estimates of the form

$$\hat{y}_i = m(y_1) + s(y_1)(x_i - \alpha(x))/\beta(x)$$

where

$$i = (n_1+1), (n_1+2), (\dots), (n_1+n_2)$$

in which y_i represents the estimated monthly rainfall at the target station in period i , and $\alpha(x)$ and $\beta(x)$ are quantities which can be computed from the observed values. The methods only differ in the manner in which they compute $\alpha(x)$ and $\beta(x)$ which in turn are functions of the following:

$$m(y_1) = \frac{1}{n_1} \sum_{i=1}^{n_1} y_i, \quad s(y_1) = \left[\frac{1}{n_1} \sum_{i=1}^{n_1} (y_i^2 - m(y_1)^2) \right]^{1/2}$$

$$m(x_1) = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i, \quad s(x_1) = \left[\frac{1}{n_1} \sum_{i=1}^{n_1} (x_i^2 - m(x_1)^2) \right]^{1/2}$$

$$m(x_2) = \frac{1}{n_2} \sum_{i=n_1+1}^{n_1+n_2} x_i, \quad s(x_2) = \left[\frac{1}{n_2} \sum_{i=n_1+1}^{n_1+n_2} (x_i^2 - m(x_2)^2) \right]^{1/2}$$

$$m(x) = \frac{1}{n_1+n_2} \sum_{i=1}^{n_1+n_2} x_i, \quad s(x) = \left[\frac{1}{n_1+n_2} \sum_{i=1}^{n_1+n_2} (x_i^2 - m(x)^2) \right]^{1/2}$$

$$r = \left[\frac{1}{n_1} \sum_{i=1}^{n_1} x_i y_i - m(x_1) m(y_1) \right] / s(x_1) s(y_1)$$

For a schematic explanation of notation n_1 , n_2 , y_1 , y_2 , x_1 , x_2 , refer to Figure 4.4.

The six models considered by Zucchini and Hiemstra (1984) calculated $\alpha(x)$ and $\beta(x)$ in the following manner;

Model 1

$$\alpha(x) = m(x_1) \text{ and } \beta(x) = s(x_1)/r$$

Model 2

$$\alpha(x) = m(x_1) \text{ and } \hat{\beta}(x) = s(x_1)$$

Model 3

$$\alpha(x) = m(x) - a s(x)/b \text{ and } \beta(x) = s(x)/b$$

where

$$a = (rn_2/(n_1 + n_2))(m(x_2) - m(x_1))/s(x_2)$$

$$b^2 = [(n_1-1)+(n_2-1)r^2 s(x_2)^2/s(x_1)^2+(n_2-1)c(1-r^2)+ \\ (n_1n_2r^2(m(x_1)-m(x_2))^2/[(n_1+n_2)s(x_1)^2])]/(n_1+n_2-1)$$

$$\text{and } c = n_2(n_1-4)(n_2-1)/[(n_2-1)(n_1-3)(n_1-2)]$$

Model 4

$$\alpha(x) = m(x) \text{ and } \beta(x) = s(x)$$

Model 5

$$\alpha(x) = m(x) + n_2r\{m(x_1) - m(x_2)\}/(n_1+n_2)$$

$$\alpha(x) = s(x) + n_2r\{s(x_1) - s(x_2)\}/(n_1+n_2)$$

Model 6

$$\alpha(x) = k_1 m(x_1) + k_2 m(x_2)$$

$$\alpha(x) = k_1 m(x_1) + k_2 m(x_2)$$

where

$$k_1 = (n_1+rn_2)/(n_1+n_2)$$

$$k_2 = (1-r)n_2/(n_1+n_2).$$

Zucchini and Hiemstra (1984) have the following comments to offer on these models.

Model 1

The simple linear regression model minimises the sum of the squared errors; however, the variance of the resulting estimate of $\hat{y}_i$ is lower than the population variance of the y_i . $\text{Var}(\hat{y}_i) \approx r^2 \cdot \text{Var}(y_i)$ therefore a value of $r = 0,8$ leads to a 36 per cent reduction in variance. The variance is an important parameter in hydrology since the variability of

rainfall affects the variability of streamflow and hence is important in determining the size of storage requirements and operating rules for reservoirs. A systematic bias of this magnitude in the variance is obviously undesirable if the augmented monthly rainfall is to be used in the models whose output is, in turn, used in the determination of storage requirements.

Model 2

In essence this expression amounts to a rescaling of the target station's record with the mean and standard deviation of the concurrent part of the record. No systematic bias is introduced into the augmented record by this model and it has the ability to "correct" the mean and standard deviation to a certain extent according to the non-concurrent part of the target station's record. A disadvantage of this model is that it does not have a minimum sum of squared errors.

Model 3

Matalas and Jacobs (1964) developed this model. The "unbiasedness" of the estimate of monthly rainfall which this model produces is attained at the expense of increasing the standard error of the estimates. According to Zucchini and Hiemstra (1984) simpler models which admit a small bias perform much better than Model 3, for example Models 4 and 5.

Model 4

This model standardises the total target station's record according to its mean and standard deviation and rescales it according to the mean and standard deviation of the control station. The model does not allow the observed mean and variance at the target station to be altered at all,

even if they were based on only a few observations and no matter how strongly these values are correlated to those of the control station. In this sense the model is conservative. By contrast, Model 2 allows these parameters to be altered to a maximum degree (also independently of the correlation).

Model 5

Model 5 behaves like Model 4 if the correlation coefficient is low, whereas it approaches Model 2 increasingly as the correlation coefficient increases, in a sense thus being a mixture of Models 2 and 4. An analogous adjustment is made to take account of the sample sizes n_1 and n_2 .

Model 6

Model 6 is a mixture of Models 2 and 4 and is thereby similar to Model 5.

Although Zucchini and Hiemstra (1984) do not recommend that these models be used for isolated "gap filling" for reasons of computational cost, it was decided to use these methods since the monthly rainfall records used in this project required gap filling as well as augmenting.

4.3 Optimal interpolation techniques

Several models which perform optimal interpolation were investigated. Each model contains slightly different assumptions. These assumptions as well as the advantages and disadvantages of three different models which were investigated are discussed. The methods all assume that the value at any given station is correlated to the corresponding values at neighbouring stations. This correlation structure is used to assign coefficients to a linear model which uses the neighbouring stations as predictors. Two

features of these methods are that the probable error of prediction may be evaluated at each point on the surface and the models are exact interpolators. The three methods which were investigated are:

- (a) Universal Kriging,
- (b) Ripley's method, and
- (c) Gandin's method.

These methods are presented briefly in the following sections and are described in more detail in Appendix B.

4.3.1 Universal Kriging

A basic assumption in the Kriging technique is that the process being modelled is a random surface which is generated by some probability density function. In order to make statistical inferences, some assumptions need to be introduced into the model. The random surface is assumed to satisfy the intrinsic hypothesis, namely, that the first and second moments of the difference

$$D(x,h) = Z(x+h) - Z(x)$$

where

$Z(x)$ and $Z(x+h)$ = the value assumed by the surface at the positions x and $(x+h)$ respectively,

depend only on the distance, h , between the points, and not on the relative positions of the points on the surface.

According to Davis (1973), the major advantage of the Kriging technique is that it provides an estimate of the probable standard error in the interpolation value, thus enabling the spatial distribution of this error to be mapped.

4.3.2 Ripley's method

Ripley's (1980) method assumes that there are many realisations of the same random surface. The covariances and means may be calculated at the sample points directly and no assumptions of stationarity need to be made, although it is assumed that the same process generated all the surfaces. As with Kriging, in the presence of a drift, the estimated mean is first removed from the surface.

An implicit assumption in this method is that all the measured points are useful as predictors at every point on the surface. However, this assumption causes the model to carry a number of redundant predictors when estimating the value at a point.

Furthermore, covariance matrix of the predictors (K) is likely to be ill-conditioned, which makes the solution sensitive to small perturbations in K . This in turn is estimated with a degree of uncertainty, and therefore this method will give unreliable results if K is ill-conditioned (Seed, et al, 1983).

4.3.3 Gandin's method

The problem of having redundant predictors included in the model may be overcome by the use of the interpolation technique proposed by Gandin (1965) and cited by Creutin and Obled (1982). In this method three fundamental assumptions are made, viz:

- (a) The events are realisations of a unique random process or have come from the same population.
- (b) Weak stationarity or homogeneity is assumed when choosing the correlation function.

(c) The mean and standard deviation fields are available.

Two well known techniques for reducing ill-conditioning in least squares regression are the use of normalised variables and of subset selection (Neter and Wasserman, 1974). Gandin's method is far less susceptible to ill-conditioning than Ripley's (1980) method (Seed, 1987). Gandin's method has an advantage over the classical least squares regression in that all the data are used to calculate the means and covariances. In addition a positive definite function may be fitted to the correlation-distance relationship between gauges to estimate the correlations not available directly from the data.

In a comparison of Gandin's method and Kriging, Creutin and Obled (1982) found that although the Kriging method had less constraining assumptions, it did not perform as well as Gandin's method.

4.4 Multiple regression

In least squares analysis there is a difference between numerical ill-conditioning and statistical multicollinearity. Ill-conditioning of a matrix will make the solution of the set of equations unstable. Multicollinearity will be statistically important before ill-conditioning becomes a numerical problem. Multicollinearity has the effect of increasing the variance of the regression parameters, although it does not affect the goodness of fit of the model. A method of reducing the multicollinearity in the model is to select a subset of the predictors (Weisberg, 1980). Subset selection is more difficult with the sets of rainfall data in this project because predictors cannot be viewed in isolation. Predictors must be assessed in the context of the predictors already included in the subset, since the significance of a predictor depends on the other predictors already present in

the model.

The selection procedure should be able to select several subsets, which, although not necessarily the best possible subset, are sufficiently good to be used for predicting new values. This is necessary because the predictor stations in the subset may have gaps in their records for a particular year to be synthesised, and therefore the next best subset with no missing data would have to be used. For this reason all possible subset selection were used to find several subsets that were acceptable for prediction.

There are many statistics which may be used to rank the subset in some logical ranking of 'goodness' in order to find the 'best' subset.

For example the Mallows Cp statistic

$$C_p = \frac{R_{SSP}}{\hat{\sigma}^2} + 2p - n$$

where

- R_{SSP} = the residual sum of squares for the subset size p
 $\hat{\sigma}^2$ = the unbiased estimate of the variance of the errors using the full model and
n = the number of cases in the regression
p = subset size

selects that subset with the minimum mean square error (Weisberg, 1980).

When the independent variables are random and not fixed, as is the case with monthly rainfall totals, the statistic that minimises the mean square error for prediction is the S_p criterion

$$S_p = SSE_p / (n-p-1) (n-p-3)$$

where

SSE_p = sum square errors for the p predictor subset
(Sparks, Zucchini and Coutsourides, 1985).

This criterion was used to rank the subsets for synthesising missing data.

Least squares regression assumes that the variance of the residual or error term is constant for all the cases used in the regression. For monthly rainfall, this assumption may not be correct and the variance of the error is dependent on the magnitude of the response variable. A technique to overcome this problem is to stabilise the variance through the use of the square root, logarithm or inverse transformation (Weisberg, 1980). The square root transformation was selected for this study.

4.5 Comparison of Gandin's method and multiple regression

A comparison of Gandin's method and least squares best subsets multiple regression was conducted by Seed (1987). A summary of his technique and results is presented below.

The two methods were compared using the data from 32 selected stations with more than 50 years' data in the Welding and Havenga (1974) regions 20 Ala and 20 Alc as shown in Figure 4.5.

The first step in the comparison was to use the square root transform suggested by (Weisberg, 1980), on the raw monthly data in order to reduce the influence of months with exceptionally high rainfall. Thereafter a set of predictor stations was selected for each target station using the following algorithm:

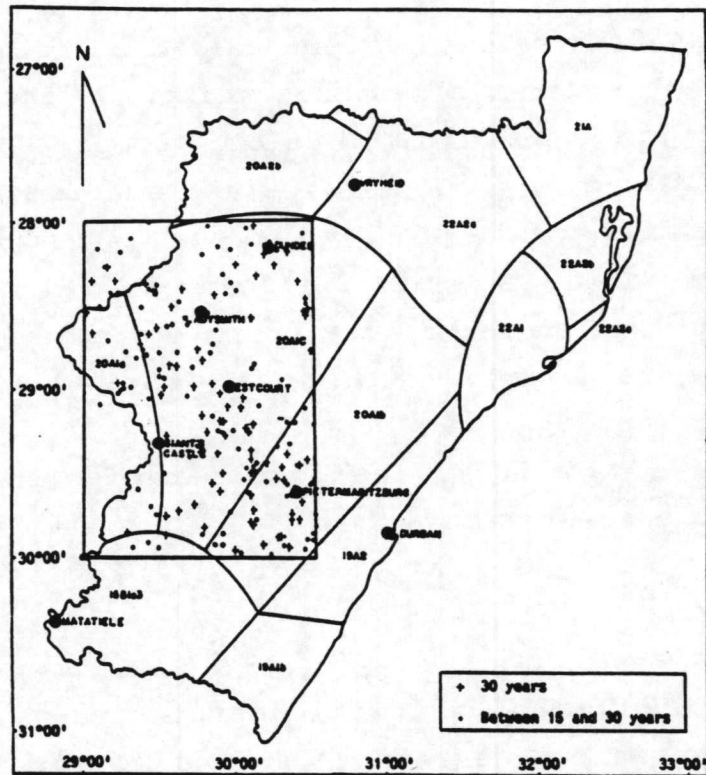


Figure 4.5 The study area in relation to Welding and Havenga (1974) rainfall regions for Natal (after Seed, 1987)

- (a) identify those predictor stations with a correlation greater than 0.7 with the target station,
- (b) rank these predictor stations in descending order to concurrent years of data with the target station,
- (c) delete the stations with the least amount of concurrent data until all stations remaining have a minimum of 30 years of complete data,
- (d) if fewer than three predictor stations are found, skip to the next target station,
- (e) at each station tested, records were selected at random and removed from the data set or "hidden".

The hidden data were then synthesised using Gandin's method and a least squares all-subset selection procedure written by Sparks and Rust (1983). Using a comparison between hidden and synthetic data at selected stations the following were calculated:

- (a) the ratio of synthetic:observed rainfall for every month of synthetic rainfall,
- (b) the ratio of synthetic:observed mean for each station,
- (c) the ratio of synthetic:observed standard deviation for each station and,
- (d) the mean standard error of prediction for each station.

The distribution of the ratios synthetic:observed for the mean and standard deviation of January rainfall are shown in Figure 4.6, 4.7 and 4.8 respectively. It is apparent from Figure 4.6 that Gandin's method produced better synthetic:observed ratios than the best subsets multiple regression.

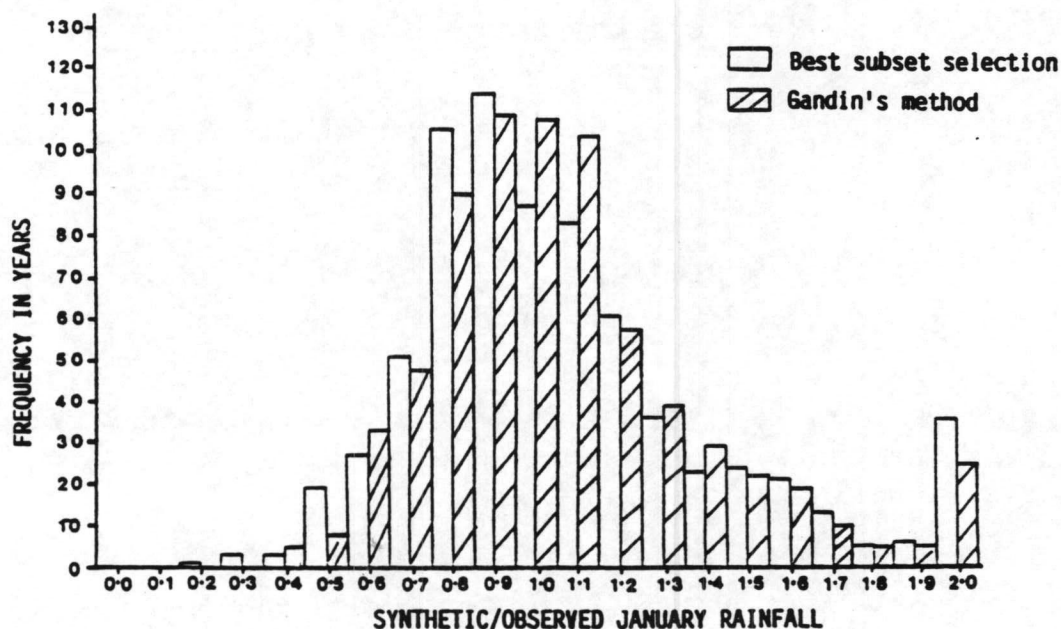


Figure 4.6 Ratio of synthetic:observed January rainfall for best subset multiple regression and Gandin's method (after Seed, 1987)

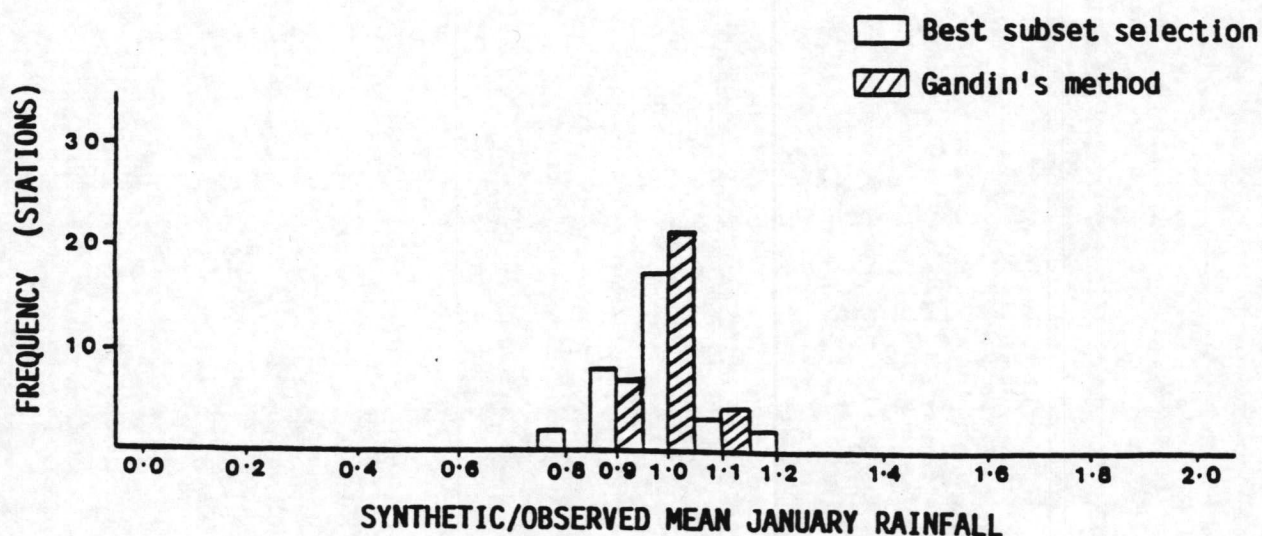


Figure 4.7 Ratio of synthetic:observed mean January rainfall for best subset multiple regression and Gandin's method (after Seed, 1987)

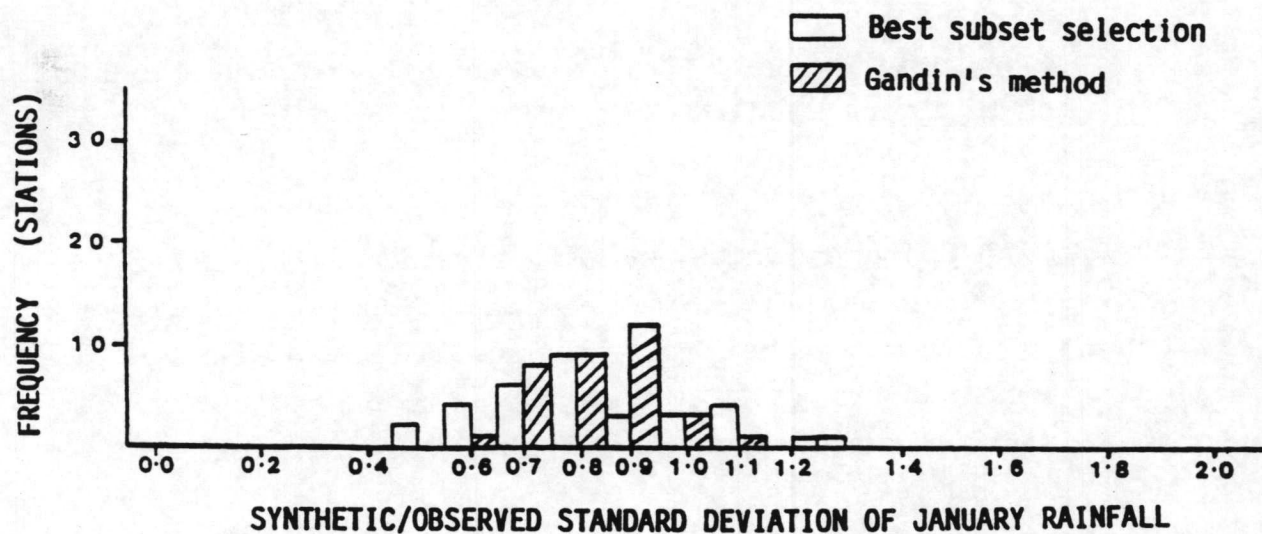


Figure 4.8 Ratio of synthetic:observed standard deviation of January rainfall for best subset multiple regression and Gandin's method (after Seed, 1987)

In addition, Gandin's method produced estimates with a lower mean square error for 23 of the 32 stations tested.

A study of Figures 4.6 to 4.8 led to the conclusion that both of the methods tested were able to preserve the mean of the hidden data. However, the model proposed by Gandin had the following advantages:

- (a) the prediction errors were consistently smaller,
- (b) the distribution of the ratio synthetic:observed mean had a smaller range and
- (c) the distribution of the ratio synthetic:observed standard deviation had a smaller range, and
- (d) the mode classes for both the mean and standard deviation distributions were larger.

On the basis of these advantages, Gandin's method was selected for further comparison with the models reported by Zucchini and Hiemstra (1984).

4.6 Comparison between Gandin's method and the Zucchini and Hiemstra models

Seed (1987) showed that the method of Gandin was superior to Ripley's method, Universal Kriging and multiple regression as discussed in Section 4.5. It remained to test Gandin's method against the six models presented by Zucchini and Hiemstra (1984) to establish the most appropriate model with which to synthesise monthly rainfall data.

The four areas shown in Figure 4.9 were selected as test sites for the comparison of these methods. The areas were chosen for their diversity of rainfall regimes, both within and between the areas. Tests were conducted on January and July monthly data in each case in order to establish the performance of the synthesising techniques, in respective

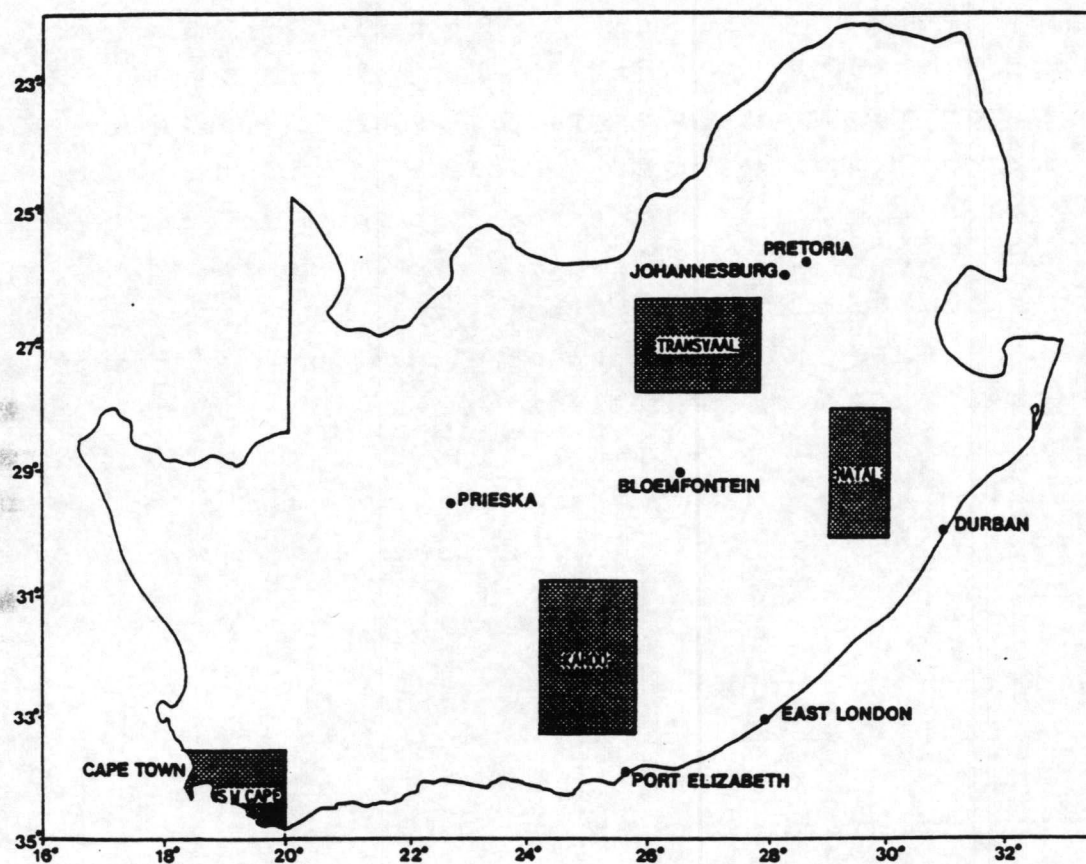


Figure 4.9 Location of the areas selected for testing Gandin's method against the models presented by Zucchini and Hiemstra (1984)

wet and dry seasons. The results of these tests are presented in Table 4.3.

The performances of the models presented by Zucchini and Hiemstra (1984) were tested using the technique described in Section 4.2.3 and illustrated in Figure 4.4.

However, for these techniques to be comparable with Gandin's method, which used all possible stations, up to five stations were used to synthesise as many months as possible using the models of Zucchini and Hiemstra (1984). The proviso was that the control station should have a correlation coefficient of at least 0,5 with the target station before it was used. The most suitable control station according to the 6:3:1 scoring system described in Section 4.2.2 was used first and thereafter the next best and so on until either up to five control stations had been used or the correlation coefficient of the station pair dropped below 0,5. Only stations with 30 or more years of data were selected for this analysis and thereafter each station pair which was used had to have at least 30 years of concurrent record.

The data used to establish the parameters for Models 1 to 6 was not used at all in the evaluation of the model performance. However, with Gandin's method it was not possible to separate the data which was used in the generation of the interstation correlation matrix from the data which was used to test the model. Thus all the data at each station was used to establish the interstation correlation matrix. A section of the data from each station (at least 15 years) was "hidden" and the synthetic data for this period was used to establish the ratio of the mean and the standard deviation of the synthetic:observed monthly rainfall. The wrap around mechanism described in Section 4.2.3 was also used. Gandin's method therefore received an unfair yet unavoidable advantage over the models reported

by Zucchini and Hiemstra (1984).

No one model was consistently "the best" in both wet and dry months and in all four areas. This meant that a system had to be devised whereby the models which performed well consistently could be identified and a rank order of overall performance established. The system which was devised involved the establishment of the frequency distribution of the ratio synthetic:observed mean and standard deviations. The frequency of occurrence of the ratios which fell into the classes 0,7 to 1,3 were obtained. These frequencies (expressed as percentages) and presented in Table 4.3, were summed for all four test areas and seasons. The model which achieved the highest percentage sum was considered to be the best overall. The Model 5 reported by Zucchini and Hiemstra (1984) and described in Section 4.2.3 was selected on this basis. It was not considered sufficient to look at the 0,7 to 1,3 portion of the frequency distribution of the ratio synthetic:observed only but also at the tails of the frequency distribution, since the performance of the model across the whole frequency distribution is important. This analysis revealed that here the Model 5 of Zucchini and Hiemstra (1984) performed better than Gandin's method in that the frequencies of occurrence of observations at the tails of the distribution were less for Model 5 than for Gandin's method.

4.7 Replacement of some missing months by the mean value for that month

The synthesis was carried out on a monthly basis. In other words each column of monthly data at the stations was synthesised independently of the data in other months. The control stations were not necessarily the same for each column of monthly rainfall values to be synthesised. Once all 12 months at each station had been synthesised the

Table 4.3 Model test results, percentage of points in class 0,7 - 1,3 of ratio
(synthetic:observed)

4 - 36

	SW CAPE 31 stations		KAROO 32 stations		TRANSVAAL 34 stations		NATAL 33 stations		TOTAL	AVERAGE	RANK ORDER
MEANS											
	JAN	JUL	JAN	JUL	JAN	JUL	JAN	JUL			
MODEL 1	85	100	86	81	99	78	96	77	702	88	2
MODEL 2	81	98	67	62	88	72	90	71	629	79	5
MODEL 3	67	88	62	29	83	38	90	47	504	63	7
MODEL 4	82	92	88	59	96	60	97	62	636	80	4
MODEL 5	85	98	86	80	98	78	97	83	705	88	1
MODEL 6	77	96	89	72	98	19	96	73	620	78	6
GANDIN	92	95	91	70	99	50	98	77	672	84	3
STANDARD DEVIATIONS											
	JAN	JUL	JAN	JUL	JAN	JUL	JAN	JUL			
MODEL 1	78	83	55	70	64	81	49	74	554	69	5
MODEL 2	77	98	73	73	79	77	80	83	640	80	2
MODEL 3	42	25	38	24	30	36	53	29	277	35	7
MODEL 4	70	94	67	53	81	70	86	79	600	75	4
MODEL 5	85	96	78	75	81	81	79	88	663	83	1
MODEL 6	70	2	59	69	17	14	4	53	288	36	6
GANDIN	60	92	85	72	93	50	87	92	631	79	3

cross addition of monthly values in each year was performed in order to establish the number of newly synthesised years. Unfortunately in many cases not all 12 monthly values were present in order to make up the annual total. Since it would be unwise to discard 9, 10 or 11 "good" months of rainfall data because 3, 2 or 1 month of data were missing from a year after the synthesis procedure had been conducted, it was decided to replace the then still missing month with the mean value for that month.

The increase in the number of complete chronological years of data which are generated using this method was marked. For example for Welding and Havenga region 1 the average number of complete years per station in the original data was 34. After synthesis by Gandin's method the average number of complete years rose to 37. However, if the mean monthly value replaced missing data for up to a maximum of three months in any one year then the average number of complete years per station rose to 47. Therefore it was considered desirable to implement this procedure in the final synthesis, the data thus synthesised were flagged appropriately.

Tests, the results of which are shown in Table 4.4 indicate clearly that it was preferable, both in terms of the mean and the standard deviation to replace the missing months described above with the mean monthly value than to discard an entire year. This procedure was thus adopted.

Table 4.4 Effect on the MAP and the standard deviation of annual totals if a limited number of missing months are replaced by their observed means

95 per cent confidence limits

MISSING MONTH REPLACED BY MONTHLY MEAN										YEAR DELETED
NO. OF MISSING MONTHS	SUMMER			WINTER			ANY SEASON			
	1	2	3	1	2	3	1	2	3	
PERCENTAGE OF DATA ALTERED										
RATIO (ALTERED:OBSERVED) MEAN ANNUAL PRECIPITATION (SW CAPE)										
5	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98
	1.00	1.00	1.00	1.00	1.01	1.01	1.00	1.01	1.01	1.02
10	0.99	0.99	0.99	0.99	0.98	0.98	0.98	0.98	0.99	0.94
	1.00	1.01	1.01	1.01	1.01	1.02	1.01	1.01	1.01	1.02
15	0.99	0.97	0.98	0.98	0.97	0.97	0.97	0.97	0.96	0.94
	1.00	1.01	1.01	1.01	1.02	1.02	1.01	1.01	1.01	1.04
20	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.94
	1.01	1.01	1.01	1.01	1.02	1.02	1.01	1.01	1.02	1.04
RATIO (ALTERED:OBSERVED) STANDARD DEVIATION OF ANNUAL TOTALS (SW CAPE)										
5	0.98	0.96	0.95	0.95	0.94	0.92	0.95	0.91	0.91	0.89
	1.02	1.03	1.02	1.02	1.02	1.03	1.02	1.02	1.03	1.02
10	0.94	0.94	0.91	0.92	0.86	0.82	0.87	0.88	0.86	0.81
	1.01	1.02	1.05	1.04	1.05	1.04	1.03	1.03	1.03	1.04
15	0.93	0.93	0.94	0.87	0.82	0.76	0.85	0.84	0.77	0.52
	1.02	1.02	1.05	1.03	1.04	1.01	1.05	1.04	1.05	1.06
20	0.94	0.92	0.91	0.90	0.87	0.83	0.91	0.89	0.85	0.79
	1.02	1.03	1.05	1.05	1.03	1.04	1.02	1.01	1.02	1.12
RATIO (ALTERED:OBSERVED) MEAN ANNUAL PRECIPITATION (KAROO)										
5	0.99	0.98	0.98	0.99	0.99	0.98	0.99	0.98	0.97	0.98
	1.01	1.01	1.01	1.00	1.01	1.01	1.01	1.00	1.01	1.01
10	0.97	0.97	0.97	0.99	0.98	0.98	0.98	0.98	0.98	0.95
	1.01	1.01	1.02	1.01	1.01	1.01	1.01	1.01	1.02	1.04
15	0.97	0.98	0.95	0.97	0.97	0.98	0.97	0.96	0.96	0.92
	1.01	1.02	1.03	1.01	1.01	1.02	1.01	1.01	1.03	1.04
20	0.98	0.97	0.96	0.99	0.98	0.97	0.98	0.98	0.97	0.92
	1.02	1.02	1.03	1.01	1.01	1.02	1.01	1.02	1.02	1.06
RATIO (ALTERED:OBSERVED) STANDARD DEVIATION OF ANNUAL TOTALS (KAROO)										
5	0.96	0.91	0.90	0.97	0.96	0.96	0.96	0.93	0.91	0.92
	1.02	1.02	1.06	1.02	1.01	1.02	1.01	1.04	1.02	1.01
10	0.86	0.78	0.74	0.92	0.88	0.85	0.95	0.85	0.81	0.74
	1.03	1.03	1.04	1.01	1.02	1.02	1.03	1.03	1.02	1.04
15	0.89	0.87	0.84	0.96	0.94	0.90	0.91	0.91	0.88	0.83
	1.04	1.05	1.06	1.03	1.05	1.04	1.04	1.05	1.05	1.06
20	0.86	0.81	0.79	0.92	0.92	0.91	0.93	0.86	0.83	0.77
	1.04	1.03	1.09	1.03	1.03	1.03	1.04	1.05	1.09	1.09

Table 4.4 (contd)

95 per cent confidence limits									
NO. OF MISSING MONTHS	MISSING MONTH REPLACED BY MONTHLY MEAN								
	SUMMER			WINTER			ANY SEASON		
	1	2	3	1	2	3	1	2	3
PERCENTAGE OF DATA ALTERED	RATIO (ALTERED:OBSERVED) MEAN ANNUAL PRECIPITATION (NATAL)								
5	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98
	1.01	1.01	1.01	1.00	1.00	1.00	1.00	1.01	1.00
10	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	1.00	1.01	1.02	1.00	1.00	1.01	1.00	1.01	1.02
15	0.99	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.98
	1.01	1.02	1.02	1.00	1.01	1.01	1.01	1.01	1.01
20	0.99	0.98	0.97	0.99	0.99	0.99	0.99	0.99	0.98
	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
	RATIO (ALTERED:OBSERVED) STANDARD DEVIATION OF ANNUAL TOTALS (NATAL)								
5	0.95	0.96	0.93	0.95	0.93	0.92	0.93	0.92	0.94
	1.02	1.03	1.02	1.01	1.02	1.02	1.01	1.01	1.05
10	0.92	0.91	0.90	0.89	0.91	0.88	0.94	0.91	0.90
	1.03	1.05	1.02	1.02	1.01	1.01	1.01	1.02	1.01
15	0.92	0.88	0.84	0.90	0.88	0.89	0.94	0.89	0.86
	1.03	1.02	1.01	1.02	1.02	1.04	1.03	1.02	1.02
20	0.91	0.88	0.85	0.94	0.91	0.91	0.95	0.90	0.90
	1.04	1.02	1.04	1.02	1.03	1.04	1.04	1.04	1.06
	RATIO (ALTERED:OBSERVED) MEAN ANNUAL PRECIPITATION (TRANSVAAL)								
5	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98
	1.00	1.01	1.01	1.00	1.00	1.01	1.00	1.01	1.01
10	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.98
	1.01	1.01	1.02	1.00	1.01	1.01	1.01	1.01	1.01
15	0.98	0.98	0.98	0.99	0.99	0.98	0.99	0.98	0.98
	1.01	1.02	1.03	1.01	1.01	1.01	1.01	1.01	1.01
20	0.98	0.97	0.96	0.98	0.98	0.98	0.99	0.97	0.97
	1.01	1.02	1.02	1.01	1.01	1.01	1.01	1.02	1.02
	RATIO (ALTERED:OBSERVED) STANDARD DEVIATION OF ANNUAL TOTALS (TRANSVAAL)								
5	0.96	0.95	0.90	0.96	0.96	0.97	0.96	0.94	0.94
	1.02	1.05	1.03	1.01	1.02	1.03	1.02	1.03	1.04
10	0.89	0.88	0.89	0.98	0.96	0.95	0.93	0.90	0.88
	1.02	1.03	1.02	1.01	1.01	1.01	1.02	1.02	1.02
15	0.93	0.86	0.85	0.97	0.92	0.92	0.89	0.89	0.87
	1.05	1.03	1.10	1.02	1.01	1.01	1.04	1.06	0.03
20	0.87	0.83	0.78	0.95	0.91	0.90	0.95	0.84	0.77
	1.09	1.04	1.04	1.06	1.04	1.03	1.05	1.05	1.04

CHAPTER 5

AUTOMATION OF THE REGRESSION PROCEDURE FOR RELATING MEAN ANNUAL PRECIPITATION TO PHYSIOGRAPHY

The large volumes of data, numbers of variables and possible regression models made it imperative to develop techniques to automate the management of the procedures leading up to and including the regression procedure for relating MAP to physiographic variables. An intricate system involving numerous major procedural steps and incorporating 27 programs in FORTRAN or Statistical Analysis System (SAS), was developed to perform the abovementioned functions.

The discussion in this chapter will include:

- (a) an outline of the procedural steps in the regression process,
- (b) the techniques of automated data management of the variables used in the regression analysis,
- (c) the treatment of residuals,
- (d) the response surface outside the predictor space,
- (e) weighting and transforming the regression variables,

Throughout the system there were points at which human intervention and checking was necessary. In systems and data sets as large as the ones which were being manipulated in this project it would be easy to "lose track of things" if the entire procedure was automated. Errors could become compounded and or lost if checks, either automatic or manual were not introduced at specific points in the process. Despite the fact that extensive checks, as dis-

data entered the regression procedure it was still possible for data errors to be present. These errors would manifest themselves at various stages in the procedure to be outlined and would be trapped, corrected or removed. The theme of care, error checking and the "feel" for what is correct and natural, especially in terms of the final isohyets is prevalent throughout the discussion which follows. It was in this aspect of the research work that the judgement and the science moulded together.

The process of achieving the best regression expression for the rainfall surface was therefore iterative in the sense that the isohyetal maps were actually produced and then studied before accepting, rejecting or modifying the regression expression. Outliers, anomalies and hyper- or insensitivities of the surface were thus detected.

5.1 Procedural steps in the regression process

Step 1

The corner co-ordinates of the study area were defined and the search and extract program described in Section 3.4 was used to retrieve the rainfall statistics and monthly data for each station located in the area of interest from the DATABASE file. Southern Africa was divided into the 34 regions shown in Figure 5.1 for the purpose of establishing regression equations for MAP. Data from all stations were extracted, irrespective of record length, and the output file (EXTRACT) was stored temporarily until the regression expression for that area was finalised. EXTRACT was used as a quick reference for any queries which arose as a result of comparing the rainfall statistics to be mapped with those at nearby stations.

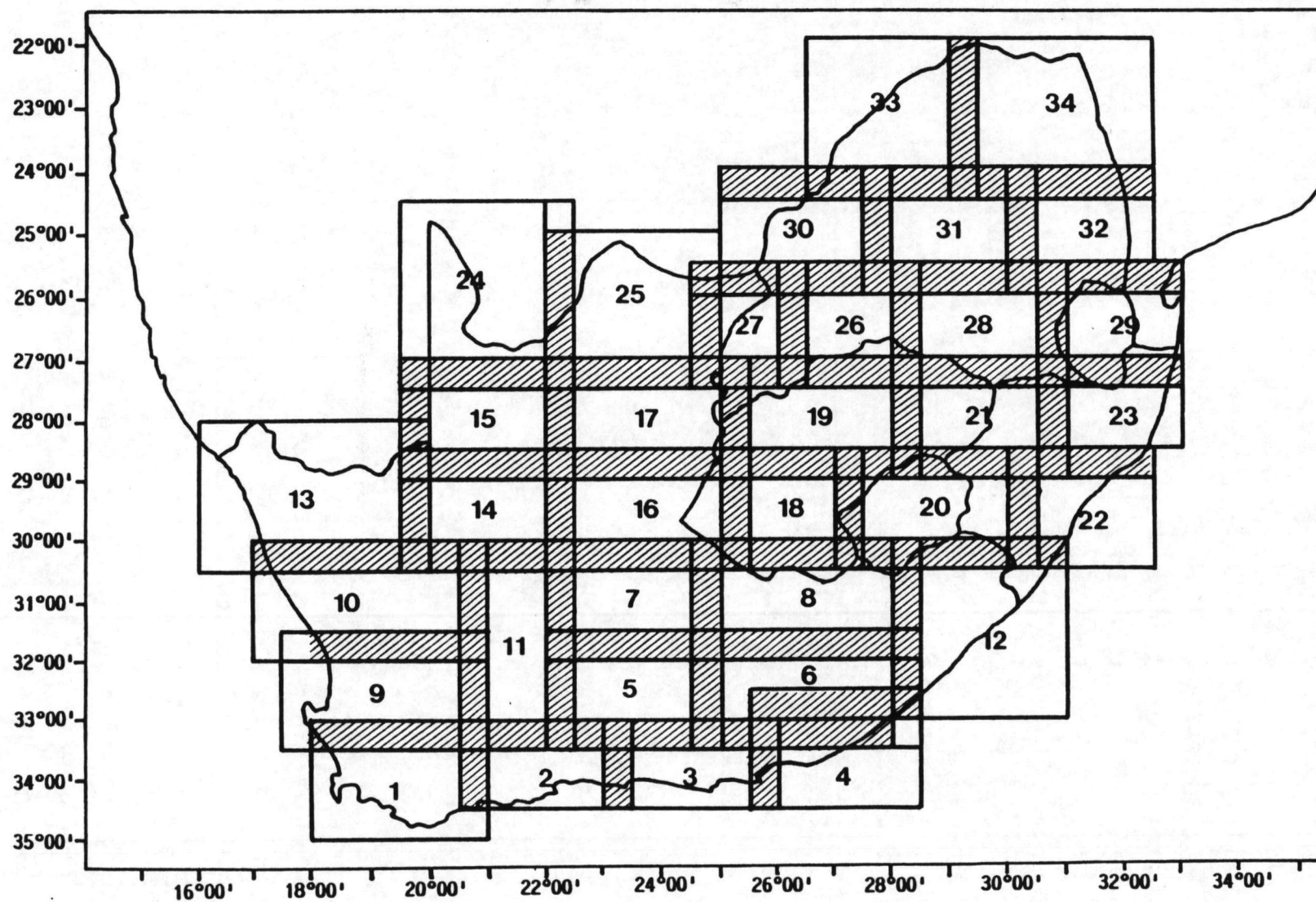


Figure 5.1 Regional boundaries used in the regression analyses

Step 2

A program extracted the station number, latitude, longitude, altitude, number of years of record, MAP and standard deviation of annual totals from the EXTRACT data file and sorted these stations into descending order by record length. The minimum record lengths used in each area are those given in Section 4.1 Figure 4.3. These minimum record lengths were adhered to as far as possible. However, in certain selected cases it was necessary to relax these criteria for one or two stations in an area in order to obtain data in "vital" areas, such inclusions having been motivated by the need for adequate spatial and altitudinal coverage of stations in the area. The aim was to endeavour to obtain the entire grid point variable vector space within the station point variable vector space. This ideal, which was never attained, was desirable since the regression expression would then be interpolating and very little extrapolation beyond the regression vector space would be required.

Step 3

If two stations had the same location, problems were experienced when the residual analysis was conducted. Therefore a program was run to identify the stations which had the same latitude and longitude. The altitudes, MAP's and standard deviations of annual total rainfall at these stations were then scrutinised and if the altitudes of the two stations were almost equal then the station with the longest record length was accepted. In addition the exact (to within 0,1 minute) location of the stations were sought from INDEX and the stations plotted on the respective map of scale 1:250 000. Thereafter one of the stations was "relocated" temporarily to the latitude and longitude nearest the other station so as to resolve the contradiction.

Step 4

The gridded image of latitude, longitude, altitude and distance from the sea values for the study area was extracted from SAGRID. A border strip, with a width of 2 minutes of a degree was added to the study area to enable the grid points on the edge of the study area to be the centre of a complete 5 minute by 5 minute mask when the roughness, exposure and aspect indices were calculated.

Step 5

The altitudes on SAGRID were replaced by those from INDEX i.e. rainfall station altitudes, where appropriate. The rationale behind this step was to prevent unnecessarily large residuals being calculated when the regression expression was applied to the point altitude at the station point.

Step 6

The locations of all stations in the study area were plotted along with the station altitude, number of years of record and MAP. These plots were consulted closely when the regressions were conducted to ensure that the physical relevance of the station MAP in the context of its spatial and altitudinal position, was borne in mind constantly.

Step 7 - 9

The roughness, exposure and aspect indices were calculated for each point in the SAGRID image of the study area. These coefficients are described in Section 2.5.

Step 10

The indices calculated in Steps 7 - 9 were merged with the file extracted in Step 2. These data were then ready for the regression procedures using SAS.

Step 11

A number of transformations were performed on the raw data set of variables. The following transformations were performed on each of the independent variables:

- . cubes
- . square roots
- . logarithms
- . reciprocals.

In addition the cross products of every combination of pairs of variables were calculated. Analyses based on the Box and Cox (1964) techniques, were performed on the dependent variable i.e. MAP, for each region. In most cases no transformation of the dependent variable was necessary.

Clarke (1986), suggested that the independent variables, viz, latitude, longitude, continentality, roughness, exposure and slope aspect be forced into the regression, since these are the primary variables and have a physical meaning.

Thereafter the stepwise regression procedure using forward selection was engaged to develop a series of models. The final selection was then made from these models, using the procedure described in Step 12.

Step 12

A FORTRAN program extracted the coefficients from the SAS

output in Step 11 and set up the models for the next of the SAS procedures which processed every model generated in the previous steps and calculated the predicted error sum of squares (PRESS) statistic, described by Younger (1979), for each model. The model with the lowest PRESS statistic was then selected. The most appropriate model for this work is not necessarily the one which fits the data best but that one which predicts best; a procedure for which the PRESS statistic gives an answer by removing one station MAP at a time and recalculating the coefficients of the model without the hidden data. Thereafter the residual is calculated at the hidden station. The technique is repeated for each model and for each station in the data set.

Step 13

A provisional selection of the final model to be used to estimate MAP was made using the PRESS statistic calculated in Step 12. The residuals were then calculated, plotted and checked to see whether they indicated model failure in the spatial or altitudinal context, by revealing clusters of stations with high residuals. Once satisfied that the surface represents the spatial distribution of rainfall well, Step 14 was commenced.

Step 14

A gridded image of percentage residuals was interpolated onto a regular grid from the random scatter of station points. The regression image was multiplied by the image of percentage residuals as described in Section 6.3. A check was then conducted on the gridded digital image to ensure that the MAP surface fitted the data exactly at the station points.

Step 15

In the case of coastal or border regions, all points outside the northern border of the Republic of South Africa or off the coast were blanked out except for a 2 minute boundary zone. Values in this zone ensured that the interpolation and contouring was not distorted at the borders or coastline.

Step 16

A table of the regression variables and the coefficients, which were used in the final models for each region, was retained on computer file.

Step 17

The gridded image file of estimated MAP was then processed through the contouring routine in the SACLANT GRAPHICS PACKAGE (SACLANT), described by Diederiks (1979), which interpolated from the 1 minute grid onto a 0,25 minute by 0,25 minute grid before plotting isohyets of MAP.

5.2 Examination and treatment of residuals

An important practical consideration in the application of trend surface analysis is the identification and possible omission of stations with abnormal rainfall records. A large residual which is anomalous with the residuals in the vicinity of the station is generally an indication of anomalous data, either in the regression variables or in the rainfall data itself, unless there were obvious topographical reasons for such an estimated or real value. All such high residuals were examined individually at the stage that they were produced by SAS and if there was doubt as to their validity they were allowed to remain in the data set until the map of residuals was studied in order to establish an opinion on the residual in the context of other residuals in the vicinity. Clusters of high

residuals of the same sign indicated model failure in an area. In such cases the physiography and the regression variable data as well as the model were examined to determine the cause of the cluster of high residuals.

Also at the regression stage it was necessary to seek regions of the regression space which exhibited a dearth of rainfall stations and hence generally wide confidence levels in the regression estimate. The technique reported by Neter, Wasserman and Kutner (1983) and discussed in Section 6.2 was used to identify such regions. In some instances the record length limits presented by region in Appendix A and discussed in Section 4.1, were relaxed to include stations of shorter record length in crucial areas for example Jonkershoek.

5.3 The response surface outside the predictor space

It was found generally that polynomials of high order yielded progressively increased adjusted multiple correlation coefficients (r^2) and decreased the errors due to sum of squared differences. However, there are dangers inherent in using polynomials of too high an order, especially with an inadequate station network. The isohyetal maps plotted from such expressions may produce quite unrealistic and even physically impossible point values of rainfall when extrapolating beyond the regression data space (Schulze, 1979). This lack of "control" over the response surface once the predictors move outside the regression data set is a real problem. An investigation into the suppression of extrapolated estimates when the predictor variables yielded estimates which were higher or lower than the highest value in the regression data set, was therefore instituted.

Schulze (1979) used an estimated value of

$$P = HD + 3(\sqrt{DIF})$$

where

P = adjusted estimate of the point value

HD = highest data input value of MAP

DIF = difference between the estimated point value and the highest data input value of MAP for estimated points which were higher than any MAP in the data set and

$$P = LD - 3(\sqrt{DIF})$$

where

LD = lowest data input value of MAP

The suppression factor is largely subjective. Hence, after a number of tests, the following procedure was adopted for suppression of high and low values of MAP which fell outside the predictor space.

Cut-off levels were fixed at 10 per cent above and below the highest and lowest value of MAP in the regression data set. Estimated MAP values which exceeded these limits were removed from the gridded data set of estimated MAPs. Interpolation was then used to estimate the MAP value at the point using a 5 x 5 minute mask of points surrounding the removed point. In this way the surrounding MAP values were allowed to influence the value at the high or low point and apart from the cut-off level of 10 per cent, no arbitrary value was assumed.

5.4 Weighting and transforming the regression variables

Since variance of the estimate of MAP is different at each station, this variable variance necessitated the use of the weighted least squares regression technique. In addition the confidence of the estimate of MAP varies with record

length as described in Section 4.1. Record length was therefore also brought into the weighting factor which was derived from the product of the record length and inverse of the variances. The model therefore attenuates those cases in which the record length is short and or the variance is high.

A number of transformations were performed on the variables and new variables were created by performing cross products and raising the independent variables to powers. With regard to the dependent variable viz map the techniques of Box and Cox (1966) were used to determine whether any transformation on the MAP was appropriate. Their technique was applied to each of the 34 regions in which regression equations were determined. No transformation of the MAP was indicated by this technique in any of the regions according to the Box and Cox technique.

5.5 Discussion

This section has dealt primarily with the techniques and procedures for obtaining the best regression equation relating MAP to physiographic variables, and transformations and cross products of these variables. In the process of obtaining the regression equation for the rainfall surface it was necessary to plot isohyets of MAP. The latter process introduces several problems which have been alluded to in the preceding section and will be discussed more fully in Chapters 6 and 7. Brief mention has been made of the composite residual/regression surface technique which enables the final surface to fit the observed rainfall exactly. This aspect is discussed in Section 6.3.

The problem of continuity at regional boundaries has also been mentioned and the regions which were used have been presented. A more detailed discussion of the solutions to this problem is given in Section 7.3.

In addition to the regression approach the technique of direct interpolation was also investigated in this study and compared with the regression technique in section 6.1.

The confidence of the estimate of MAP at non station points is an important aspect which is discussed in Section 6.2.

CHAPTER 6

MAPPING STATISTICS OF PRECIPITATION

The complete automation of the process of mapping rainfall statistics is not possible due to the many subjective decisions which are required (McKay, 1976). However, the computer is a useful aid to decision making by assisting in evaluating, preparing and controlling the quality of the data as well as in its plotting. According to McKay (1976), the difficulties in interpreting topographical effects have been a major deterrent to the more complete automation of computerised mapping. Experience in this project has confirmed the views of McKay (1976) who notes that the problem of mapping rainfall in mountainous areas is formidable. Bearing in mind the many mountainous areas in Southern Africa it becomes evident that the mapping was no easy task. Compromises, assumptions and some subjectivity were required to produce maps which are useful to the average user and which have scientific merit.

The techniques of mapping fall into two broad categories, viz, trend surface analysis based on multiple regression and direct interpolation. The merits of trend surface analysis and direct interpolation are investigated in this chapter. A combination of the regression surface and the surface generated by interpolating between the residuals was considered finally to be the most pragmatic solution. The MAP surfaces only have real merit if their quality is known. Therefore, residual analyses and estimation of confidence levels of regression estimates were performed on the rainfall data. Furthermore the difficulty of attaining continuity of the MAP surface at the boundaries of each regional regression surface is discussed and solutions are presented.

The surface of MAP provided the foundation for the digital mapping of mean monthly precipitation and the variability of monthly and annual precipitation at a resolution of 1 minute of a degree and such digital images were produced. The latter maps were produced at a scale of 1:1 000 000.

6.1 Comparison of techniques of mapping

The mathematical and statistical techniques which are employed to distribute point rainfall estimates of MAP spatially may be divided into two main groups, viz:

- (a) direct interpolation and
- (b) trend surface analysis based on multiple regression of physiographic variables related to the station

The appropriateness of the technique is dependent on a combination of the physiographic features and the distribution of rainfall stations in the area in which the technique is to be employed. The north eastern Orange Free State (OFS)/southern Transvaal (TVL) and the Drakensberg region of Natal are areas with very different rainfall station distributions and physiography. These areas were therefore selected to compare the two techniques mentioned above. The areas are shown in Figure 6.1.

Mandeville and Rodda (1970) state that the employment of interpolation techniques, which rely on the fundamental assumption of the juxtaposition of similar amounts of rainfall, should be confined to flatter areas where rainfall gradients are not marked. These techniques are considered to be particularly suited to areas where the distribution of rainfall stations is good in relation to the complexity of the terrain and the anticipated rainfall surface. Direct interpolation techniques were therefore expected to perform well in the OFS/TVL test area.

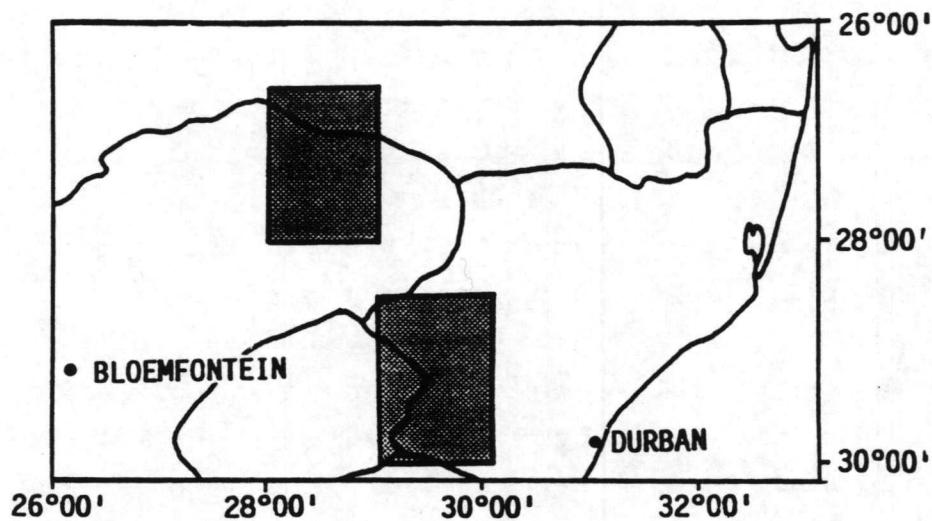


Figure 6.1 Location of the OFS/TVL and Drakensberg study areas

The Natal Drakensberg and its foothills are important areas of water production in South Africa. The areal estimation of rainfall is therefore of major importance to water resources planners in this region. Here the distribution of raingauges is sparse in relation to the complex physiography and the variability of the rainfall surface. It was therefore considered to be suitable to undertake rainfall mapping in this area using techniques which are able to accommodate these complexities. The influence of physiography on the spatial distribution of MAP has been well known for a long time, as is evidenced by the studies reviewed in Section 2.1. It was anticipated that trend surface analysis would be the best technique to use in this area. Extensive background to the technique of trend surface analysis is not provided here since the technique has been documented *inter alia*, by Schulze (1976, 1977, 1979) and Hughes (1982) as well as having been dealt with in some detail in the preceding chapters.

It was considered reasonable to assume that each technique would perform well under certain circumstances. However,

the question as to which technique was best in each area in turn raised the question as to whether a joint trend surface/interpolation approach would be appropriate. A combination of techniques was indeed chosen as the final mapping procedure. A further point to bear in mind is that the "true" rainfall surface is never known. The traditional tests for goodness of fit can therefore not be applied to the observed versus estimated surfaces.

It was not possible to test the performance of the interpolation technique by the traditional method of "hiding" stations' MAPs and then assessing the goodness of fit. The interpolation technique was sensitive to the location and magnitude of MAP. In certain instances the removal of a station made little or no difference to the interpolation estimate and in other cases the differences were marked. One means of gaining a subjective assessment of the interpolation techniques was to investigate the nature, spatial location and extent of the differences between the estimates of MAP produced by the interpolation and regression techniques.

Such a study was conducted in two areas which differ markedly with respect to their physiographic features as is illustrated in the three dimensional depictions in Figures 6.2 and 6.3. The vertical scales are the same in both these diagrams. A quantitative assessment of the differences is presented by the histograms in Figures 6.4 and 6.5 which show the distribution of station altitudes and grid point altitudes. It is significant that 20 per cent of the altitude points in the Drakensberg area are above the highest rain station. Although it is not possible to determine the true MAP at these altitudes, owing to the absence of stations, the differences in estimates of MAP, using regression and interpolation, are marked as Figures 6.6 and 6.7 illustrate.

For simplicity's sake the only independent variables used in the regression equation for trend surface analysis, in this preliminary investigation were latitude, longitude and altitude. There are indeed several other variables which may be used to estimate MAP, for example, continentality, roughness, aspect in relation to the rainbearing wind direction and exposure.

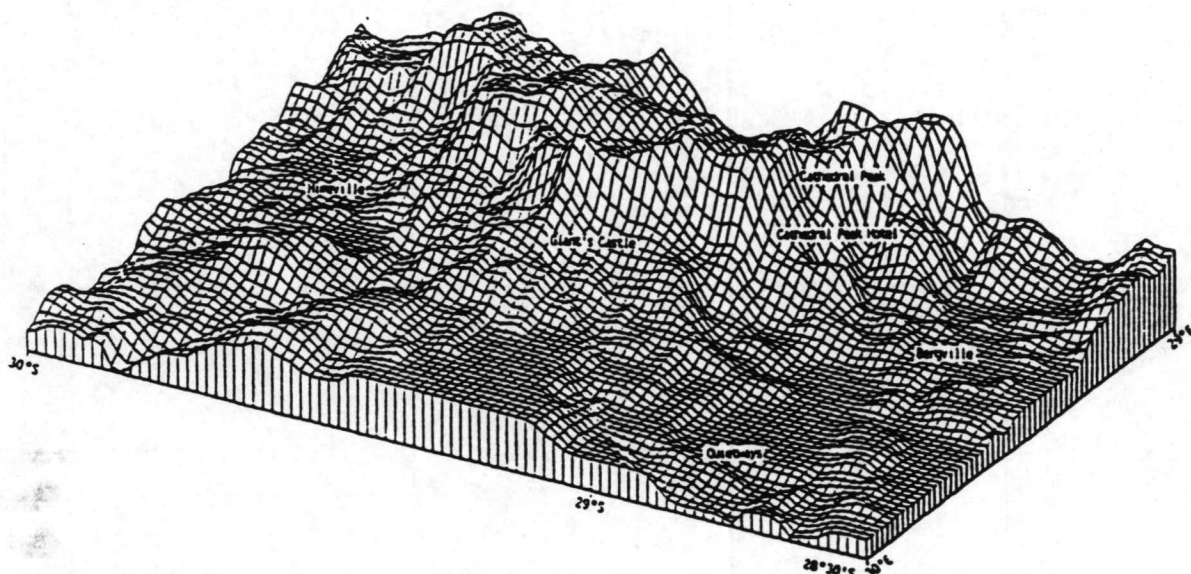


Figure 6.2 Three dimensional representation of the physiography of the Drakensberg study area

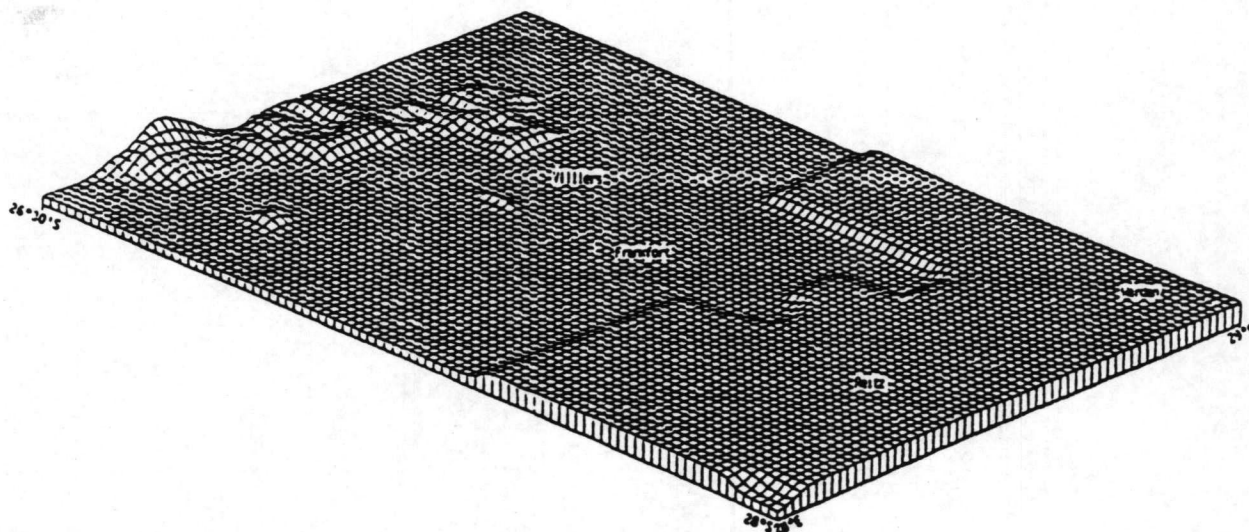


Figure 6.3 Three dimensional representation of the physiography of the OFS/TVL study area

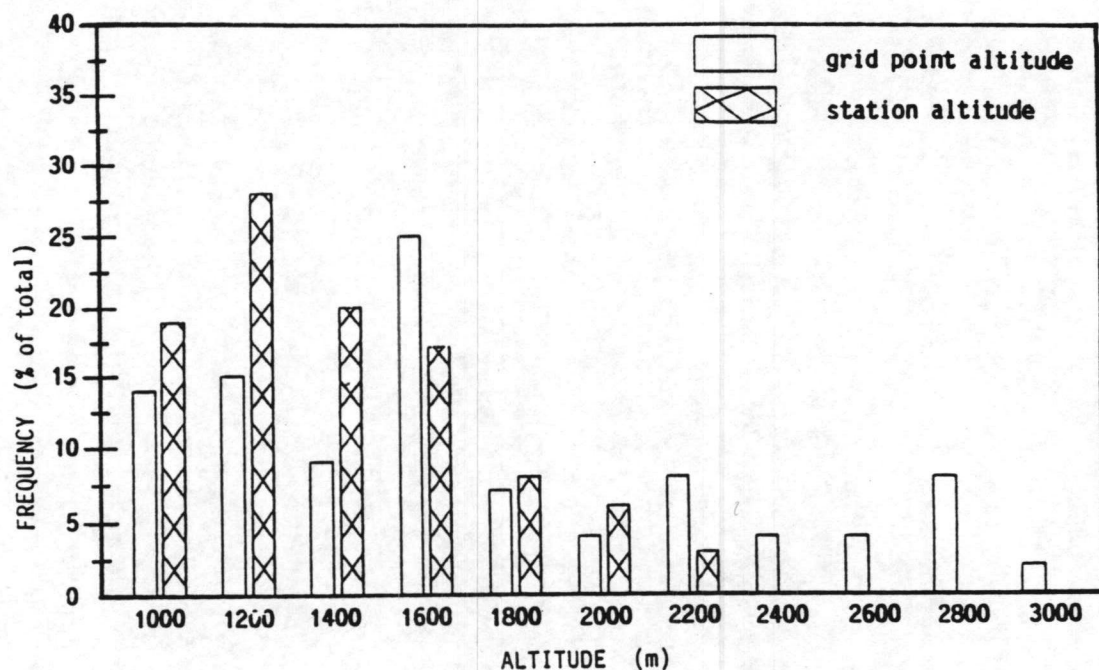


Figure 6.4 Frequency distribution of grid point and station point altitudes in the Drakensberg study area

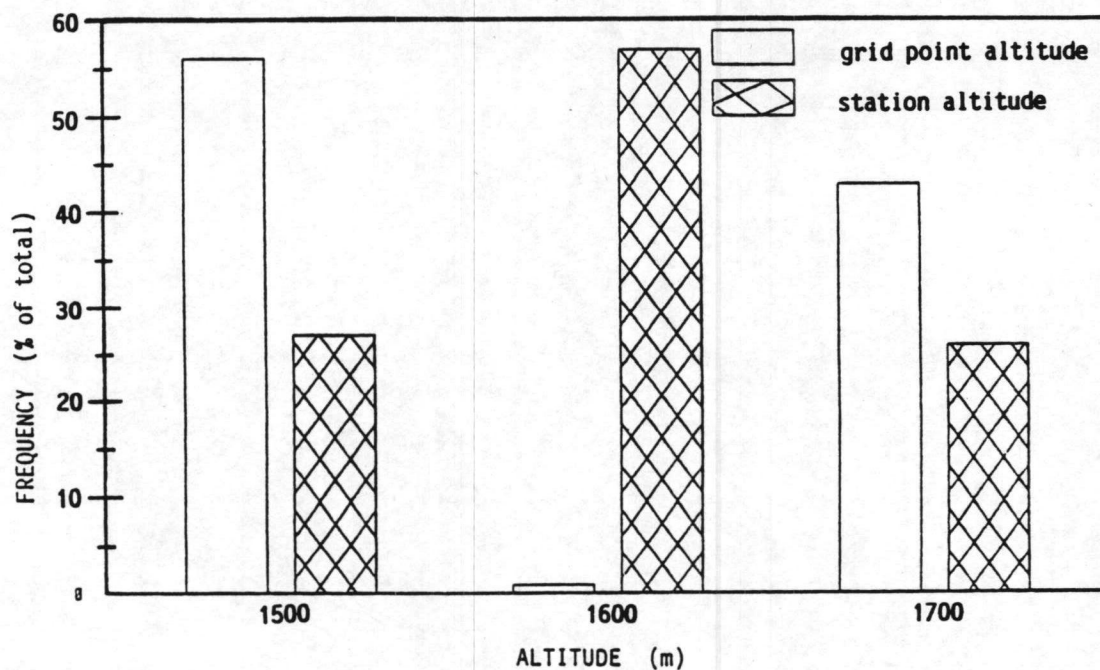


Figure 6.5 Frequency distribution of grid point and station point altitudes in the OFS/TVL study area

tervals provided the framework on which to map surfaces of MAP using interpolation (I) and regression (R). The differences between the two techniques are expressed as a percentage of the regression/trend surface analysis estimates in the Drakensberg and as a percentage of the interpolation estimate in the OFS/TVL. The differences are mapped in Figures 6.6 and 6.7, which also show the location of the rainfall stations used in the analyses. The shaded portions in Figure 6.7 depict areas in which the difference in MAP estimates by the two techniques is greater than 20 per cent.

It is evident from Figure 6.6 that there is little difference between the estimates of MAP in the OFS/TVL. The area is flat and the spatial distribution of stations is good. In addition there is no strong correlation between MAP and location with altitude. This is evident from the correlation matrix shown in Table 6.1. The adjusted r^2 for regression of longitude, latitude and altitude with MAP is also very low for the OFS/TVL with only 12 per cent of the variation of MAP explained by latitude, longitude and altitude. In certain parts of the Drakensberg area where the rainfall station distribution is sparse in relation to the physiographic complexity of the area, the estimates differ markedly.

The particular advantage of the trend surface analysis used is that it is able to fit a complex surface where no stations exist. Generally stations are not found in deep valleys and on high hills or mountains. The interpolation technique does not have the ability to fit surfaces with the same degree of flexibility as trend surface analysis in areas which are not gauged. This point is most aptly illustrated in the profile diagram shown in Figure 6.8. The profile is along line AA' in Figure 6.7. The interpolation

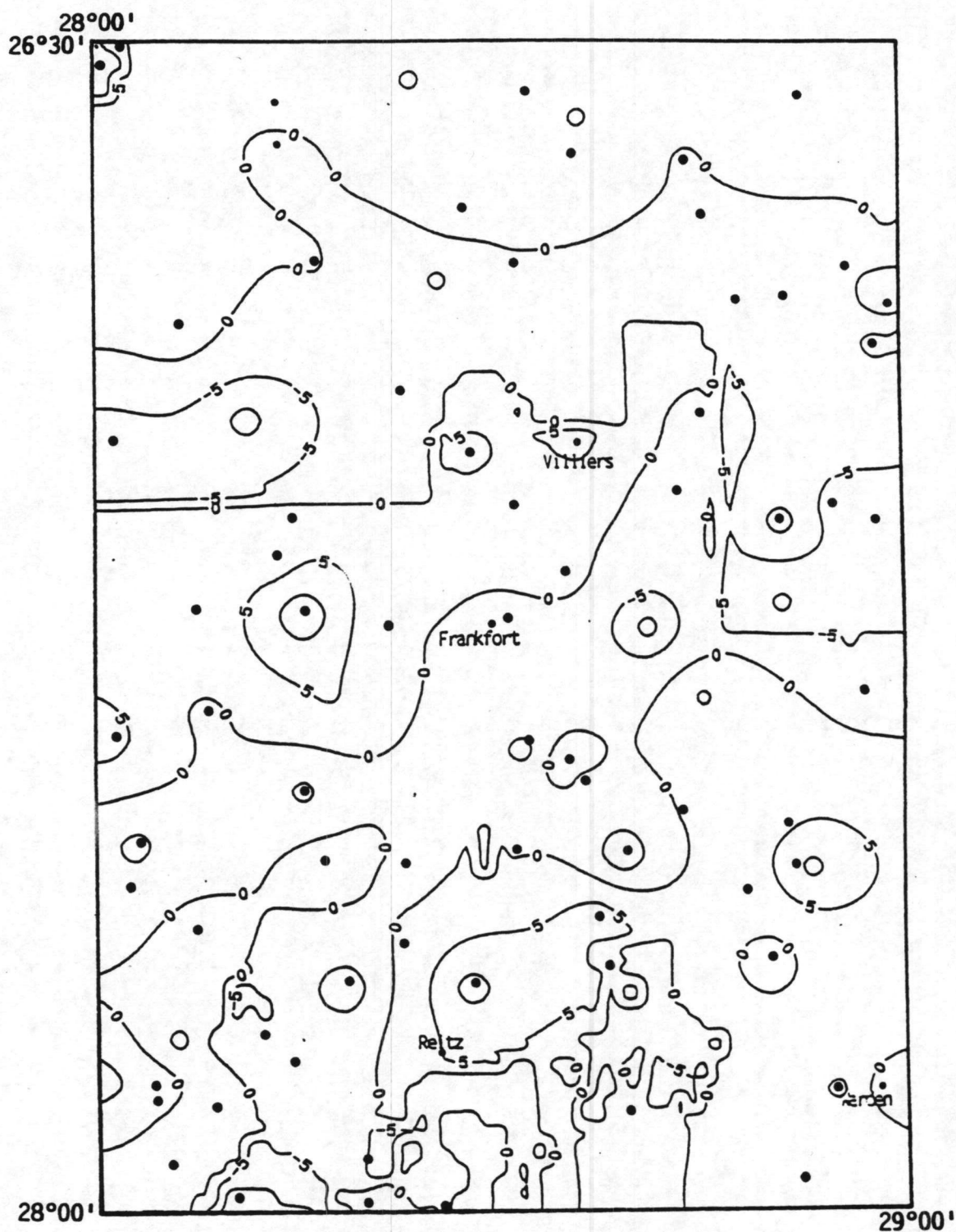


Figure 6.6 Percentage difference between the estimates of MAP using interpolation (I) and regression (R) in the OFS/TVL study area (Difference = $((I - R) / I * 100)$)

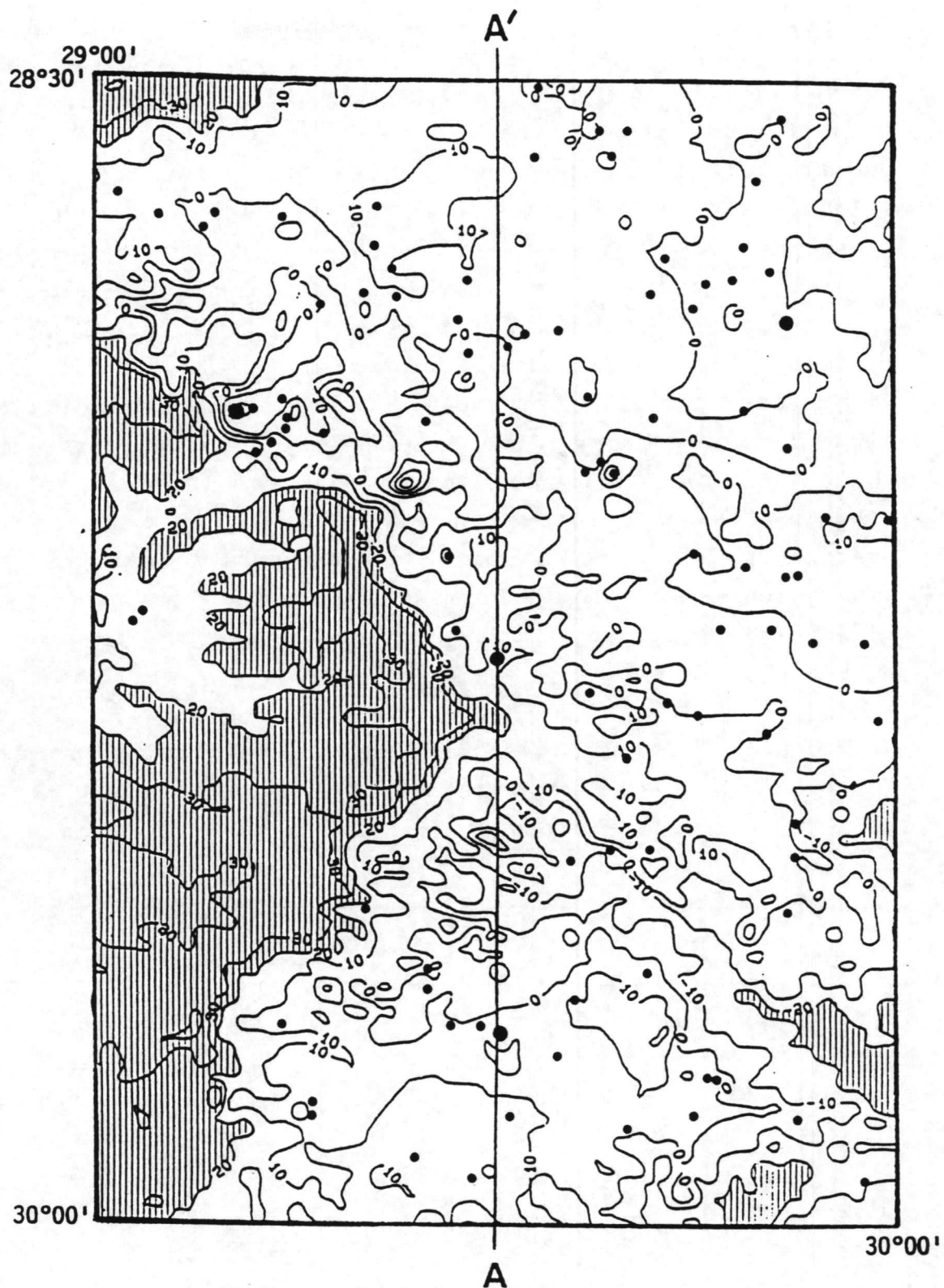


Figure 6.7 Percentage difference between the estimates of MAP using regression (R) and interpolation (I) in the Drakensberg study area (Difference = $((R-I)/(R * 100))$)

technique, by its very nature, fits the recorded data at the station points. The interpolation line profile is not a straight line between the station points due to the influence of the station points to the side of the profile line which are not shown on Figure 6.8. These stations do, however, have an effect on the spatial interpolation. The regression estimates are obviously not as accurate as the interpolation estimates at the station points. However, it is evident that the interpolation estimates have no "feel" for the topography. The MAP is very strongly correlated with altitude in this area as may be seen in Table 6.1 and the regression expression explains 73 per cent of the variation in MAP. The plot of observed vs estimated MAP using regression in the Drakensberg area is shown in Figure 6.9. The regression model used in this example predicts at the higher MAPs.

Table 6.1 Correlation matrices and adjusted r^2 from regression in the Drakensberg and OFS/TVL study areas

DRAKENSBERG					OFS/TVL			
	LONG	LAT	ALT	MAP	LONG	LAT	ALT	MAP
LONG	1.000	.241	-.208	-.458	1.000	-.143	.026	.296
LAT		1.000	.517	.370		1.000	.585	-.069
ALT			1.000	.800			1.000	.185
MAP				1.000				1.000
	r^2 adjusted = 0.735				r^2 adjusted = 0.119			

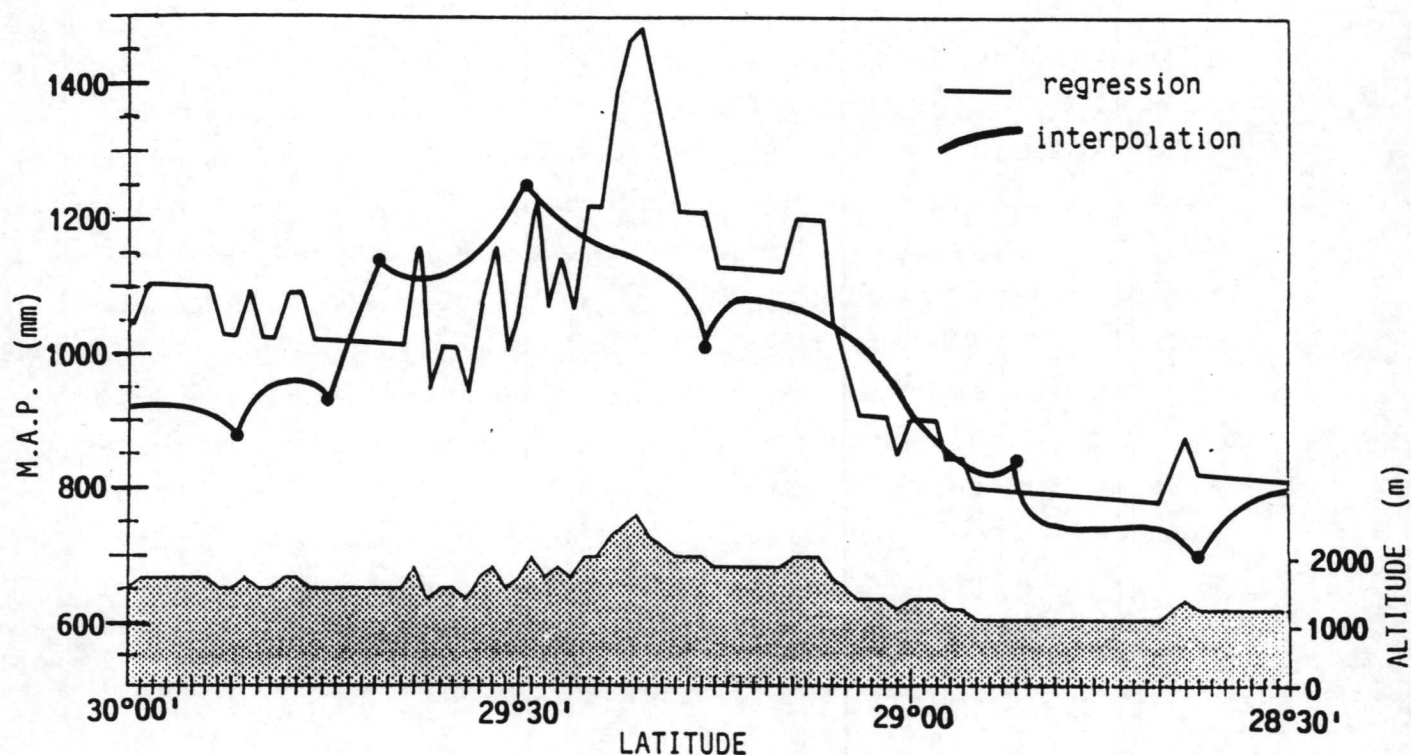


Figure 6.8 Profile analysis to illustrate the performance of the regression and interpolation techniques in the Drakensberg (Profile along section AA' of Figure 6.7)

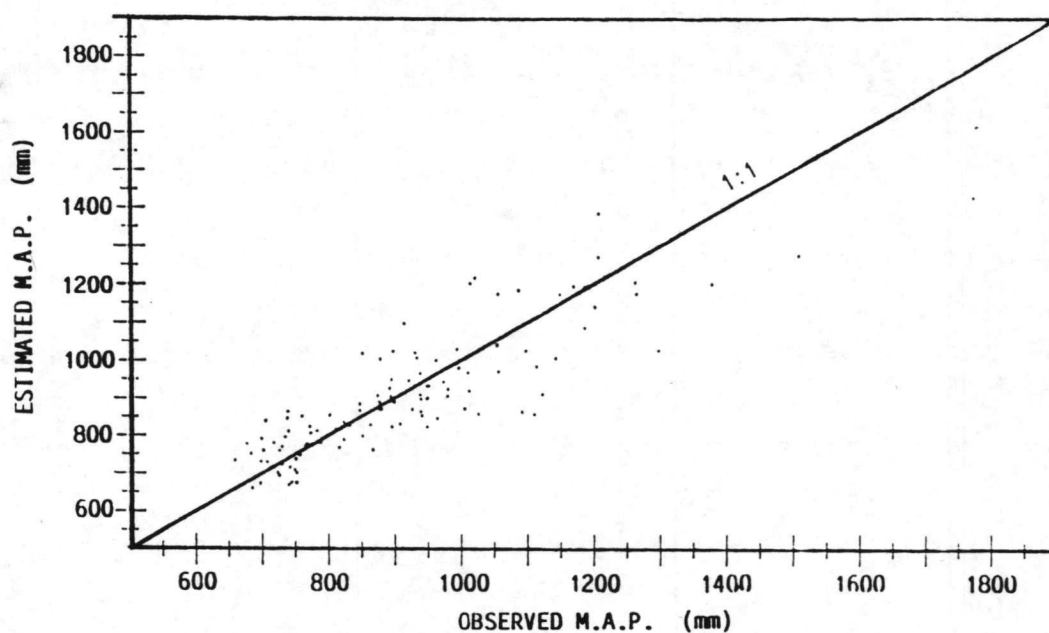


Figure 6.9 Observed versus estimated MAP using regression, in the Drakensberg area

It is evident from the results that each technique has its place. The direct interpolation technique performs adequately in the flatter OFS/TVL where the distribution of stations is good and no clear topographically based rainfall trends exist. However, the interpolation technique performs unsatisfactorily in the undulating and mountainous terrain of Natal where the raingauge network is sometimes sparse. This phenomenon is also illustrated in Figure 6.23 in which direct interpolation is applied in a mountainous area of the south western Cape. The ability of the trend surface analysis technique to account for a change in rainfall with a change in physiography means that the important effect of topography on rainfall is reflected more accurately in these areas.

The interpolation technique does, however, display a greater sensitivity to local anomalies in the regional trends of rainfall distribution because the generated rainfall surface is "forced" to fit the individual station points. There against multiple regression seeks a single expression for the best fit through all the points in a region. Inevitably the local sensitivity, in particular to "anomalous" changes in a rainfall surface, must suffer. Experience with potential users of these maps led to the conclusion that the surfaces had to "fit" at the station points in order to enjoy any form of credibility.

A compromise solution requiring the insertion of observed rainfall values at the station grid points and the rainfall values which were estimated via regression at all the other grid points was considered. Thereafter the contour interpolation routines are applied to the grid of point values. In this way it was hoped that the local sensitivity of the rainfall surface would be retained in the vicinity of stations and at the same time the regional trends with respect to factors such as topography would be preserved in areas which are not gauged.

The employment of the abovementioned technique was a move in the right direction but it was still not entirely satisfactory since irregularly scattered single points in the grid, i.e. the stations points, do not have much of an influence on the surface except at or very near the station point. The additional weighting of station points was considered as was the possibility of "expanding the influence" of the station point measurements by replacing the regression estimates adjacent to the station point by the station value. Uncertainty surrounding the spatial extent of this data replacement and the additional weight, if any, that such points should carry in the final surface, in addition to the fact that such a technique would bring in an "insensitivity to topography", led to the rejection of these ideas. However, another solution along similar lines was adopted and is discussed in Section 6.3.

6.2 Confidence of estimates and estimates of error in mapped values of MAP

At the commencement of this section it is necessary to clarify the difference between the confidence of the estimate of MAP and the error of the estimate of MAP. It should be realised that the daily rainfall amounts, from which the monthly and annual totals were aggregated, are only indices of rainfall in the first instance. In addition the estimate of the population MAP is only a statistic based on a sample of annual totals. The analysis presented in Section 4.1 and Appendix A shows that the sample estimate of the population mean improves with record length. These estimates of MAP are then used in regression equations to express MAP as a function of physiographic variables.

There is a probability distribution associated with the magnitude of the response variable for each predictor vari-

able value in a linear regression expression, this phenomenon being illustrated in Figure 6.10.

In order to obtain an estimate of the confidence limits of the estimate of MAP produced by least squares regression, the technique presented by Neter, Wasserman and Kutner (1983) was employed. An example of the results of such an analysis is presented in Figure 6.12. The MAP surface to which Figure 6.12 refers is presented in Figure 6.11. These figures are windows on the MAP surface at scale 1:250 000 in the Babanango area in Natal.

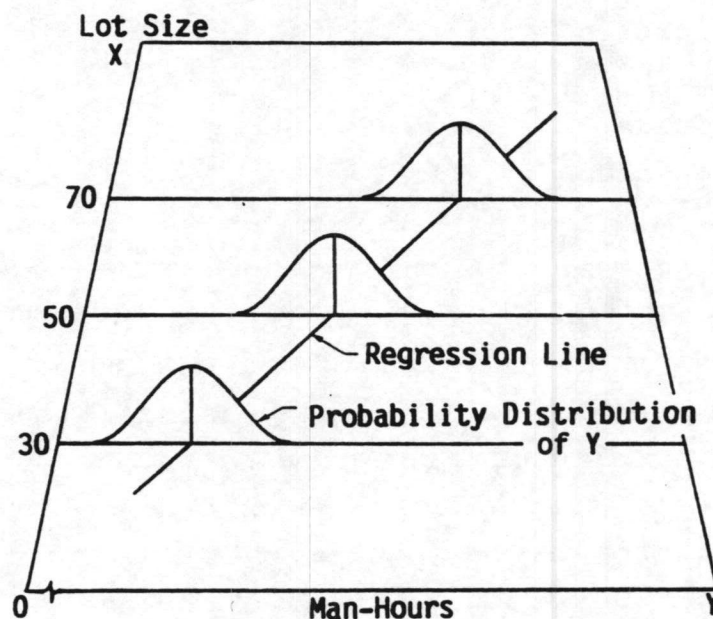


Figure 6.10 Pictorial representation of a linear regression model (Neter, Wasserman and Kutner, 1983)

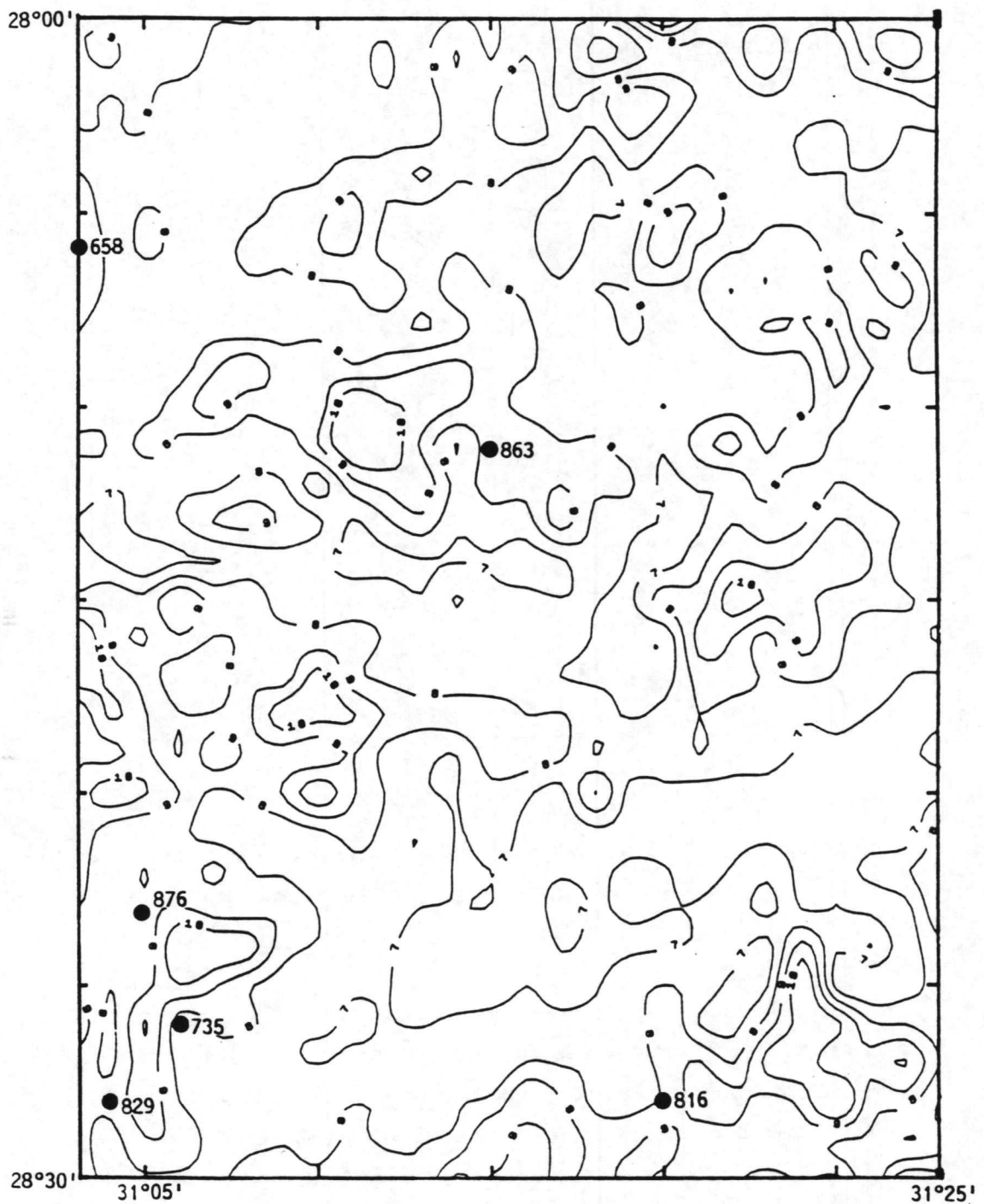


Figure 6.11 MAP in the Babanango area of Natal, predicted by the least squares regression technique (units: mm x 100)

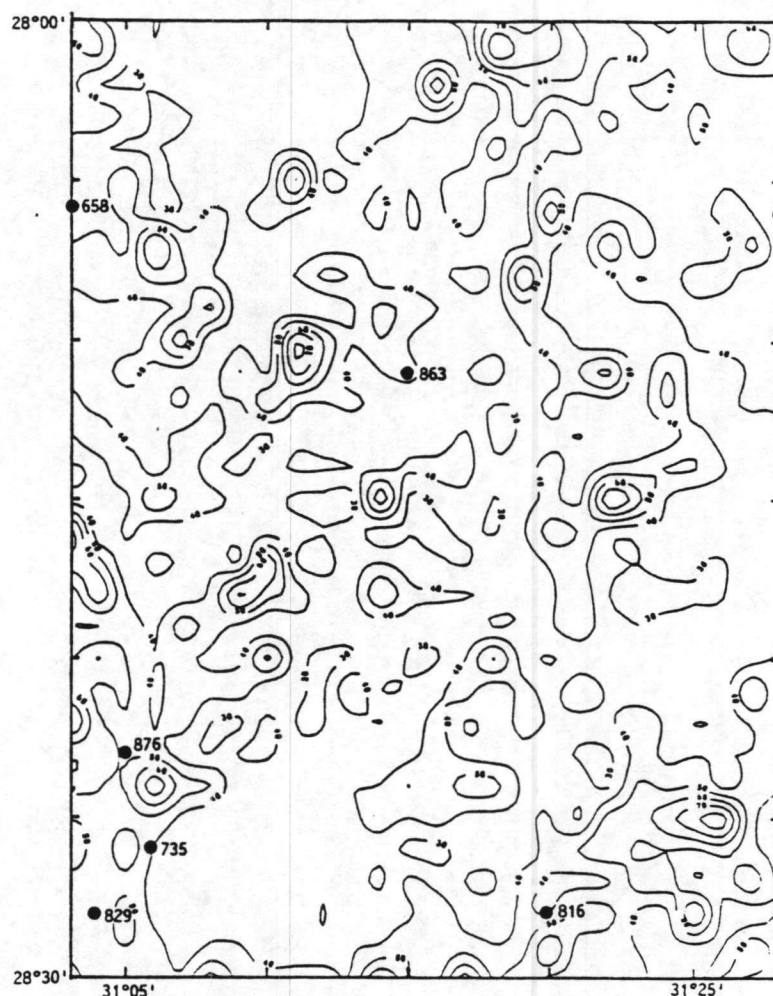


Figure 6.12 The 90 per cent confidence limits of predicted MAP in the Babanango area of Natal, using least squares regression (Units: mm)

The confidence limits of a prediction is a useful technique to identify those areas in the predictor space which do not have response variables in the regression data set or areas where the regression surface does not fit the data well. Areas which fall outside the predictor space will have wide confidence limits associated with the estimates of MAP. This technique is better than the straightforward interpolated surface of residuals since it is sensitive to the confidence of prediction over the whole area and not simply at the station points.

The above paragraphs discussed the confidence of the estimate of MAP. The error of the estimate of MAP, also known as the residual, on the other hand, should only be calculated at the station points which had sufficient record to be included in the regression data set. Ideally the residuals should be small with positive and negative residuals scattered about randomly in space. Clusters of residuals of similar sign and large magnitude indicate model failure. If a direct interpolation is conducted to distribute the station point residuals onto a regular grid, then areas where the regression model is estimating too high or too low are depicted. The technique which was finally employed used this residual surface to modify the regression surface which is sensitive to topography and form a composite regression/residual surface. The confidence limit estimates discussed earlier are no longer valid for this composite surface. The confidence limits of the prediction technique does, however, still provide some indication of the location and magnitude of uncertainty in the precipitation surface.

It is pertinent at this stage to introduce the concept commonly referred to as "error" maps and the practical implementation of the information contained therein. There is no doubt that a knowledge of the quality of an estimate enhances the usefulness of that estimate. With this in mind it was evident that there was a need for some form of "error" map for the MAP statistics. However, having produced such an "error" map it is relevant to question just how this map would be used in practice. The non-scientific user would probably not understand the full significance of an "error" map and therefore be confused by the inclusion of such a map and in addition not know how to implement the information contained therein.

The "error" map may either be a map of confidence limits of the estimate or alternatively it could be the residual map.

There is a great deal of confusion and ignorance, even among scientists, on this point and it was for this reason that the earlier paragraphs were devoted to this issue.

In order to convey the "error" information on the isohyetal maps some form of shading according to class intervals of errors would be required. It must be presumed that the informed user would wish to use such shading to interpret more fully or even adjust for possible errors in isohyetal values.

Preliminary versions of the MAP maps which were based solely on regression were sent to potential users for perusal. Almost without exception these users were concerned about the fact that the isohyets did not "fit" at the station points. This is not surprising as the regression surface is meant to produce a best fit over the whole region and will not necessarily pass exactly through any of the data points used in the regression. However, this is a fundamental yet subtle statistical point and one which will not easily be accepted or understood by the observers at approximately 10 000 rainfall stations around the sub-continent, especially when they know that their data were used in order to produce the isohyetal maps.

To summarise then, "error" maps at a scale of 1:250 000 were considered to be too expensive and time consuming to produce and it was thought that they would not be correctly interpreted. The types of "error" maps which could be produced are of the confidence interval type as shown in Figure 6.12, or interpolated maps of the residuals. The latter are created by calculating the residual between the regression estimate and the measured value at the station points and then employing interpolation techniques in order to map these randomly spaced residuals onto a regular grid before producing contour lines of residuals. Only stations with rainfall records above the predetermined length would

be included in the residual analysis.

6.3 The composite regression/residual surface

In view of the discussion in the above Sections, it was considered preferable to adjust the MAP surface according to the residual map at the research mapping stage and then simply present the final estimate of the isohyetal map. A great deal of additional cost and time in the final production of these maps was thus saved. After consultations with Adamson (1986) and Ryan (1986), the solution which was finally adopted involved the production of a regression surface, on the 1 minute grid, in the normal way. The percentage residual at each station was then calculated and an estimate of these percentage residuals was produced, at each grid point, by direct interpolation. Each grid point on the regression surface was then multiplied by one minus the corresponding grid point value on the percentage residual surface in order to produce a new surface which reflected the general shape and sensitivity of the MAP surface to topography, as determined by the regression surface and in addition this new composite surface fitted the measured MAP exactly, at the station points. In this way the general regional trend in MAP was modified to provide local sensitivity to measured MAP values at the rainfall stations used in the analysis.

The profile presentation of the interpolation, regression and regression/residual combination in Figure 6.13 emphasises the merits of each method. We believe that the combined regression/residual technique, which was discussed above, is preferable to interpolation or regression alone, despite the fact that the confidence limit map is thus less meaningful. The "perfect" fit at the station points and the increased local sensitivity produced by the regression and residual mapping technique were considered to outweigh the abovementioned disadvantages. According to Ryan (1986)

and Adamson (1986), this technique is acceptable, as long as the technique is reported clearly and no attempt is made to present the "final" map as the product of trend surface analysis by multiple regression only.

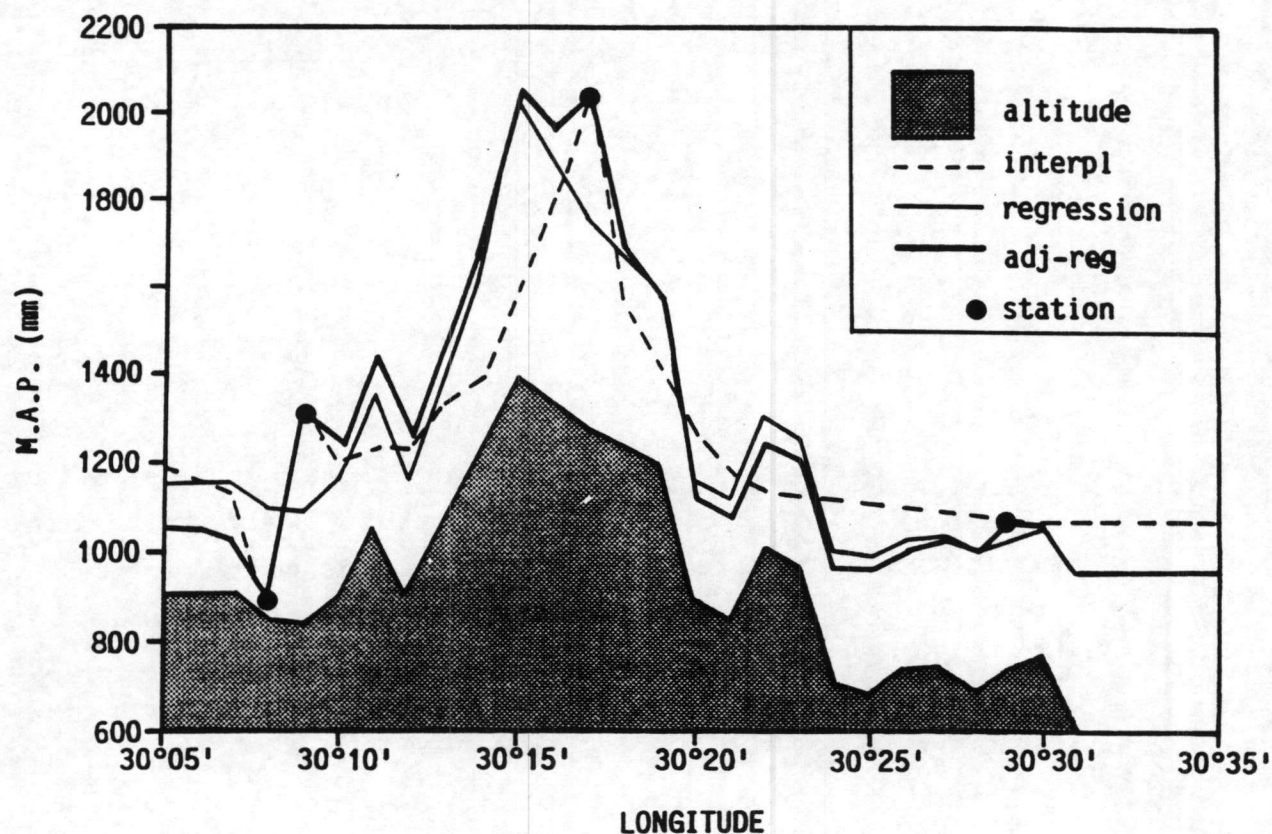


Figure 6.13 Profile of MAP surfaces along the section AA' in Figure 6.14 (Study area: Venda)

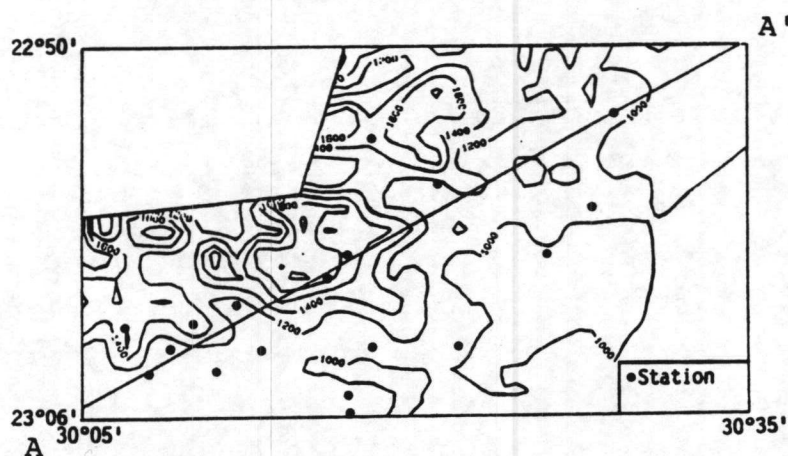


Figure 6.14 Isohyets of MAP, obtained by multiple regression only (Study area: Venda)

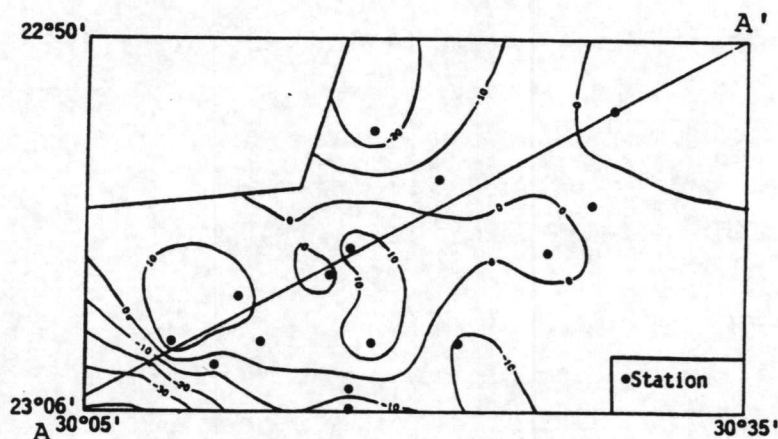


Figure 6.15 Residual surface of MAP (Study area: Venda)

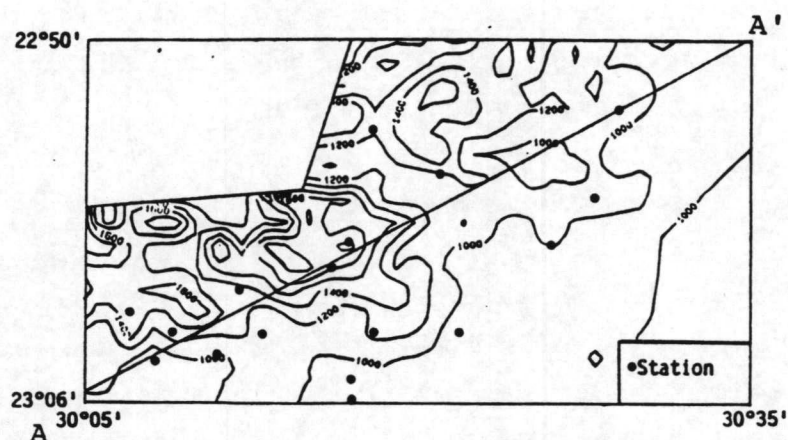


Figure 6.16 Regression surface adjusted according to the residual surface (Study area: Venda)

The employment of the technique of combining the regression and residual surfaces means that strictly speaking it is no longer possible to establish the confidence of the estimate. It is, however, possible to gain reasonable insight into the confidence of the estimate. A digital image depicting the confidence of the estimate of MAP from regression was therefore produced on the one minute grid. A further way in which the user may gain information on the "goodness of fit" is through the length of record analyses presented in Appendix A. The map fits the surface of station MAP estimates exactly. The user may therefore gain

some idea of the confidence of the estimate of MAP in the vicinity of the station by consulting Appendix A and, at the same time, establishing the length of record which was used at the station. The latter may be obtained from the file INDEX referred to in Chapter 3.

McKay (1976) states that it is most important that all aspects of the data, the quality and statistical character be thoroughly understood before mapping proceeds. Furthermore, the resulting map or supporting text should identify clearly the nature of the data, its quality and the period of records so that the user will not assign greater accuracy to his interpretation than is warranted. It may be seen that every endeavour was made to attain these goals, so aptly expressed by McKay (1976).

6.4 Selection of stations for use in mapping

The increase in the application of water balance techniques on a regional basis has stimulated interest in the extrapolation in space of point measurements of rainfall. Point measurements of precipitation which were formerly accepted as accurate absolute values are now generally recognised as indices (McKay, 1976). The emphasis has, therefore, shifted from focusing on accurate point estimate to a recognition that some degeneration in the quality of the estimate of rainfall at stations is more than compensated for by the effect of this rainfall station in the spatial estimation of the rainfall statistic, for example MAP. The effect of inclusion of a single rainfall station in a data set may have a significant effect on the regression function for MAP or the interpolation result. Stations which are at a higher altitude or more isolated than the rest of the rainfall data set in that area, are particularly important in this respect.

Boughton (1981) made the observation that in mountainous

terrain there is, generally, a bias towards location of raingauges at lower altitudes where rainfall variability is usually less pronounced. He attributed this bias to the agricultural development of flatter country and the easier access to raingauges. Boughton (1981) concluded that although no complete generalisations could be made, it was likely that there would be a tendency to underestimate catchment mean rainfall in mountainous areas. In this project, the regression technique used which relates MAP inter alia to altitude, addresses this problem. However, if there are no high altitude stations the trend surface analysis based on multiple regression cannot be applied with confidence in the high altitude areas.

Studies were conducted, region by region, to establish the "representativeness" of the altitude of rainfall stations which were used in the regression with respect to grid point altitudes. An example of one of these analyses is presented in Figure 6.17 in the form of cumulative frequency plots of rainfall station altitudes and grid point altitudes. These plots were consulted along with the "length of record" versus "quality of mean tables" presented in Appendix A and discussed in Section 4.1 when stations were being considered for admission to the regression data sets.

As the mapping exercise progressed the need for subjective judgement on the question of rainfall station entry into the data set for regression analysis became evident. The estimation of the errors of interpolation at confidence levels of 80 and 90 per cent revealed that it was preferable to use a short term record in some areas than to interpolate from remote points or estimates based on regression, especially when the physiographic attributes of the station in question lay outside the bounds of the physiographic attributes of any other stations in the

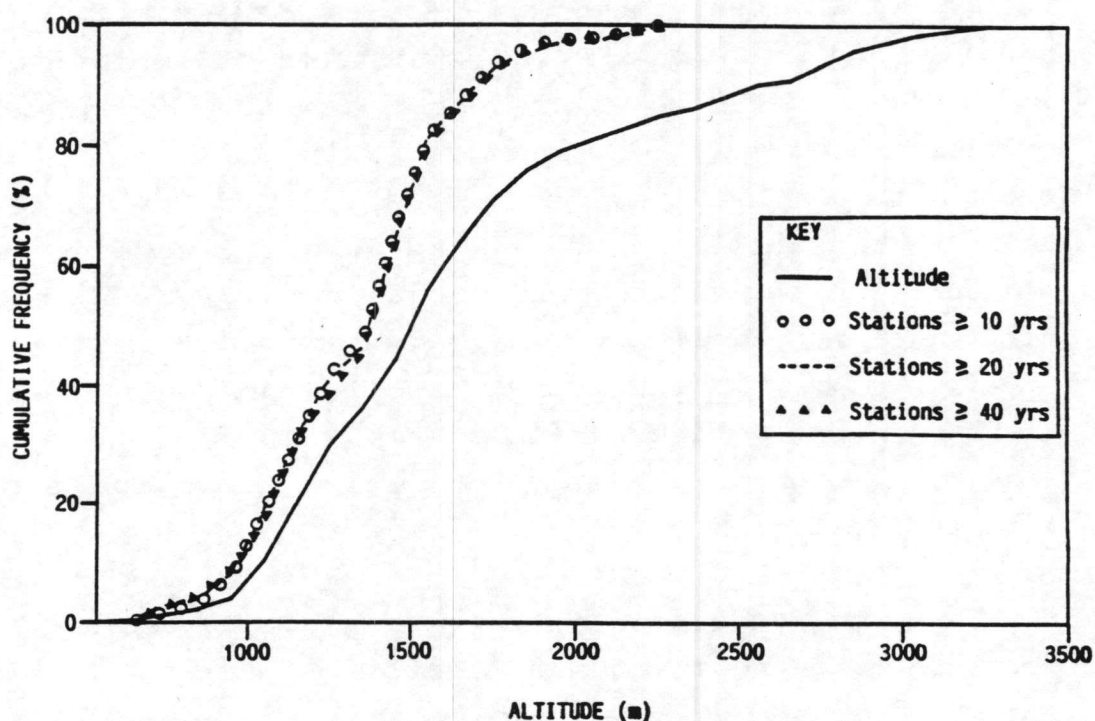


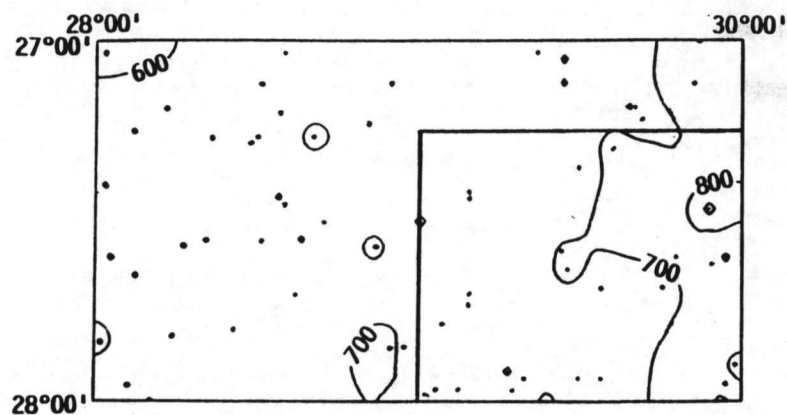
Figure 6.17 Cumulative frequency distributions of rainfall station altitude and grid point altitude for record lengths of more than 10, 20 and 40 years in the Natal Drakensberg

regression data set. The record length criterion was relaxed to allow stations to enter the data set after due consideration of the error estimates in the vicinity of the station, the error estimate of the mean for that record length (Appendix A) and the spatial and altitudinal situation of the station in relation to nearby stations.

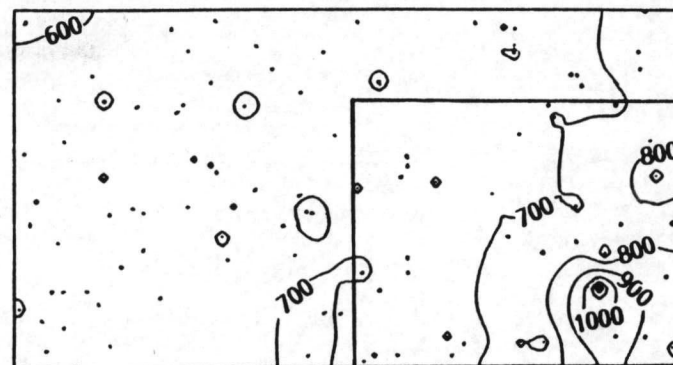
In the absence of a formal report on the methods employed to produce the 1:250 000 MAP maps currently in use it remains to speculate on the research methods used to obtain these isohyets. Figure 6.18 shows clearly that the 1:250 000 MAP maps currently in use were based on data from stations with less than 40 years' record at the time. It

is believed widely that the producers of the existing isohyetal maps employed rescaling or normalisation to a base period of 40 years viz. 1921-1960 to obtain estimates of MAP. In addition a decision must have been taken regarding the effect of altitude on MAP. The following example in the Frankfort/Standerton/Volksrust area serves to illustrate this observation. A study of the topographic and isohyetal map (Figure 6.18) of the area confirms the strong association which must have been established between MAP and altitude. It is evident from the rainfall record lengths noted in Figure 6.18 that the existing MAP map in this area could not have had any mathematical basis, if only stations with more than 40 years of record were used.

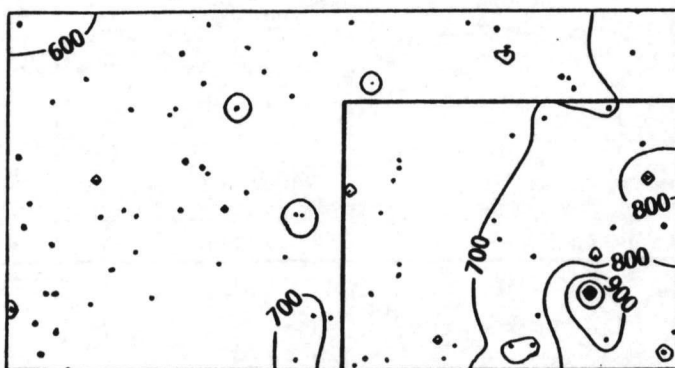
It is also apparent from Figures 6.19 (a - d) that the "higher" rainfall areas coincident with the range of hills in the east of this study area are not represented realistically by the stations with more than 40 years of record. Clearly some relaxation of the 40 year criterion is necessary in areas such as these with complex physiography. A relaxation of the 40 year restraint may not degrade the estimate of the mean by too much, whereas the exclusion of some vital stations, such as the ones shown in these examples, will certainly lead to isohyetal estimates of inferior quality. The decision whether to use a 40, 30, 20 or even 15 year record length for the rainfall, therefore, is largely subjective. In this project it was decided that the spatial and altitudinal distribution of stations was of prime importance and hence the record length criterion was often relaxed to achieve a better spatial distribution.



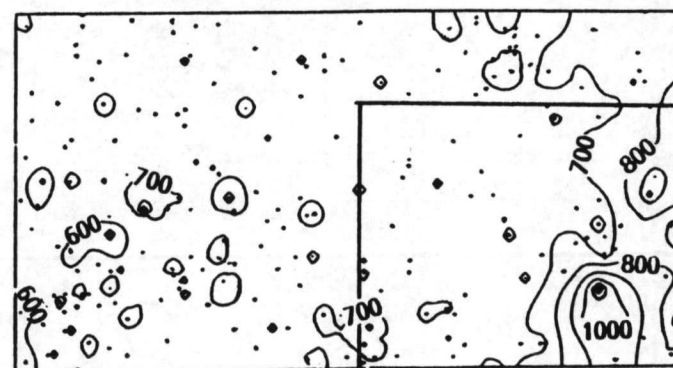
6.19 (a) Station records >40 years



6.19 (c) Station records >20 years
(as of 1984)



6.19 (b) Station records >30 years
(as of 1984)



6.19 (d) Station records >1 year
(as of 1984)

 Outline of area shown in Figure 6.18

Figure 6.19 Isohyets of MAP in the Frankfort area, drawn using polynomial interpolation

6.5 Example of the procedure for generating a MAP surface

A region in the south western Cape was selected to serve as an example of the MAP surface generation procedure. This area is a particularly complex topographic area and it was anticipated that the MAP surface would be equally complex. The location of this region is shown in Figure 6.20.

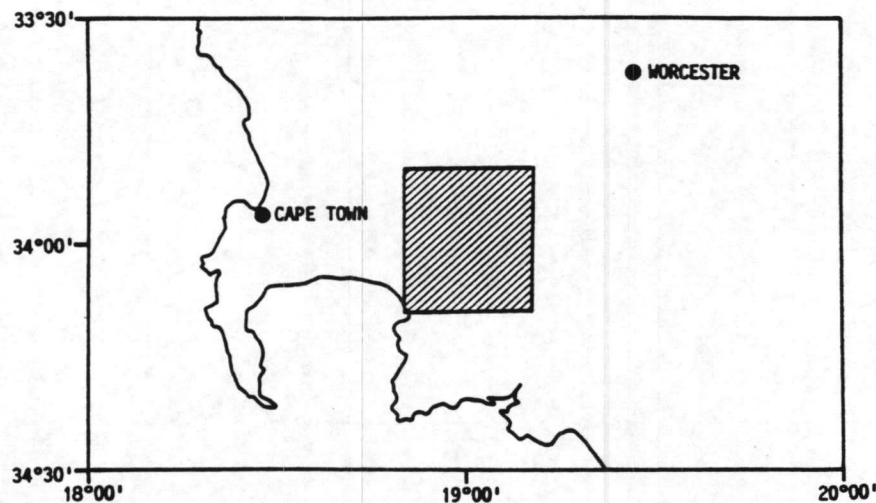


Figure 6.20 Location of the south western Cape region used in this study

Owing to problems of scales all the subsequent Figures presented in this example will display only a small window on the study area shown in Figure 6.20. The three dimensional plot of part of the study area shown in Figure 6.21 illustrates the complex nature of the terrain.

Only stations with 20 or more years of data were used in this analysis. This was in accordance with the length of record criteria discussed in Section 4.1 and presented in Table 4.1 and Appendix A. The locations of the 28 stations thus chosen and which are located in the abovementioned window are shown in Figure 6.22

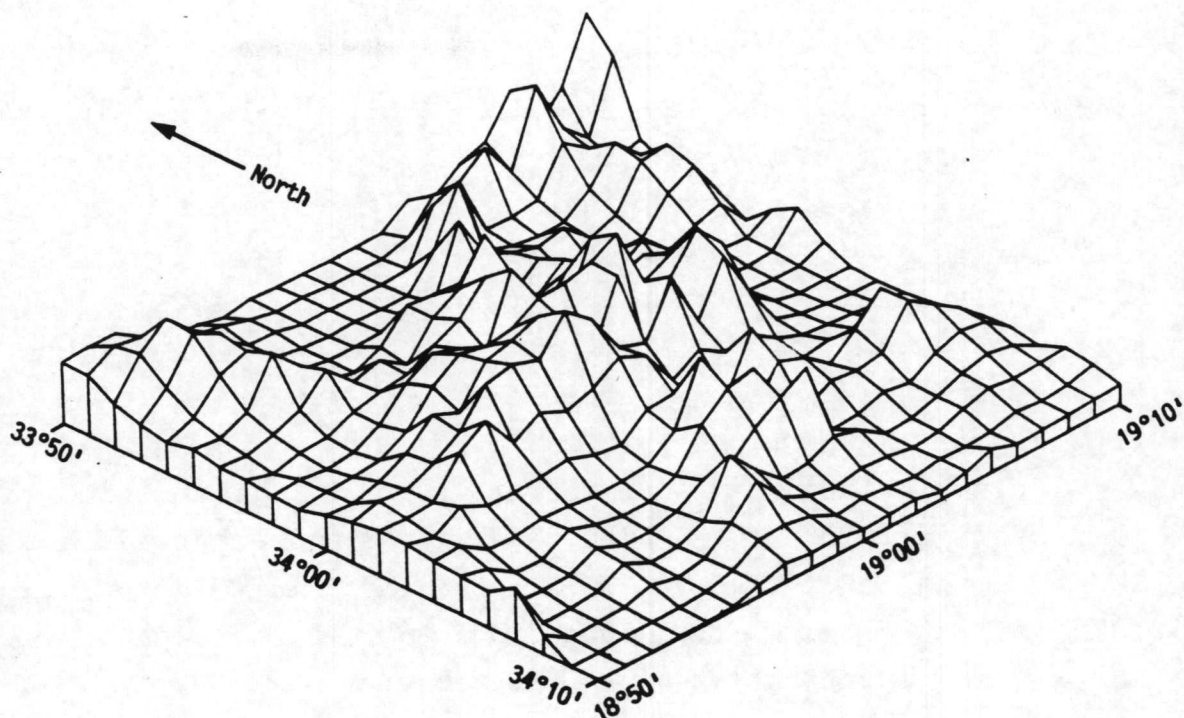


Figure 6.21 Three dimensional plot of altitudes in a part of the south western Cape region

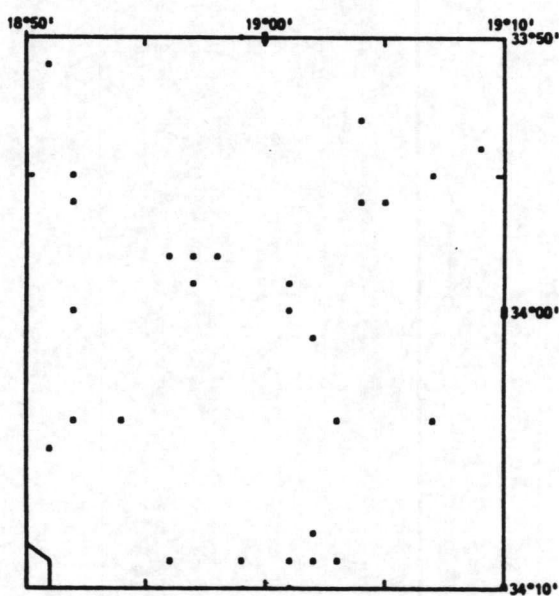


Figure 6.22 Location of the 28 stations (>20 years) used in this area for the analysis of MAP in the south western Cape region

The isohyets of MAP from direct interpolation of the station MAPs are shown in Figure 6.23. The isohyets show no sensitivity to the topographical changes and reflect only broad regional trends of MAP. This insensitivity is probably enhanced by the fact that most of the farm homesteads and hence rainfall stations are in the valleys where the rainfall is lower on the hills and mountains. The regression technique, by contrast, is not as reliant on a good distribution of stations with respect to altitude, since a few high altitude stations in a region can suffice for the regression technique, whereas with interpolation each piece of high ground must contain a rainfall station in order to gain representation in the isohyetal map.

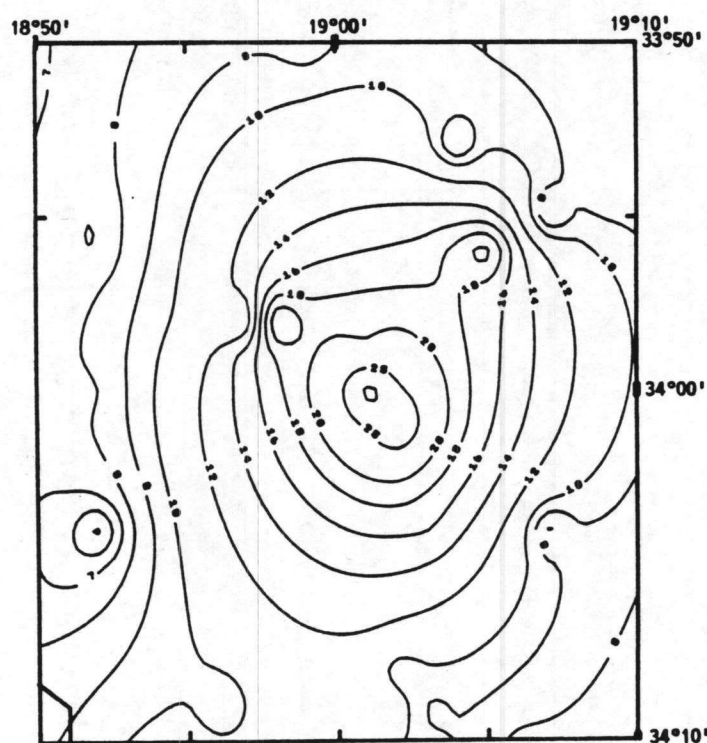


Figure 6.23 Isohyets of MAP by direct interpolation between the station MAPs in the south western Cape region (Units: mm x 100)

Most of the MAP surfaces which resulted from the regression technique showed a high degree of sensitivity to topographical changes in addition to other trends. This sensitivity is well illustrated in Figure 6.24.

After the MAP surface was generated by the regression analysis, the residual surface was generated by interpolation from the scatter of station point residuals. The residual surface is presented in Figure 6.25.

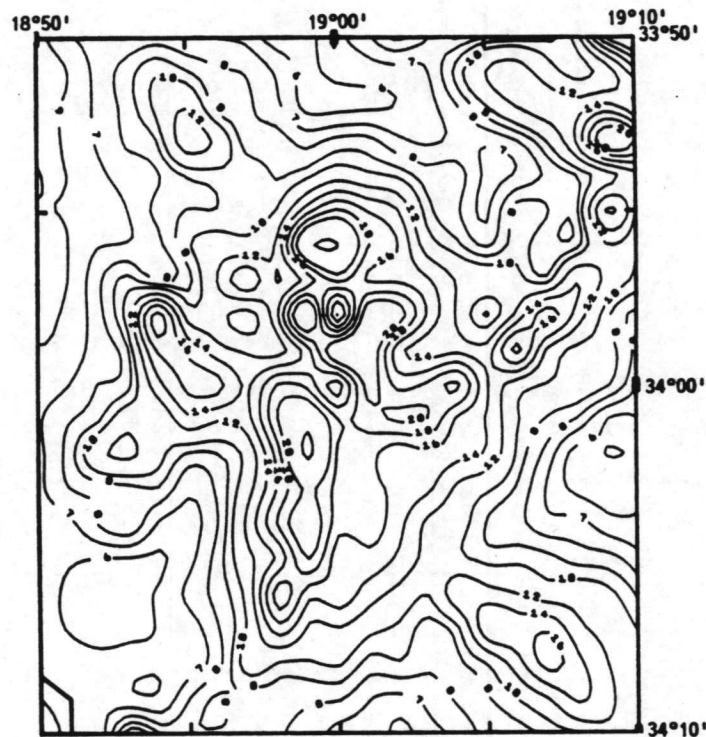


Figure 6.24 Isohyets of MAP resulting from the regression technique in the south western Cape region (Units: mm x 100)

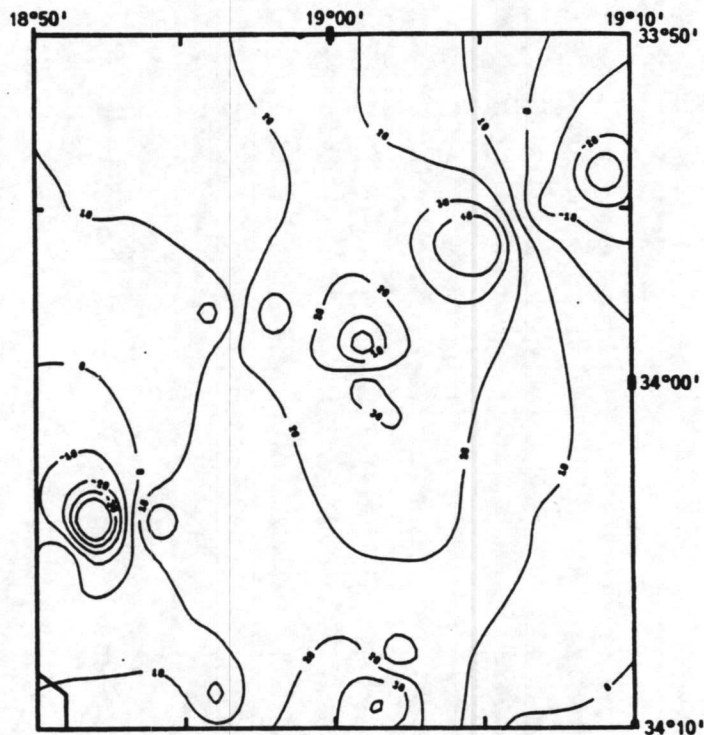


Figure 6.25 Residual surface by interpolation from the station point residuals in the south western Cape region (Units: percentage)

Finally, the technique described by Neter, Wasserman and Kutner (1983) and presented in Section 6.2 was employed to estimate the 90 per cent confidence limits of prediction. The surface of these confidence limits is shown in Figure 6.26. This surface of confidence limits was used to assist in the decision whether one or more stations at critical locations but with a record length greater than 20 years in the case of the south western Cape should be included in the regression analysis to improve the confidence of the estimate in certain areas.

The residual surface and the surface resulting from regression (Figure 6.24) were combined as described in Section 6.3.

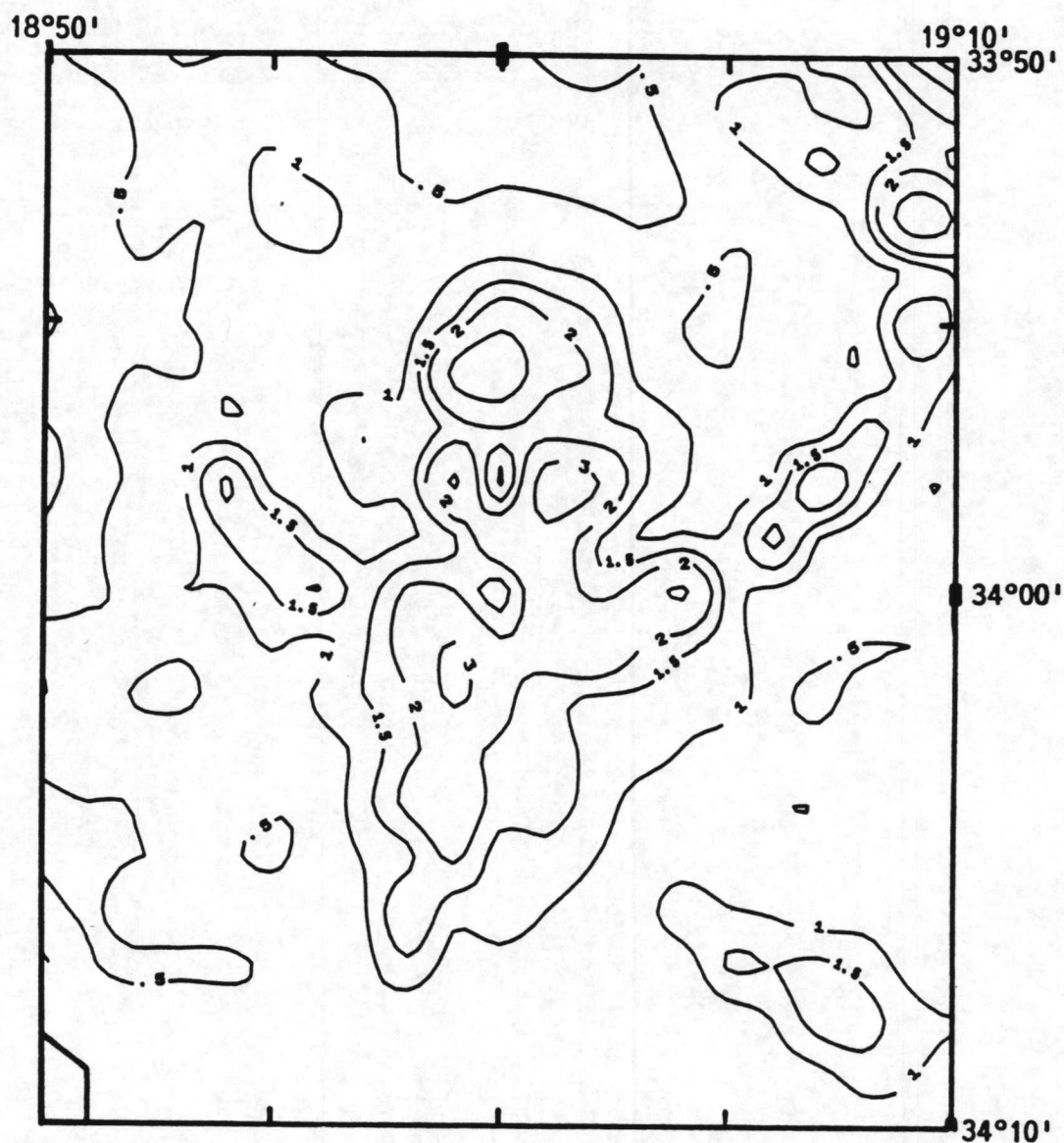


Figure 6.26 The 90 per cent confidence limits of prediction of MAP using least squares regression, expressed as a percentage of the predicted value in the south western Cape region (Units: mm x 100)

The product is a surface (Figure 6.27) which follows the trend exhibited by the regression surface but also fits the observed MAP exactly at the rainfall station points which were used in the analysis. The resultant surface therefore has both the predictive capability of regression analysis in ungauged areas as well as the local sensitivity of direct interpolation in the vicinity of rainfall stations.

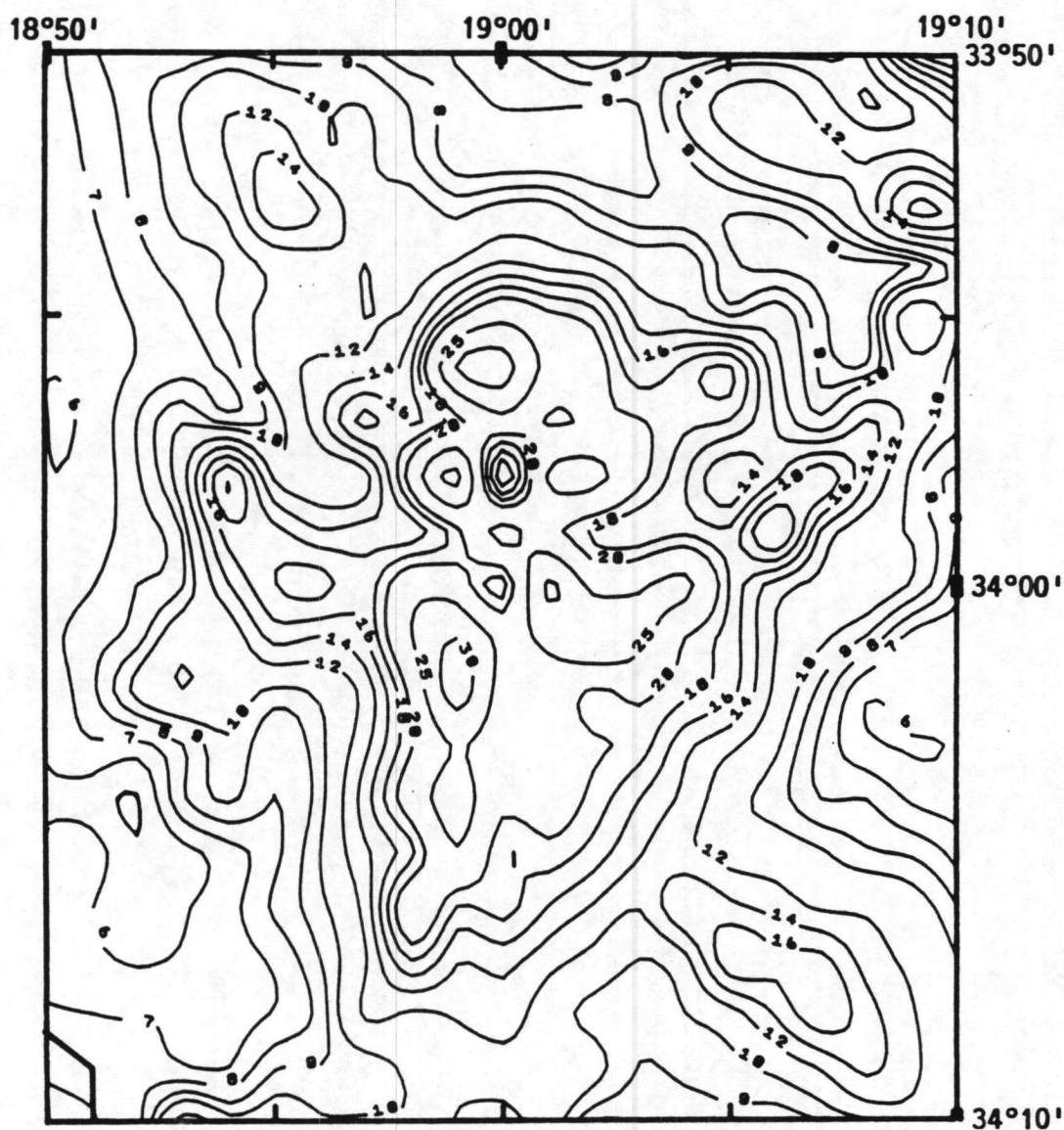


Figure 6.27 Regression and residual surfaces of MAP combined in the south western Cape region
(Units: mm x 100)

6.6 Mapping mean monthly and other rainfall statistics

The surface of MAP provided the foundation for the mapping of mean monthly precipitation and the variability of monthly and annual precipitation. The amount of work involved in producing separate regression expressions for each of the monthly surfaces for the abovementioned statistics was prohibitive. Alternative techniques were therefore sought.

For mean monthly rainfall a technique which established the ratio between the mean monthly rainfall and the MAP at each station point was chosen. These ratios were interpolated onto a regular grid (1' x 1') using SACLANT and then multiplied by the MAP image. The resultant image was one of mean monthly rainfall which fitted the observed data exactly at station points and followed the general shape of the MAP surface in areas between stations.

The important aspect of sensitivity to local differences in the ratio of mean monthly rainfall to MAP was retained by this technique. Mean monthly rainfall is highly correlated with MAP and therefore the application of a fixed ratio in each region would probably have been considered acceptable had it not been for the refinement introduced by the interpolation technique.

The variability of monthly and annual rainfall is not as straightforward to map since the variability is related inversely to mean monthly rainfall and to MAP. An analysis was performed on each of the mean monthly and MAP data sets in each region to regress the coefficient of variability (CV) of monthly and annual rainfall on the mean monthly rainfall and MAP. The residuals between the observed and estimated CV surfaces were then calculated and the regressed surface of CV was adjusted in the same manner as the MAP surface described in Section 6.3.

CHAPTER 7

HOMOGENEOUS RAINFALL REGIONS

The primary reason for regionalisation was to provide a logical, physically based division of the rainfall stations so that the regression analyses could proceed using areas which were both manageable and meaningful. The meteorological systems which generate rainfall over Southern Africa vary from one part of sub-continent to another. It is therefore useful to delimit regions of homogeneous precipitation in order to establish the factors and processes governing MAP and other precipitation statistics separately, for each region, thereby allowing for different regression expressions or even techniques to be applied in various regions and for greater accuracy of estimation of rainfall surfaces within a region. At the same time it was also appropriate that a re-evaluation of the distribution of homogeneous precipitation regions be conducted in conjunction with this precipitation mapping project.

7.1 Existing delimitations of rainfall regions

Southern Africa has been classified into various rainfall regions at different times in the past. The most notable of these regionalisations have been the South African Weather Bureau (SAWB) (1960) and SAWB (1972) classifications and that by Welding and Havenga (1974). After careful consideration of the techniques employed in the aforementioned classifications it was decided, initially, to adopt the Welding and Havenga (1974) rainfall regions for use in this project.

The Welding and Havenga (1974) regions were preferred to those of the SAWB (1960, 1972) for several reasons. First, the Welding and Havenga (1974) regional classification used 1549 stations for the period 1931-1966, which is more than double the number used by the SAWB (1972) for the same period. Secondly, Welding and Havenga (1974) used monthly data in a correlation procedure which was followed by the hierarchical classification technique of McQuitty (1960). This technique is considered to be statistically more sensitive than that of the SAWB (1972), which used the distribution of mean monthly rainfall. Buys et al. (1979) warn that the mean monthly precipitation value masks too much information. Welding and Havenga (1974) took cognisance of topographic features, as did SAWB (1972). Welding and Havenga (1974) delimit 114 regions compared to the 34 of the SAWB (1972) and in Natal the Welding and Havenga (1974) regions show reasonable agreement with those proposed by Schulze (1983). The overall goal of the Welding and Havenga (1974) classification was to identify similar regions for agricultural purposes. It is therefore reasonable to assume that cognizance was taken of the situation "on the ground", as it were, and that it was not purely an exercise in which the correlations between monthly precipitation were the sole criterion in the delimitation. The statistics did, however, reveal that the stations whose monthly rainfall is well correlated tend to be clustered together in proximity to one another.

Unfortunately the Welding and Havenga (1974) regionalisation did not take cognizance of the MAP at stations since their classification was based purely on interstation monthly rainfall correlation. In addition the only map depicting the Welding and Havenga (1974) regions is that which appears in their publication and it is at a scale of 1:5 000 000.

7.2 Method of delimitation of homogeneous rainfall regions

The altitude image of South Africa and the techniques developed during this project provided an excellent opportunity to enhance considerably the resolution and detail in the delimitation of rainfall regions. Special emphasis was given to altitude and MAP in this exercise. Welding and Havenga (1974) had shown that the monthly rainfall totals at stations which are in close proximity, are usually well correlated. Early in the project it was considered to use groups of these more detailed proposed regions and assign them to one of the major regions in the Welding and Havenga (1974) classification and in this way refine the boundaries of the Welding and Havenga regions for use in the regression analysis and mapping. The inclusion of MAP, altitude and a coarse estimate of aspect (deduced from the altitude image) as criteria for classification, have made the resulting regions more relevant for agricultural and hydrological purposes since temperature and hence potential evapotranspiration can also be inferred from altitude and aspect (Linacre, 1977 and Schulze, 1983).

A classified image of altitude covering Southern Africa was printed at a resolution of 1 minute of a degree. Superimposed on this image were the locations of all stations with 10 or more years of record. The MAP at each station was noted at the station point as shown in Figure 7.1. Thereafter followed the careful, patient and very time consuming process of delimitation of homogeneous rainfall regions based, inter alia, on MAP, altitude and aspect whilst keeping the regions reasonably small and consulting 1:250 000 topographic maps in order to provide a more detailed interpretation of the classified altitude map. The following criteria were used in the division of rainfall region boundaries.

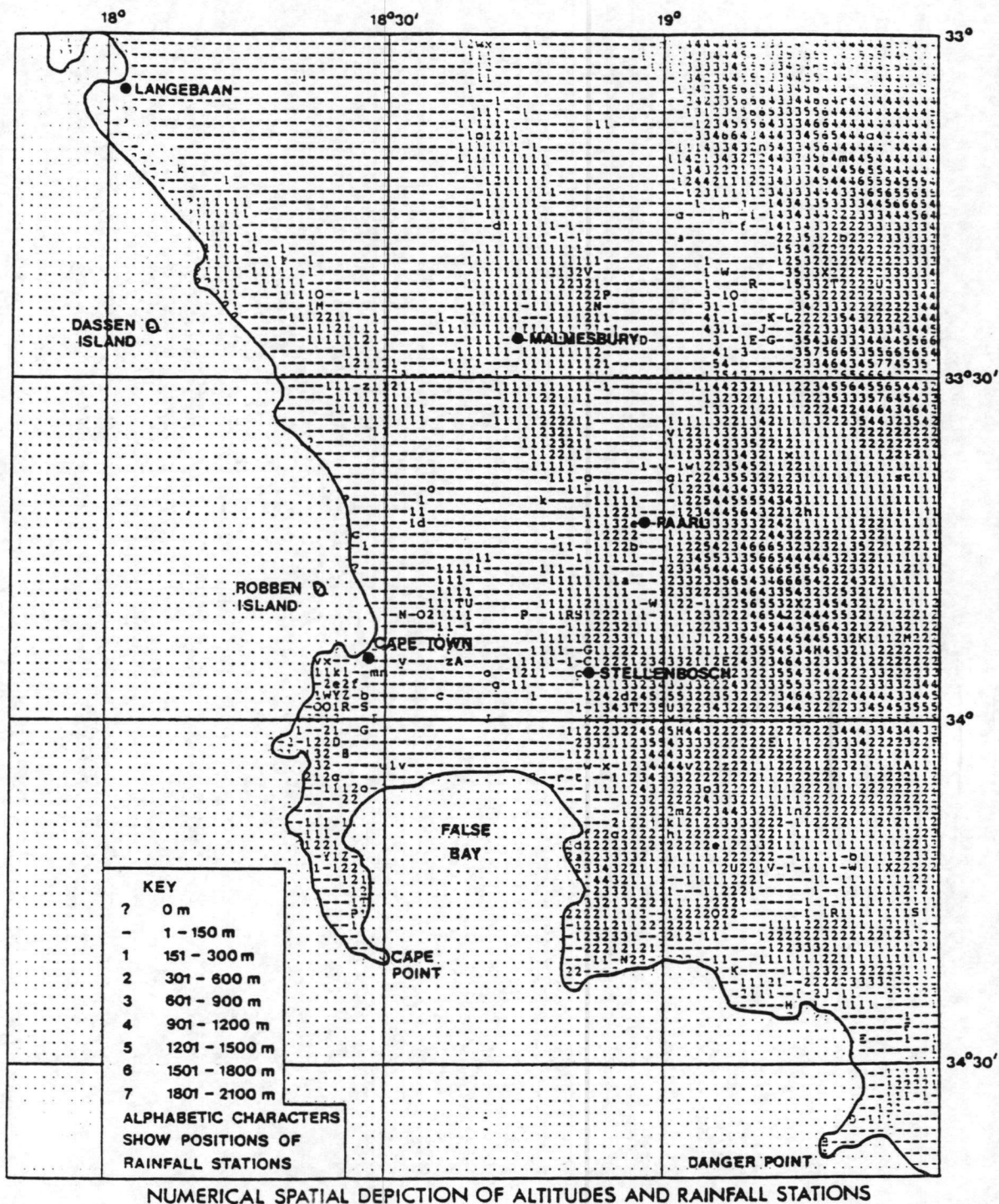


Figure 7.1 Classified image of altitude and the location of stations with >10 years of record for part of the south western Cape

- (a) Mean Annual Precipitation (MAP) :
Stations with 10 or more years of rainfall record were plotted along with their MAP. Boundaries were drawn so as to group stations of similar MAP and since regions were made reasonably small these were likely to have similar weather patterns and strong interstation monthly rainfall correlations. Most regions selected have a range in MAP of less than 100 mm.
- (b) Altitude :
Since rainfall and temperature both vary with altitude an effort was made to restrict the range of altitudes covered by any one region. Consequently, many of the boundary lines were drawn to follow the breaks in "contours" of the classified image of altitude.
- (c) Aspect :
The direction of movement of weather systems can have a marked influence on rainfall amounts, especially in mountainous areas. An attempt was made therefore to limit regions to the same side of mountain ranges and allowances were made for rainshadows.
- (d) Terrain :
In mountainous areas, regions were reduced in size to allow for more rapid spatial variations in rainfall, temperature and aspect. Regions tended to be larger in flat areas. The lack of rainfall stations in Lesotho made it an exception to this trend.
- (e) Agricultural activity :
In areas where the limited rainfall would curtail agricultural activity, regions were delimited larger, for example in the northern Cape.

In this manner, 712 regions were delineated. The boundaries of these regions were digitized to a resolution of 1 minute of a degree and stored for computerised plotting. An example of this detailed regionalisation is presented in Figure 7.2 and the complete map of regions in South Africa appears in Appendix D. It may be argued that 712 regions is too fine a delimitation. However, it was decided to proceed with the process of human pattern recognition based on the above criteria. Regions may always be grouped subsequently by the user for specific tasks.

7.3 Problems at regional boundaries

The identification of the region in which the station and grid points were to be placed was a task which had to be automated in view of the large number of stations and grid points in the analysis. A computer program was developed to perform this regionalisation. In addition it was imperative that a re-classification of stations and grid points be effected with ease, if the regional boundaries were altered for any reason. These boundaries had to remain flexible since the boundaries could be changed in the event of subsequent regression analysis revealing a need for adjustment.

An anticipated problem with regionalisation was that of continuity at the boundary of each region. The program therefore contained the facility to incorporate stations and grid points on the periphery of a region. Stations and grid points could thus be allocated to more than one region. The overlapping or border zone was necessary to assist in the smoothing processes involved in mapping precipitation at regional boundaries.

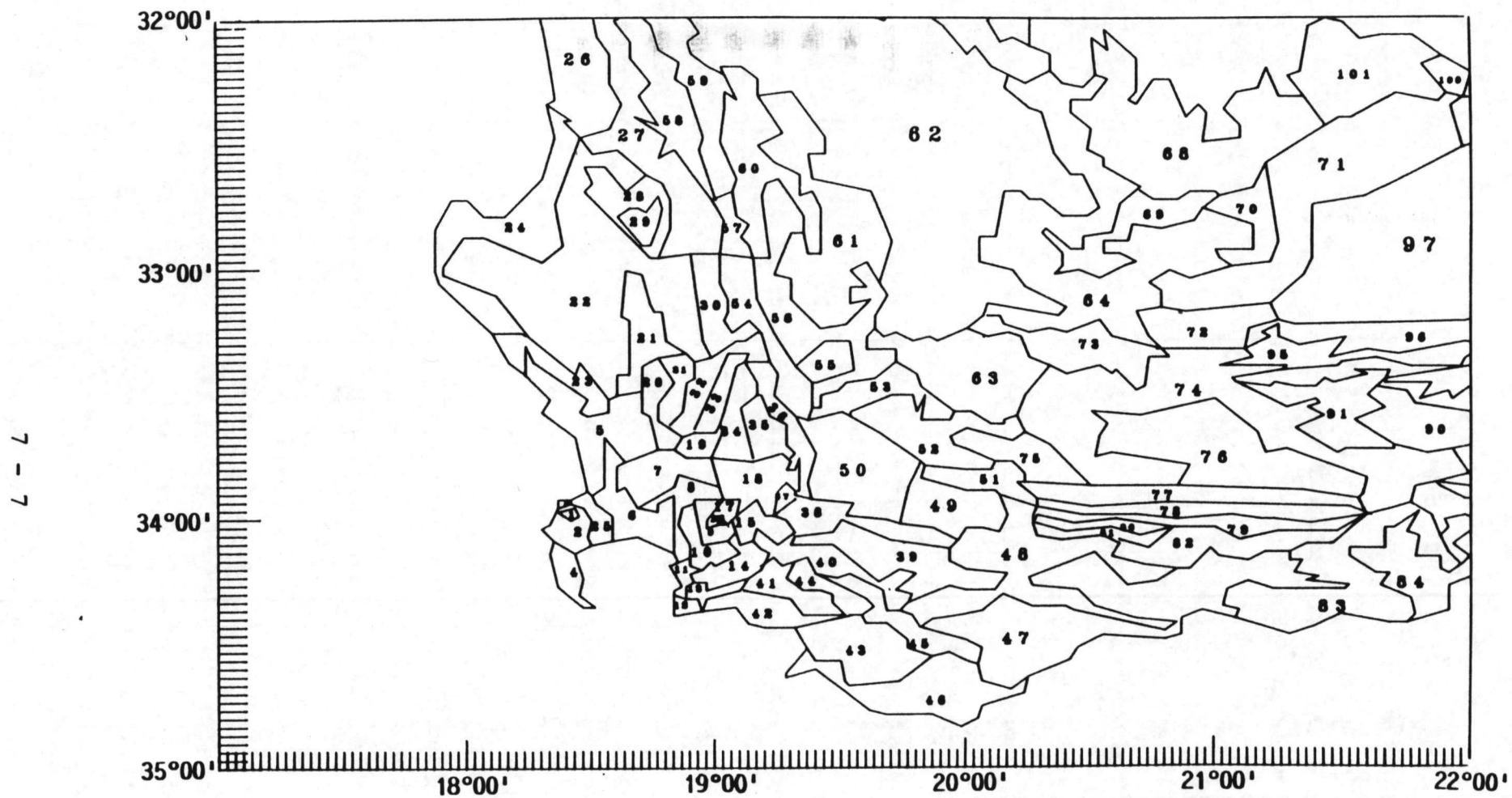


Figure 7.2 The detailed delimitation of rainfall regions in the south western Cape

As stated at the beginning of this chapter the primary reason for regionalisation was to provide a logical, physically based division of the rainfall stations so that the regression analyses could proceed using areas which were both manageable and meaningful. However, when it came to matching or merging the rainfall surfaces generated by the regression analyses at these regional boundaries, major problems were encountered. The employment of overlapping border zones did not produce the desired continuity of MAP surfaces at regional boundaries. At this stage the inadequacy of the resolution on the Welding and Havenga (1974) regional boundaries also became apparent. The vexing problem of obtaining a smooth transition of MAP estimates at regional boundaries received a great deal of attention. Eventually, it was found that fairly large regions should be used for the regression as shown in Figure 5.1. Where possible these regression regional boundaries were confined to areas in which there were no rapid changes anticipated in trends in MAP.

An example of the problem of merging isohyets at regional boundaries is shown schematically in Figure 7.3 for an area where the rainfall surface for the lowlands "meets" that for the highlands on top of the escarpment. In this example the range of altitudes of rainfall stations in the lowlands is not great. This also applies within the highlands. Therefore altitude was not strongly represented in the regression expressions for each of these regions. Consequently, neither of the expressions were able to generate the abrupt transition that was apparent at the escarpment.

The use of the arithmetic mean of the two surfaces at each grid point in order to join them was considered but then rejected as a possible solution because the topography criss-crossed the boundary line, as shown in Figure 7.3. This implied that those highlands which were in the lowland

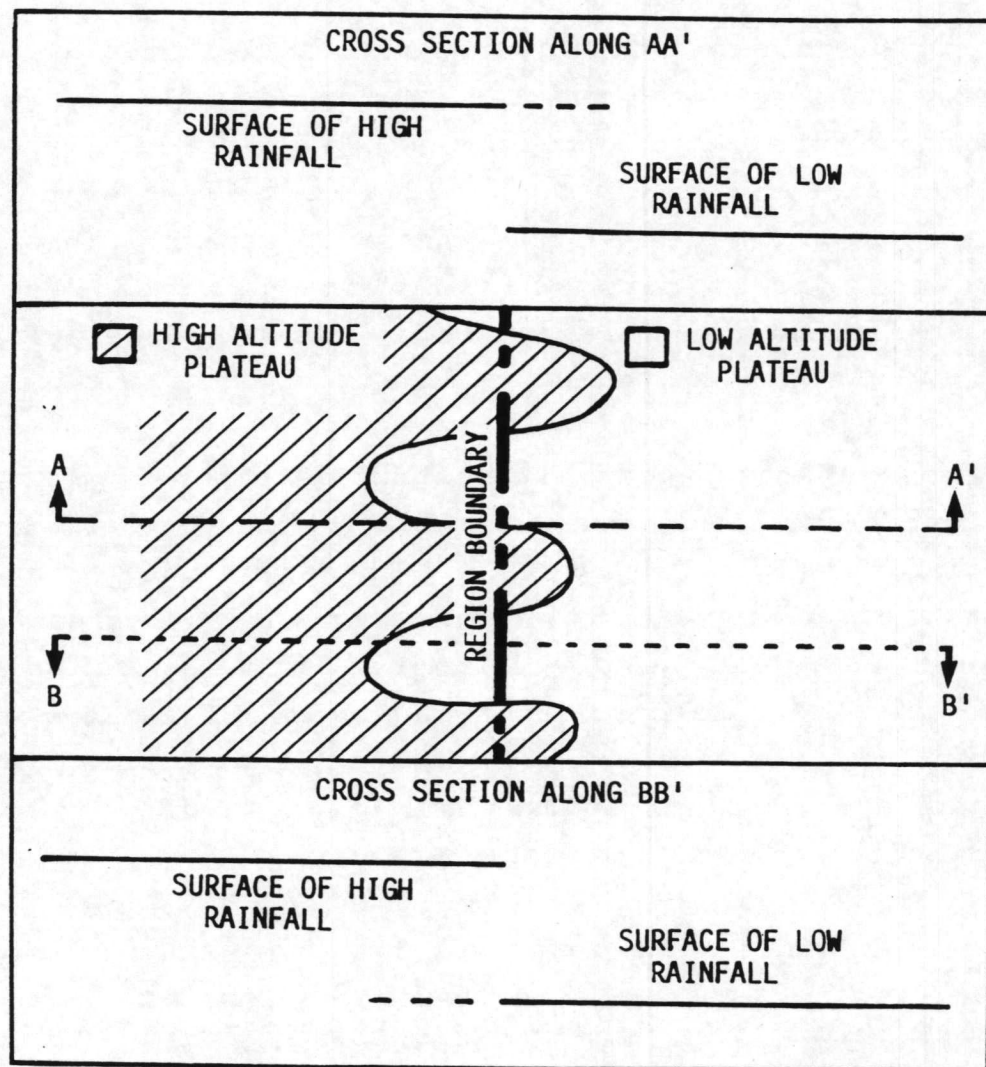


Figure 7.3 Illustrations of the problem of merging rainfall surfaces at regional boundaries

region because of the delimitation were more accurately represented by the highland rainfall surface, and vice versa. The results of a direct averaging technique would have been insensitive to this phenomenon.

To surmount the abovementioned problems, it was decided to use much larger regions and to ensure that the regional boundaries did not occur on a topographical divide but rather on sections of country where similar topography existed on both sides of the regional boundary. This system proved successful. The need for detailed regional delimitations for this aspect of the work therefore fell away, to a large extent.

In most of the regions it was necessary to try many different regression expressions before the boundaries merged satisfactorily. The process is, however, complex and time consuming since most regions are bounded by three to four other regions and all the boundaries have to merge smoothly. The example shown in Figure 7.3 and the example described below are typical of the type of problems encountered and of the solutions employed in order to provide continuity at regional boundaries.

Figure 7.4 shows the merger of two regions just inland of Mossel Bay in which the regression expression for the surface to the east of AA' had more emphasis placed on the altitude term than did the surface to the west of AA'. The distinct change in sensitivity of the isohyets to altitude is therefore evident in the east. A new regression expression in which altitude was less prominent, was chosen for the eastern section. This change in expression was achieved with only a negligible drop in the adjusted r^2 and the PRESS statistic and the two regions were merged satisfactorily as shown in Figure 7.5.

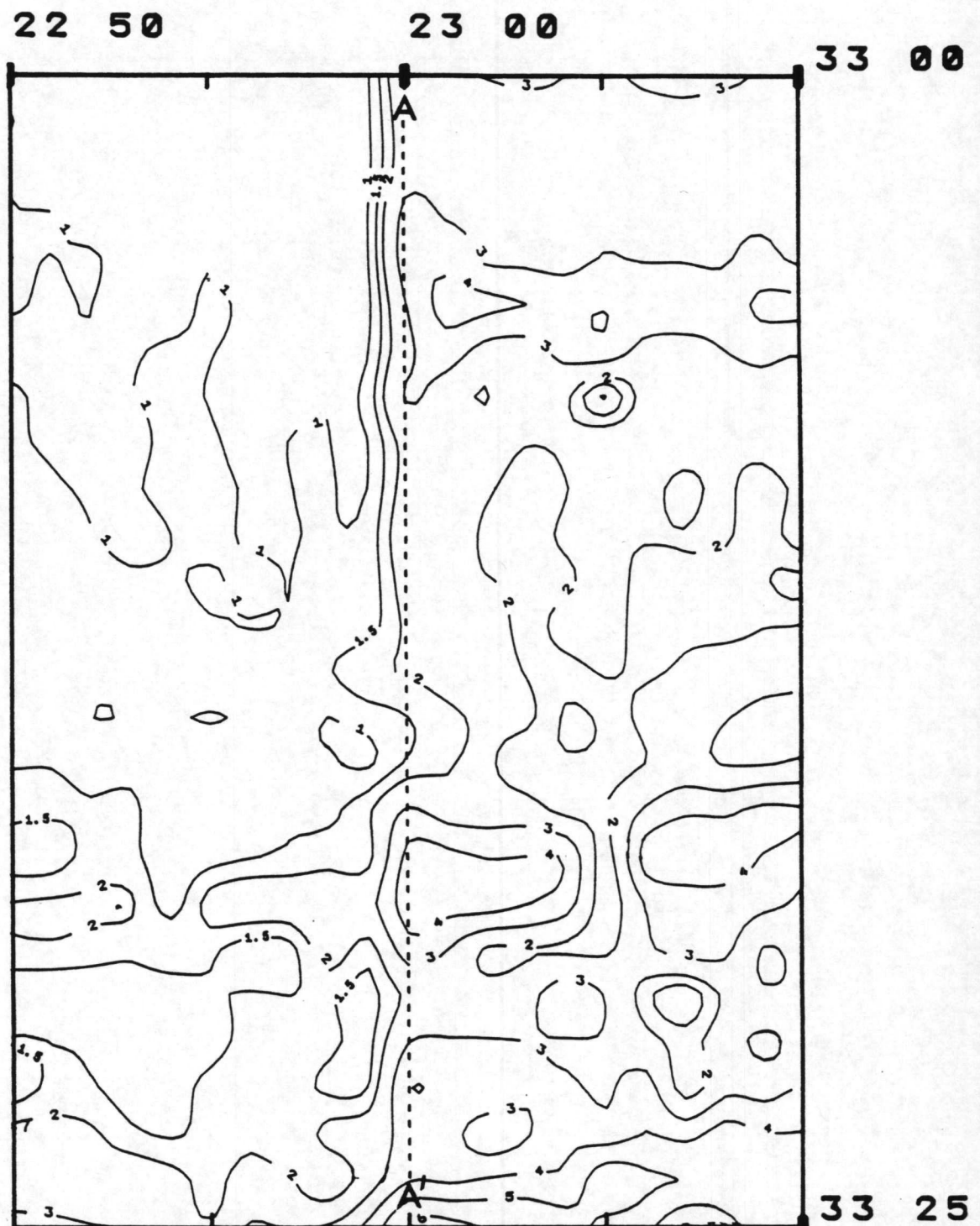


Figure 7.4 Merging of two surfaces of MAP in which the altitude term is more strongly represented in the region to the east of AA'

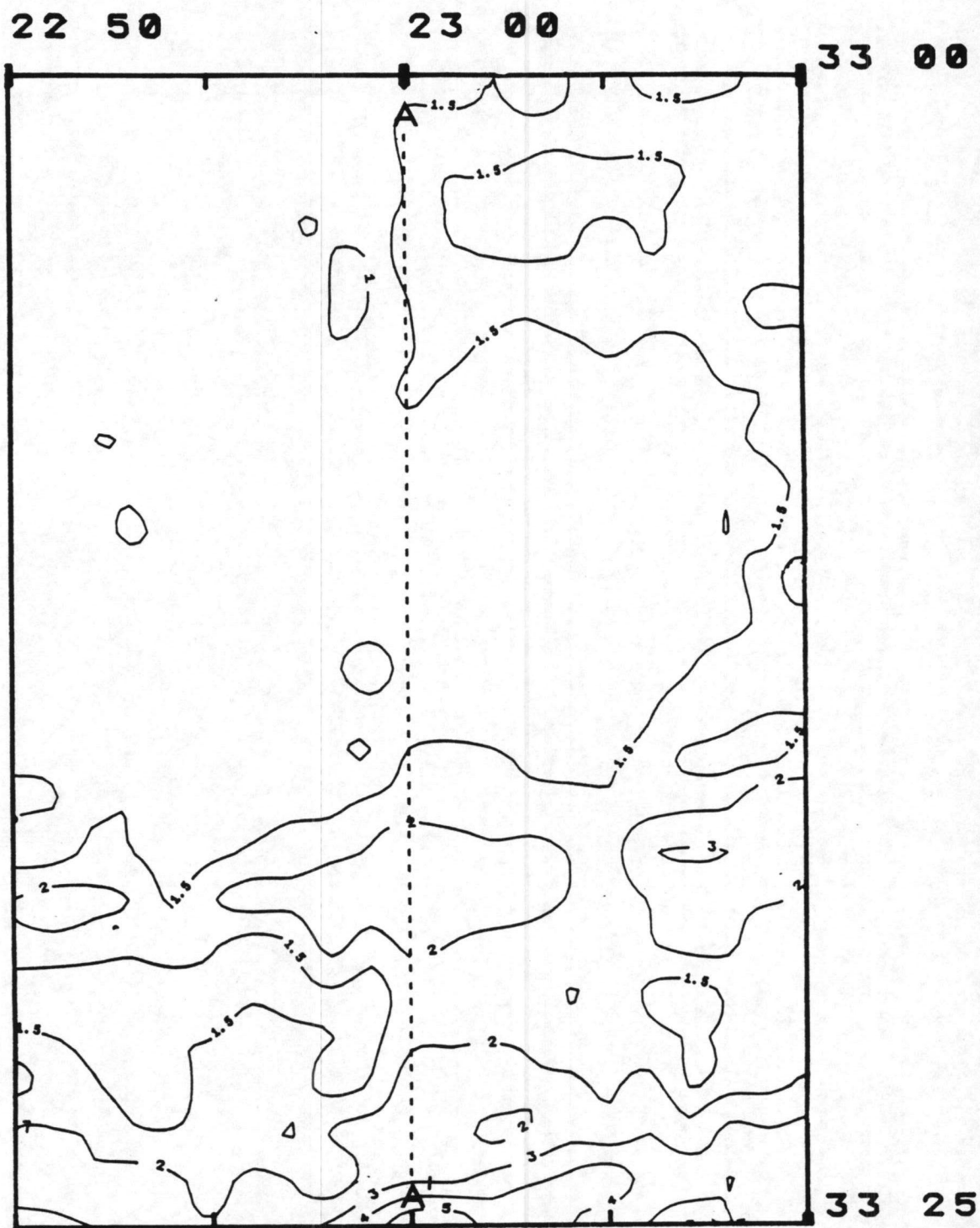


Figure 7.5 Continuity at the boundary after adjusting the altitude term in the region to the east of AA'

7.4 Conclusions

The re-evaluation of the distribution of homogeneous rainfall regions reported above fulfilled one of the major objectives of the project. However, these regions were not used in the regression modelling and final mapping exercise for the reasons outlined. It was always the intention to use these regions, not necessarily as individual regions but perhaps as groups of regions. The problem of merging MAP at regional boundaries and the large border zone that was necessary to achieve this made the detailed delimitation of regions insignificant for this particular aspect of the project. The regionalisation is however, useful for other applications, inter alia, for delimiting regions of equal agricultural potential and planning requirements, from the climatological point of view. These regions have been used by Dent, Schulze and Angus (1987) and Schmidt and Schulze (1987a) in a detailed analysis of soil moisture deficit for irrigation planning in Southern Africa and for the regionalisation of potential runoff depths and peak discharges for small catchments in Southern Africa respectively.

CHAPTER 8

DISCUSSION AND CONCLUSIONS

8.1 Merits of isohyetal maps vs digital images of rainfall

Isohyetal maps of MAP provide an efficient and effective means of communicating important and frequently applied information to a wide range of users. In addition, maps are inexpensive, portable and durable. However, printed isohyetal maps do have several disadvantages which are becoming more apparent as a result of advances in computer technology. Some of these disadvantages are;

- (a) the scale of the map is fixed,
- (b) the isohyetal interval is fixed,
- (c) it is time consuming to estimate catchment mean areal rainfall from these maps,
- (d) the capture of isohyetal information from maps and into computer compatible form is tedious,
- (e) the information content on the isohyetal map is fixed insofar as one cannot manipulate these data without first capturing them on computer.

Digital images of the isohyetal maps were created in the process of producing the maps. These images have been retained on computer tape and are available on request. The MAP and mean monthly precipitation maps as well as monthly and annual percentiles, skewness coefficients and coefficients of variability are also available in digital form at a spatial resolution of 1 minute of a degree. The digital images have a number of advantages, namely:

- (a) the images may be printed to maps at any desired scale,
- (b) the grid point estimates ensure that no information content is lost due to isohyetal intervals,
- (c) estimates of catchment or areal mean rainfall can be calculated rapidly,
- (d) the grid point information can be subjected to any number of manipulations with relative ease,
- (e) the data are readily transportable (all the abovementioned images for Southern Africa can be contained on one magnetic tape).

8.2 Other uses of the products of this study

The combination of these images of precipitation statistics, the programs for interpolation and contouring as well as the monthly data random access file DATABASE, INDEX and the grid of altitude points provides a powerful and valuable new information base to the water research and development community of Southern Africa. The validity of this assertion is evidenced by the list of organisations presented in Appendix C, which have already benefited from these data and techniques.

An example of the type of study which can be performed with relative ease, using these programs and data is the study of drought in Southern Africa by Dent et al, 1987. In that study the data set of monthly rainfall records and the SACLANT interpolation package were used in conjunction with a series of techniques, developed to depict spatially, the extent, relative severity and location of the areas most affected by the recent drought in Southern Africa.

The confidence with which the mapped estimates of MAP and mean monthly rainfall may be accepted is important. Confidence levels at station and non-station points are presented in different forms. The 80 and 90 per cent con-

confidence levels of the MAP and mean monthly rainfall at stations may be estimated using Appendix A. The isohyetal surfaces pass through these station points exactly, therefore it may be assumed that the confidence levels of the estimate in proximity to that point would be approximately the same. This is especially so in areas where the isohyetal surface is not complex. Regarding non-station points which are reasonably remote from the station, either spatially or altitudinally, the confidence limits of regression maps, an example of which is presented in Figure 6.12, give some indication of the confidence that may be attached to these estimates. However, it should be noted that the combination of the regression and residual surfaces has altered the regression surfaces and therefore these confidence limits are no longer strictly true.

Some concern has been expressed by potential users who previewed the maps, that the small isohyetal lines were not labeled. This is a recognised problem, but, the plotting program decides automatically when the lines are large enough to be labeled. Smoothing the surface in order to eliminate these small isohyetal features is possible. However, the entire image would then be smoothed and this would detract from the required sensitivity of the isohyets in areas of complex topography. In addition, the smoothed surface would not fit at the station points. Fortunately the maps are printed as overlays to the topographical maps and in most areas MAP and mean monthly precipitation are correlated closely with altitude. Topographic features can therefore be used to deduce directional trends in the isohyets.

It was not one of the objectives of this project to make recommendations on network efficiency. However, a careful visual analysis of the isohyetal maps, on which are printed the rainfall stations used in the analysis, will reveal areas where the isohyetal surface is complex yet the gauge

density is sparse. Scrutiny of the confidence level maps will reveal generally that the estimates in such areas have wide confidence limits. These then are the areas in which the networks could be improved. However, this is "easier said than done", since it was probably inaccessibility and economic factors which caused a dearth of station in such areas in the first instances. Of particular note in this regard are many of the mountain catchment areas and the National States, ironically, the very areas where these data are needed most. The platform provided by this project could be used to assist in future research to study network efficiencies and to make recommendations on the placement or removal of stations in any future rationalisation programme on the national rain gauge network.

The magnitude of this mapping project and the large volumes of data which had to be manipulated and managed, dictated that most processes were automated. Such automation was an objective of the project since it was recognised that certain areas of the country may need to be remapped in ten to fifteen years time when more data are available. The high degree of automation achieved in project means that not only has the project produced valuable data sets but a range of useful techniques have emanated from this research.

The data base of rainfall and physiography which has been collected, collated and/or generated forms an important platform for a wide range of agrohydrologic and water resources engineering tasks. For example, the synthesised monthly rainfall data are considered a valuable product of this project. These data will be very useful for water resources studies such as those carried out using the Pitman Monthly Model (Pitman, 1973) the output from which forms the basis of the Surface Water Resources of South Africa series, (Pitman et al, 1981).

Another application of the physiographic data base has been the preparation of preliminary maps of monthly mean maximum and minimum temperatures and thereafter estimated evaporation surfaces for Southern Africa, which has been developed using these data (Clemence et al., 1985; Dent, Schulze, Lynch and Maharaj, 1987). Such information is useful in the area of irrigation planning and hydrologic modelling of monthly water yield. The addition of soils information to these physiographic data will make them a most valuable tool for the spatial analysis of agrohydrologic and water resources engineering problems. The potential is present for these data to be used to build "intelligence images" for satellite imagery interpretation.

8.3 Recommendations for future research

A number of applications for the information, data and techniques developed during this project have been presented above. In addition to these a platform has been created for future research in several directions.

Since the agricultural sector is a major user of water in Southern Africa the study of water resources cannot be separated from agrohydrological research. Research into the agrohydrology and climatology of Natal by Schulze (1983) has been well received and requests for similar research in other parts of Southern Africa have been numerous. The study by Schulze (1983) used a grid size of 10 km whereas it is now possible to employ a grid size of 1 minute of a degree or approximately 1,6 km over the whole country. The potential for future research in agrohydrology based on the platform created by this project is truly exciting.

The Pitman Monthly Model of Pitman (1973) has been used extensively in water resources analyses in Southern Africa Pitman et al (1981). The monthly rainfall input to this

the pitfalls of using this model in catchments where the rainfall gradient is steep. Research into improved techniques for estimating monthly and daily rainfall surfaces is already being undertaken in the Department of Agricultural Engineering, University of Natal, the base for this research having been laid during the rainfall statistics project.

Reference has also been made to employing the results of this study in future research into techniques for analysing droughts, spatially.

Finally the physiographic and climatological infrastructure created by this project is providing an essential base for present and future research into the important field of distributed hydrological modelling for water resources assessments of large catchments in Southern Africa.

CHAPTER 9

SUMMARY

9.1 Introduction and objectives

One of the most widely used variables in agrohydrology, engineering design and water resources planning is that of rainfall and in particular mean annual precipitation¹ (MAP). In the past two decades the basic source of information on mean annual rainfall in Southern Africa² has been the 1:250 000 average annual rainfall map series compiled and drawn by the Hydrological Research Division, Department of Water Affairs in 1965 from data obtained from the South African Weather Bureau. Since then the rainfall data base has lengthened by more than 20 years and as a consequence of enhanced computer power, the techniques of analysis have become more sophisticated. The engineering and hydrological fraternities requested the Water Research Commission to initiate research into the revision of the abovementioned maps. This project to revise the temporal and spatial distribution of statistics in Southern Africa² was therefore

1 The terms precipitation and rainfall are used synonymously throughout this document to designate rainfall, since fog is not measured at sufficient sites to warrant inclusion in this study and the contribution of snow is not important to total precipitation in Southern Africa.

2 Southern Africa is taken to include the Republic of South Africa, the TBVC States as well as Lesotho and parts of Swaziland for the purposes of this document.

motivated to the Water Research Commission and commenced in July 1982. The project had the following objectives:

- (a) the re-evaluation of the distribution of mean annual and mean monthly precipitation over Southern Africa,
- (b) an investigation into the variability and other statistics of precipitation over Southern Africa,
- (c) the automated mapping of these precipitation statistics, and
- (d) a re-evaluation of the regions of homogeneous precipitation distribution.

The task of mapping the temporal and spatial statistics of rainfall in Southern Africa² is one which requires extensive data sets. The creation and/or collation, checking and management of these data formed a large part of this project. Daily or monthly rainfall records from 9 409 stations form the precipitation data set. Since physiography plays an important role in influencing the spatial distribution of long term average rainfall in many areas, a comprehensive set of raw and generated physiographic data for Southern Africa was also created. These data were collected and assembled onto a grid at a resolution of one minute of a degree. In view of the magnitude of this data set it was necessary to computerise most of the functions of checking, manipulation and management.

The rainfall data record at many of the stations was lengthened by synthesising in order to obtain a better estimate of the mean monthly and mean annual rainfalls.

A combination of the multiple regression surface of MAP and the interpolated residual surface was used to provide the final MAP image before a contouring program prepared the 70 isohyetal maps of MAP at a scale of 1:250 000. Maps of mean monthly as well as monthly and annual percentiles and

coefficients of variability (CV) were furthermore produced. The latter maps were produced at a scale of 1:1 000 000 and/or as digital images to a spatial resolution of 1 minute of a degree.

9.2 Need for a physiographic data set

Polynomial spline, multi-quadric surfaces, Thiessen polygons and the subjective interpolation of isohyets rely on a fundamental assumption of the juxtaposition of similar amounts of rainfall. Their use is therefore generally confined to flatter areas where rainfall gradients are not marked (Mandeville and Rodda, 1970). It was thought initially that the flatter areas of the Karoo, Orange Free State and the Transvaal would be mapped more suitably by such interpolation techniques. In the southern and south western Cape and the eastern seaboard of Southern Africa, where the raingauge network is sparse in relation to the complex nature of the terrain, it is more suitable to undertake rainfall mapping with techniques which are able to accommodate these physiographic complexities. The influence of physiography on the spatial distribution of mean annual precipitation (MAP) has been well known for a long time and this theory is supported by a wide range of studies in Southern Africa, inter alia, by Whitmore (1968), Torrance (1973), Schulze (1976; 1977; 1979; 1983) and Hughes (1982).

In the absence of a formal report on the techniques employed to produce the 1:250 000 maps of MAP produced in the 1960's and on inspection of these isohyets, it was presumed that a major decision had been taken with regard to the influence of altitude on MAP. This new study has made no such decision, but has rather gathered and generated physiographic information for use in multiple regression equations which relate MAP to physiographic variables.

9.3 Basic rainfall and physiographic data

The nature of this project dictates that it should utilize the most comprehensive precipitation data set available in Southern Africa. Numerous organisations and individuals contributed most generously to this rainfall data set. Substantial contributions were received from the South African Weather Bureau, the Department of Agriculture and Water Supply, the South African Sugar Association, provincial parks boards, organised agriculture, municipalities, mines and private individuals. The rainfall data base comprises of 9409 stations of which 996 have more than 50 years of good data. Most of these stations have daily records which were processed to yield monthly and annual means as well as other statistics. Extensive checks were carried out to ensure that these data were as error free as possible.

In view of the aforementioned evidence in respect of the influence of physiography on mean long term rainfall it was clear that a physiographic data base would be required. A major effort in gathering and generating physiographic data was therefore embarked upon, involving the establishment, on maps of scale 1:50 000 of the exact location and altitude of all the rainfall stations, and the preparation of an altitude data set on a grid of resolution 1 minute by 1 minute of a degree, for the whole of Southern Africa.

The grid system formed the framework for the entire physiographic and regional information system, in that at each point in the grid the variables necessary for trend surface analyses, were assigned a value. The choice of grid resolution (or grid size) is of great significance. Pawling (1973) and Schulze (1977) showed that if the grid was too coarse the resultant trend surface would lack resolution. If on the other hand the grid resolution was too fine, an unmanageable amount of data would have to be

processed. The choice of a 1 minute grid was a suitable optimum in this regard.

A grid data base for approximately 70 per cent of Southern Africa was made available by the erstwhile Hydrological Research Unit (HRU), of the University of the Witwatersrand. These data had been extracted off maps of scale 1:250 000 with contour intervals of 500 feet. Altitude data from 1:50 000 scale maps with contour intervals of 20 metres were obtained from the Soil and Irrigation Research Institute of the Department of Agriculture and Water Supply and from the Department of Land Surveying and Mapping at the University of Natal for parts of the western Cape and Natal, respectively. All these data were subjected to extensive manual and computerised checks. The grid data for the eastern Transvaal, Karoo and northern Cape as well as parts of Natal and the south western Cape were completed by the staff of this project.

9.4 Generated rainfall and physiographic data

The basic data base of rainfall, latitude, longitude and altitude was insufficient for the task of mapping rainfall statistics. A large amount of generated or synthetic rainfall and physiographic data was necessary.

It was essential to optimise the temporal and spatial distribution of rainfall records and stations. Dyer (1976), Hall (1976), Dyer and Gosnell (1978), Dyer and Tyson (1977) found an apparent persistence in periods of high and low annual rainfall in Southern Africa necessitating long record lengths to be used in this project. Matalas and Jacobs (1964) reported that it was preferable to use some synthetic rainfall data in order to increase the record length rather than to estimate the long term mean by using too short a record. After taking due cognizance of these and other research findings it was decided to synthesise

monthly rainfall in order to increase the rainfall data record length. Model 5 of Zucchini and Hiemstra (1984) was selected to perform this synthesis. Other techniques which were considered were linear and multiple regression, models proposed by Zucchini and Hiemstra (1984), Gandin's method and Universal Kriging. The generated monthly rainfall values were flagged appropriately.

The altitude data formed the foundation of a set of techniques which produced other physiographic variables at each grid point. These variables were distance to a mountain range, continentality, exposure, surface roughness and aspect with respect to rainbearing weather systems. These variables were then used as predictors in the regression analysis of MAP on physiographic variables.

9.5 Checking the rainfall and physiographic data

It became necessary to develop computerised checking procedures to screen the rainfall and physiographic data for gross errors. An example of such checks was the daily rainfall data, which was aggregated in order to produce the monthly data set, these data screened to select the daily data for each month in which more than 200 mm was recorded on one day. Each of the 3 200 occurrences of such high rainfall was checked against rainfall records from neighbouring stations, the time of the year and daily rainfall extreme value statistics presented by Adamson (1981), for feasibility, before the decision to reject or accept these data was taken.

The altitude data were classified and printed out in matrix form. The classified image was then coloured in order to facilitate the visual checking of the features according to the 1:250 000 topographic map series. Prior to this more detailed check an unclassified image was produced in order

to check that all the points were present and that no points had been duplicated.

In addition to the above checks, the altitude assigned to each rainfall station was checked against the altitude on the grid data set at that point. This provided a useful check for gross errors in data capture at 9 409 "random" points on the grid altitude data set and provided for a coarse check on each of the rainfall station altitudes.

9.6 Automated management of the rainfall data base

A project of this nature requires that the data base and particularly the rainfall data be readily accessible. The rainfall data was stored in a large random access file. An index file which contained one line of key information on each station was developed. An extraction program interfaced the index and the random access data file with the user. Data were extracted by a program which read an input request menu and various routines to carry out a range of searches. Some or all of the following criteria were able to be specified in order to extract the required data; station number or name, location in terms of latitude and longitude, altitude, starting and ending year of data and number of years of record.

9.7 Regression of MAP on physiographic variables

The large volumes of data, numbers of variables and possible regression models made it imperative to develop techniques to automate the management of the procedures leading up to and including the regression procedure for relating MAP to physiographic variables. An intricate system involving numerous major procedural steps and incorporating 27 programs in FORTRAN or SAS languages was developed to perform the abovementioned functions. Throughout the system there were points at which human intervention and

checking was necessary.

Southern Africa was divided into 34 regions for the purposes of the regression exercise. All these regions overlapped by $1/2$ degree with neighbouring regions to ensure continuity of the isohyet surfaces at regional boundaries.

The process of achieving the best regression expression for the rainfall surface was iterative in the sense that maps of residuals from regression at the station points and isohyetal maps were actually produced and then studied before accepting, rejecting or modifying the regression expression. Outliers, anomalies and hyper- or insensitivity of the surfaces were thus detected and corrected.

9.8 Techniques of mapping MAP

The techniques of mapping MAP fall into two broad categories, viz trend surface analysis based on multiple regression and direct interpolation. The merits of trend surface analysis and direct interpolation were investigated. A combination of the regression surface and the surface generated by interpolating between the residuals was considered to be the most pragmatic solution. The surfaces of MAP only have real merit if their quality is known. Therefore, residual analyses and estimation of confidence levels of regression estimates were performed on the MAP data.

9.9 Mapping mean monthly and other rainfall statistics

The surface of MAP provided the foundation for the mapping of mean monthly rainfall and the variability of monthly and annual rainfall. The amount of work involved in producing separate regression expressions for each of the monthly surfaces and for each of the abovementioned statistics was prohibitive. An alternative technique which established

the ratio between the mean monthly rainfall and MAP at station points was therefore employed. These ratios were interpolated onto a regular grid and then multiplied by the MAP image. The resultant image was one of mean monthly rainfall which fitted the observed data exactly at station points and followed the general shape of the MAP surface in areas between stations. Similar techniques were employed to produce images of monthly and annual variability of rainfall.

9.10 Delimitation of homogeneous rainfall regions

The primary reason for delimiting rainfall regions was to provide a logical, physically based division of the rainfall stations so that the regression analyses could proceed using areas which were both manageable and meaningful. The meteorological systems which generate rainfall over Southern Africa vary from one part of the sub-continent to another. It is therefore useful to delimit regions of homogeneous precipitation in order to establish the factors and processes governing MAP and other precipitation statistics separately, for each region, thereby allowing for different techniques to be applied in various regions and for greater accuracy of estimation of rainfall surfaces within a region. At the same time it was also appropriate that a re-evaluation of the distribution of homogeneous precipitation regions be conducted in conjunction with this precipitation mapping project.

The delimitation of 712 homogeneous rainfall regions proceeded in the following manner. A classified image of altitude of the entire Southern Africa was printed at a resolution of 1 minute of a degree. Superimposed on this image were the locations of all stations with 10 or more years of record. The MAP at each station was noted at the station point. Thereafter followed the careful, patient and very time consuming process of delimiting homogeneous

rainfall regions based, inter alia, on MAP, altitude and aspect whilst keeping the regions reasonably small and consulting 1:250 000 topographic maps in order to provide a more detailed interpretation of the classified altitude map.

The re-evaluation of the distribution of homogeneous rainfall regions fulfilled one of the major objectives of the project. However, these regions were not used in the regression modelling and final mapping due to the severe problems of merging MAP at boundaries of these small regions. This detailed regionalisation is, however, useful for other applications, inter alia, for delimiting regions of equal agricultural potential and planning requirements, from the climatological point of view. These regions have been used by Dent, Schulze and Angus (1987) and Schmidt and Schulze (1987).

9.11 Conclusions and recommendations for future research

9.11.1 Merits of isohyetal maps vs digital images of rainfall

Isohyetal maps of MAP provide an efficient and effective means of communicating important and frequently applied information to a wide range of users. In addition, maps are inexpensive, portable and durable. However, printed isohyetal maps do have several disadvantages, e.g.

- (a) the scale of the map is fixed,
- (b) the isohyetal interval is fixed,
- (c) it is time consuming to estimate catchment mean areal rainfall from these maps,
- (d) the capture of isohyetal information from maps and into computer compatible form is tedious,
- (e) the information content on the isohyetal map is fixed insofar as one cannot manipulate these data without

first capturing them on computer.

Digital images of all the isohyetal maps were created in the process of producing the maps. These images have been retained on computer tape and are available on request. The digital images have a number of advantages, viz:

- (a) the images may be printed to maps at any desired scale,
- (b) the grid point estimates ensure that no information content is lost due to isohyetal intervals,
- (c) estimates of catchment or areal mean rainfall can be calculated rapidly,
- (d) the grid point information can be subjected to any number of manipulations with relative ease,
- (e) the data are readily transportable (all the abovementioned images for Southern Africa can be contained on one magnetic tape).

9.11.2 Other uses of the products of this study

The combination of these images of precipitation statistics, the programs for interpolation and contouring, the monthly data file and the grid of altitude points provides a powerful and valuable new information base to the water research and development community of Southern Africa.

An example of the type of study which can be performed with relative ease, using these programs and data is the study of drought in Southern Africa by Dent et al, (1987). In that study the extent, relative severity and location of the areas most affected by the recent drought in Southern Africa were depicted.

It is not one of the objectives of this project to make recommendations on network efficiency. However, a careful

visual analysis of the isohyetal maps, on which are printed the rainfall stations used in the analysis, will reveal areas where the isohyetal surface is complex yet the gauge density is sparse. Scrutiny of the confidence level maps will reveal generally that the estimates in such areas have wide confidence limits. These then are the areas in which the networks could be improved. However, this is "easier said than done", since it was probably inaccessibility and economic factors which caused a dearth of station in such areas in the first instances. Of particular note in this regard are many of the mountain catchment areas and the National States, ironically, the very areas where these data are needed most. The platform provided by this project could be used to assist in future research to study network efficiencies and to make recommendations on the placement or removal of stations in any future rationalisation programme on the national rain gauge network.

The high degree of automation achieved in project means that not only has the project produced valuable data sets but a range of useful techniques have emanated from this research.

The data base of rainfall and physiography which has been collected, collated and/or generated forms an important platform for a wide range of agrohydrologic and water resources engineering tasks. For example, the synthesised monthly rainfall data are considered a valuable product of this project. These data will be very useful for water resources studies such as those carried out using the Pitman Monthly Model (Pitman, 1973) the output from which forms the basis of the Surface Water Resources of South Africa series, (Pitman et al, 1981).

Another application of the physiographic data base has been the preparation of preliminary maps of monthly mean maximum and minimum temperatures and thereafter estimated evapora-

and hydrologic modelling of monthly water yield. The addition of soils information to these physiographic data will make them a most valuable tool for the spatial analysis of agrohydrologic and water resources engineering problems. The potential is present for these data to be used to build "intelligence images" for satellite imagery interpretation.

9.11.3 Implications for future research

A number of applications for the information, data and techniques developed during this project have been presented above. In addition to these a platform has been created for future research in several directions.

Since the agricultural sector is a major user of water in Southern Africa the study of water resources cannot be separated from agrohydrological research. Research into the agrohydrology and climatology of Natal by Schulze (1983) has been well received and requests for similar research in other parts of Southern Africa have been numerous. The study by Schulze (1983) used a grid size of 10 km whereas it is now possible to employ a grid size of 1 minute of a degree or approximately 1,6 km over the whole country. The potential for future research in agrohydrology based on the platform created by this project is truly exciting.

The Pitman Monthly Model of Pitman (1973) has been used extensively in water resources analyses in Southern Africa (Pitman et al, 1981). The monthly rainfall input to this model is most important. Midgley et al (1983) warn of the pitfalls of using this model in catchments where the rainfall gradient is steep. Research into improved techniques for estimating monthly and daily rainfall surfaces is already being undertaken in the Department of Agricultural Engineering, University of Natal. The base for this research has been laid during the rainfall statistics

Engineering, University of Natal. The base for this research has been laid during the rainfall statistics project.

Reference has also been made to employing the results of this study in future research into techniques for analysing droughts, spatially.

Finally, the physiographic and climatological infrastructure created by this project is providing an essential base for present and future research into the important field of distributed hydrological modelling for water resources assessments of large catchments in Southern Africa.

CHAPTER 10

REFERENCES

- ADAMSON, P.T. 1981. Southern African storm rainfall. Dept. Environment Affairs, Pretoria. TR102, A1-A178.
- ADAMSON, P.T. 1986. Personal Communication. Dept. Water Affairs, Durban.
- ALEXANDER, W.J.R. 1980. Precipitation in Hydrology, 80. Dept. Civil Engineering, University of Pretoria, Pretoria. Ch 1 : 1-2
- BEDIENT, P.B., HUBER, W.C. and HEANEY, J.P. 1978. Drainage density as an index of watershed development. J.Irrig. Drain.Div., 104 : 373-387.
- BLEASEDALE, A. and CHAN, Y.K. 1972. Orographic influences on the distribution of precipitation. Proc. Symp. Distr. Precip. Mtns. Areas, WMO 326, Geneva : 322-333.
- BOUGHTON, W.C. 1981. Rainfall variability in rainfall-runoff modelling. Civil Engineering Transactions, The Institution of Engineers, Australia, 68-73.
- BOX, G.E.P. and COX, D.R. 1966. An analysis of transformations. J.R. Statist.Soc.Ser.B, 26 : 211-243
- BROCKETT, G.M. 1982. Personal Communication. Cedara Agricultural Research Institute, Cedara.
- BROMLEY, K. 1980. Seasonal rainfall prediction. The Farmer, Zimbabwe, 51 : 15-16.

- BUYS, M.E.L., FABRICIUS, A.F., VAN DEN BERGH, P. and KLOPPER, A.P.J. 1979. Analysis of rainfall in South Africa. Expectancy of monthly rainfall. Dept. Agric. Tech. Services, Tech. Comm. 148, pp 33.
- CLARKE, G.P.Y. 1986. Personal communication. Dept. Biometry and Statistics, University of Natal, Pietermaritzburg.
- CLARKE, R.T. 1976. Statistical methods for the study of spatial variation in hydrological variables. Chapter 11, Rodda, J.C. (Ed): Facets of Hydrology. Wiley, London : 299-314.
- CLEMENCE, B.S.E., DENT, M.C. and SCHULZE, R.E. 1985. Estimating the spatial distribution of pan evaporation over Southern Africa using altitude, air temperature and daylength. Proceedings Second South African National Hydrology Symposium. September 1985. ACRU Report 22 : 173-188.
- COCHRAN, W.G. 1953. Sampling techniques : New York, John Wiley & Sons (cited by Matalas and Jacobs 1964).
- COLTON, D.E. 1976. Numerical simulation of the orographically induced precipitation distributions for use in hydrologic analysis. J.App.Met., 15 : 1241-1251.
- CREUTIN, J.D. and OBLED, C. 1982. Objective analysis and comparison. Wat.Resour.Res. 18 : 413-431.
- DAVIS, J.C. 1973. Statistics and data analysis in geology. John Wiley and Sons, New York : 381-389.
- DENT, M.C. 1986. Generation of a data base required for mapping of rainfall statistics in Southern Africa.

Proc. C.I.G.R. International Symposium on Agricultural Engineering, CSIR, Pretoria, 1.70-1.76.

DENT, M.C and WILLS, H.M.M. 1986. On-line data base management system for monthly rainfall in Southern Africa. Integrated Design of Hydrological Networks. Proceedings Budapest Symposium, IAHS, Publ 158 : 381-387.

DENT, M.C., SCHULZE, R.E. and ANGUS, G.R., 1987. An analysis of regional crop water requirement deficits for irrigation planning in Southern Africa. Dept. Agricultural Engineering, University of Natal ACRU Report 287. Water Research Commission, Pretoria.

DENT, M.C., SCHULZE, R.E., LYNCH, S.D. and MAHARAJ, M. 1987. Unpublished research results pertaining to temperature maps of Southern Africa. Dept. Agricultural Engineering, University of Natal, Pietermaritzburg.

DENT, M.C., WILLS, H.M.M., SCHULZE, R E. and LYNCH, S.D. 1987. Spatial and temporal analysis of the recent drought in the summer rainfall region of Southern Africa. Water SA, 13 : 37-42.

DEPARTMENT OF AGRICULTURE AND FISHERIES. 1981. Guide to soil conservation committee members, Natal Region. Cedara Agricultural Research Institute, Cedara. Mimeographed.

DE VILLIERS, G. DU T. 1980. A short note on errors in rainfall measurement. Water SA, 6 : 144-148.

DE VILLIERS, G. du T. and BOND, R. 1982. Enkele kenmerke van die reenval van Durban en omgewing. J.Univ. of Durban-Westville, 4 : 163-173.

- DIEDERIKS, R. 1979. SACLANT graphics package. User manual. Computing Service, University of Cape Town.
- DINGMAN, S.L. 1981. Elevation : a major influence on the hydrology of New Hampshire and Vermont, USA. Hydrol.Sci.Bull., 6 : 399-413.
- DUCKSTEIN, L., FOGEL, M.M. and THAMES, J.L. 1972. Elevation effects on rainfall : a stochastic model. J.Hydrol., 18 : 21-35.
- DUNNE, T. and LEOPOLD, L.B. 1978. Water in environmental planning. Freeman, San Francisco : 71-72.
- DYER, T.G.J. 1976. Expected future rainfall over selected parts of South Africa. S.Afr.J.Sci., 72 : 237-239.
- DYER, T.G.J. and GOSNELL, J.M. 1978. Long term rainfall trends in the South African sugar industry. Proc.S.Afr. Sugar Technol. Assn., 237-239.
- DYER, T.G.J. and TYSON, P.D. 1977. Estimating above and below normal rainfall periods over South Africa 1972-2000. J.Appl.Meteor., 16 : 145-147.
- EDWARDS, K.A. 1972. Estimating areal rainfall by fitting surfaces to irregularly spaced data. Proc.Symp. Distr.Precip.Mtns.Areas. WMO 326, Geneva : 565-587.
- ENGLAND, C.B. 1971. Quantitative slope aspect determination. J.Hydrol., 12 : 262-268.
- FERGUSON, H.L. 1972. Precipitation network design for large mountainous areas. Proc.Symp.Distr.Precip.Mtns. Areas, WMO 326, Geneva : 85-105.
- FIERING, M.B. 1963. Use of correlation to improve es-

timates of the mean and variance. U.S. Geol.Survey
Prof.Paper 434-C (cited by Matalas and Jacobs 1964).

GANDIN, L.S. 1965. Objective analysis of meteorological
fields. Translated from Russian by R. Hardin, p 242,
Israel Program for Scientific Translation, Jerusalem.
(Cited by Creutin and Obled, 1982).

GREEN, G.C. 1966. The evaluation of methods of rainfall
analysis and the application to the rainfall series of
Nelspruit. Unpublished M.Sc. (Agric.) thesis, Univer-
sity of O.F.S., Bloemfontein.

HALL, M.J. 1976. Dendroclimatology, rainfall and human
adaption in the later Iron Age in Natal and Zululand.
Ann. Natal Museum, 22 : 693-703.

HALL, M.J. and BARCLAY, P.A. 1975. Methods of determining
areal rainfall from observed data. Ch.4 Chapman, T.G.
and Dunin, F.X. (eds), "Precipitation in Catchment
Hydrology". Australian Academy of Science (Canberra)
: 47-57.

HIRSCH, R. 1982. A comparison of four stream flow exten-
sion techniques. Wat.Resour.Res., 18, 4 : 1081-1088.

HUFF, F.A., CHANGNON, S.A. and JONES, M.A. 1975.
Precipitation increases in the low hills of Southern
Illinois. Part I. Climatic and network studies.
Mon. Wea. Rev., 103 : 823-836.

HUGHES, D.A. 1982. The relationship between mean annual
rainfall and physiographic variables applied to a
coastal region of Southern Africa. S.Afr.Geogr.J., 64
: 41-50.

HUTCHINSON, P. 1968. An analysis of the effect of topog-

raphy on rainfall in the Taieri Catchment Area, Otago.
Earth Sci.J., 2 : 51-68.

HUTCHINSON, P. 1969. Estimation of rainfall in sparsely gauged areas. Bull.Int.Assoc.Sci.Hydrol., 14 : 101-119.

HUTCHINSON, P. 1974. Progress in the use of the method of optimum interpolation for the redesign of the Zambian network. Geoforum, 20 : 49-61.

JONES, R.I. 1982. Cedara Agricultural Research Institute, Cedara, Personal Communication.

JONES, R.I., ARNOTT, J.K. and KLUG, J.R. 1980. Fodder production planning. Report N13/80. Cedara Agricultural Research Institute, Cedara, pp 27 and appendices.

KNISEL, W.G., RENARD, K.G. and LANE, L.J. 1979. Hydrologic data collection : How long is long enough? Speciality Conference on Irrigation and Drainage in the Nineteen Eighties, Am.Soc.Civ.Eng. Albuquerque, New Mexico.

LAWLESS, J.P. and WANG, P. 1976. Simulation study of ridge and other estimates. Communications in Statistics, A5 : 307-323 (Cited by Galpin, 1981).

LEE, R. 1972. An optographic technique for evaluating the exposure of precipitation gauge sites in mountainous areas. Proc.Symp.Distr.Precip.Mtns.Areas, WMO, 326, Geneva : 67-72

LINACRE, E.T. 1977. A simple formula for estimating evaporation rates in various climates using temperature data alone. Agric. Met, 18 : 409-424.

- LONDON, W. and EMMITT, G.D. 1986. Topographical influences on radar echo properties - implications to weather modification projects in mountainous terrain. 2nd Conference on Planned and Inadvertent Weather Modification.
- LOUW, W.J. 1982. Rawson se siklus. Weerburo Nuusbrief : 373 : 103-105.
- MANDEVILLE, A.N. and RODDA, J.C. 1970. A contribution to the objective assessment of areal rainfall amounts. J.Hydrol.N.Z., 9 : 281-291.
- MARCH, W.J., WALLACE, J.R. and SWIFT, L.W. 1979. An investigation into the effect of storm type on precipitation in a small mountain watershed. Wat.Resour.Res.15 : 298-304.
- MARKHAM, R. 1980. Predictions of annual rainfall in Southern Africa. Technical Report TWISK 152, CSIR, Pretoria: 9 pp.
- MATALAS, N.C. and JACOBS, B. 1964. A correlation procedure for augmenting hydrologic data. Statistical studies in hydrology. Geological Survey Professional Paper 434-E. United States Govt Printing Office, Washington, D.C.
- McKAY, G.A. 1976. Hydrological mapping. In: Rodda, J.C. (Ed). Facets of Hydrology. John Wiley and Sons, London: 1-35.
- McQUITTY, L.L. 1960. Hierarchical syndrome analysis. Educ. Psychol. Measur. 20 : 293-304.
- MIDGLEY, D.C. and PITMAN, W.V. 1969. Surface water resources of South Africa. University of Wit-

watersrand, HRU, Report 2/69.

MIDGLEY, D.C. and PITMAN, W.V. 1978. A depth-duration-frequency diagram for point rainfall in Southern Africa. University of Witwatersrand, HRU, Report 2/78.

MIDGLEY, D.C., PITMAN, W.V. and MIDDLETON, B.J. 1983. A guide to: Surface water resources of South Africa. ISBN 090835617X. Water Research Commission, Pretoria. pp 166.

MYBURGH, J. and HOON, J. 1985. Unpublished research results. Department of Agriculture and Water Supply, Soil and Irrigation Research Institute, Elsenburg, Cape.

NETER, J. and WASSERMAN, W. 1974. Applied linear statistical models. Richard D. Irwin Inc. Homewood, Illinois.

NETER, J., WASSERMAN, W. and KUTNER, M.H. 1983. Applied Linear Regression Models, Richard D. Irwin Inc. Homewood, Illinois.

NEVILL, E. 1908. The rainfall of Natal. Agric.J.Natal, 11 : 531-1533.

NIELD, R.E. and BOSHELL, F. 1976. An agroclimatic procedure and survey of the pineapple production potential of Columbia. Ag. Met., 17 : 81-92.

OLEA, R.A. 1974. Optimal contour mapping using universal Kriging. J.Geophys.Res., 79 : 695-702.

PAWLING, J.M. 1973. Trend surface analysis of local relief in the southern peninsula of Michigan. Prof.Geogr., 25

: 226-232.

PECK, E.L. and BROWN, M.J. 1962. An approach to the development of isohyetal maps for mountainous areas. Jour.Geophys.Res., 67, 681-694.

PITMAN, W.V. 1973. A mathematical model for generating monthly river flows from meteorological data in South Africa. University of Witwatersrand, HRU, Report 2/73.

PITMAN, W.V. 1980. Analysis of large area storms in SWA/Namibia. University of Witwatersrand, HRU, Report 1/80.

PITMAN, W.V., MIDDLETON, B.J. and MIDGLEY, D.C. 1981. Surface water resources of South Africa. University of Witwatersrand, HRU, Report 9/81.

PITMAN, W.V. and STERN, J.A. 1981. Design flood determination in SWA/Namibia. University of Witwatersrand, HRU, Report 14/81.

RAWSON, H.E. 1908. The anticyclonic belt of the Southern Hemisphere. Quart.J.Roy.Meteor.Soc., 34 : 165-185. (Cited by Dyer and Tyson, 1977).

RIPLEY, B. 1980. Spatial Statistics. Wiley, London, Ch. 4.

RODDA, J.C. 1967. The systematic error in rainfall measurement. Jour.Inst.Wat.Eng., 21: 173-177.

ROSENBLATT, J.R. 1959. On the synthetic record problem: Natl. Bur. Standards Tech. Report 1.

- RYAN, K. 1986. Personal communication. Department of Biometry and Statistics, University of Natal, Pietermaritzburg.
- SCHERMERHORN, V.P., 1967. Topography and annual precipitation. Wat.Resour.Res., 3 : 707-711.
- SCHMIDT, E.J. and SCHULZE, R.E. 1987a. Flood volume and peak discharge from small catchments in Southern Africa, based on the SCS technique. Dept. Agricultural Engineering, University of Natal, Pietermaritzburg. ACRU Report 24, Water Research Commission, Pretoria Report No. 155, TT 31/87.
- SCHMIDT, E.J. and SCHULZE, R.E. 1987b. User Manual for SCS-based design runoff estimation in Southern Africa. Design guide. Dept. Agricultural Engineering, University of Natal, Pietermaritzburg. ACRU Report 25, Water Research Commission, Pretoria. No. 155, TT 33/87.
- SCHÖNAU, A.P.G. 1982. Personal communication. Institute for Commercial Forestry Research, University of Natal, Pietermaritzburg.
- SCHULZE, R.E. 1975. Catchment evapotranspiration in the Natal Drakensberg. Unpublished Ph.D. Thesis, University of Natal, Pietermaritzburg, pp. 244.
- SCHULZE, R.E. 1976. On the application of trend surfaces of precipitation to mountainous areas. Water SA, 2 : 110-118.
- SCHULZE, R.E. 1977. Practical considerations in the application of surface fitting techniques for estimation of areal rainfall. Second Conference on Hydrometeorology, Toronto, Ontario. Publ. by American Meteorologi-

cal Society, Boston, Mass. : 104-109.

SCHULZE, R.E. 1979. Hydrology and water resources of the Drakensberg. Dept. Agricultural Engineering, University of Natal, Pietermaritzburg, ACRU Report 6. Natal Town and Regional Planning Commission, Pietermaritzburg. Report 42 pp 179.

SCHULZE, R.E. 1983. Agrohydrology and -climatology of Natal. Dept. Agricultural Engineering, University of Natal, Pietermaritzburg. ACRU Report 14. Water Research Commission, Pretoria. pp 138.

SCHULZE, R.E. 1984. An assessment of surface water resources of Natal. Dept. Agricultural Engineering, University of Natal, Pietermaritzburg, ACRU Report 20. Natal Town and Regional Planning, Pietermaritzburg. Report 63 pp 44.

SEEBER, R.J. 1983. A Southern African extreme-value model for the determination of the cumulative distribution of point rainfall rate. NITR, CSIR, Pretoria. Report TEL 192. pp 148.

SEED, A.W., DENT, M.C., SCHULZE, R.E. and MURGATROYD, E. 1983. An assessment of techniques for synthesising annual rainfall data. South African Hydrological Symposium - Proceedings. Dept. Environment Affairs, Water Research Commission, Pretoria. Maaren, H. (Ed) TR119 : 94-101.

SEED, A.W. 1987. Techniques for mapping rainfall. Unpublished M.Sc.Eng. thesis. Dept. Agricultural Engineering, University of Natal, Pietermaritzburg. pp 130.

SHAW, E.M. and LYNN, P.O. 1972. Areal rainfall evaluation

using two surface fitting techniques.
Bull.Int.Assoc.Hydrol.Sci., 17: 419-433.

SOLOMON, S.I., DENOUEVILLIEZ, J.P., CHART, E.J., WOOLLEY, J.A. and CADOU, C. 1968. The use of a square grid system for computer estimation of precipitation, temperature and runoff. Wat.Resour.Res., 4 : 919-929.

SOUTH AFRICAN WEATHER BUREAU (SAWB), 1960. Climate of South Africa. Part 5. District Rainfall. WB. Govt. Printer, Pretoria.

SOUTH AFRICAN WEATHER BUREAU (SAWB), 1972. Climate of South Africa. Part 10. District Rainfall. WB 35, Govt. Printer, Pretoria.

SPARKS, R.S. and RUST, S. 1983. Unpublished FORTRAN 77 multivariate all subsets selection program. Dept. Mathematical Statistics, University of Natal, Durban.

SPARKS, R.S., TROSKIE, L. and COUTSOURIDES, D. 1985. On variable selection in multivariate regression. Comm. Statist., Theory and Method, 14 : 1567-1587.

SPREEN, W.C. 1947. A determination of the effect of topography upon precipitation. Trans. Am. Geophys. Union, 28 : 285-290.

TORRANCE, J.D. 1973. The variation of meteorological elements with increasing altitude in the eastern border mountains of Rhodesia. Rhod.J.Agric.Res., 11 : 181-189.

WARD, R.C. 1975. Principles of hydrology (2nd Ed). McGraw-Hill, London.

WEISBERG, S. 1980. Applied linear regression. John Wiley

- and Sons, New York: 174-179.
- WELDING, M.C. and HAVENGA, C.M. 1974. The statistical classification of rainfall stations in the Republic of South Africa. *Agrochemophysica*, 6 : 5-24.
- WHITMORE, J.S. 1962a. Variability of annual rainfall in the Republic of South Africa. Dept. Water Affairs, TR 24. pp 8.
- WHITMORE, J.S. 1962b. On evidence of "persistence" in annual rainfall in the Republic of South Africa. Dept. Water Affairs, TR 25 : pp 11.
- WHITMORE, J.S. 1968. The relationship between mean annual rainfall and locality and site factors. *S.Afr.J.Sci.* 64 : 423-427.
- WHITMORE, J.S. 1972. The variation of mean annual rainfall with altitude and locality in South Africa, as determined by multiple curvilinear regression analysis. Dept. Water Affairs, Technical Report No.53, Pretoria. pp 13.
- WICHT, C.L., MEYBURGH, J.C. and BOUSTEAD, P.G. 1969. Rainfall at the Jonkershoek Hydrological Research Station. *Annals University of Stellenbosch.* 44 : pp 66.
- WORLD METEOROLOGICAL ORGANISATION. 1974. Guide to hydrological practices. WMO, Geneva, Switzerland No. 168, Ed 3.
- YOUNGER, M.S., 1979. Handbook for linear regression. North Scituate, Mass. Duxbury, U.S.A.
- ZUCCHINI, W. and HIEMSTRA, L.A.V. 1984. Part 1:

Augmenting hydrological records. Water Research Commission Pretoria Report No. 91/3/84

ZUCCHINI, W. and SPARKS, R.S. 1984a. Part II: Estimating missing values in rainfall records. Water Research Commission, Pretoria. Report No. 91/3/84.

ZUCCHINI, W and SPARKS, R.S. 1984b. Personal Communication. Dept. of Statistics, University of Cape Town.

ZUCCHINI, W. 1985. Personal communication. Dept. of Statistics, University of Cape Town.

CHAPTER 11

APPENDIX A

UPPER LIMITS OF THE PERCENTAGE ERROR OF THE SHORT TERM MEAN AND STANDARD DEVIATION

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 1		POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 75 STATIONS												
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN	AT 90% CONFD.	82.5	82.5	60.5	51.5	56.5	45.5	45.5	46.5	52.5	60.5	76.5	74.5	21.5
	80% CONFD.	65.5	64.5	47.5	40.5	43.5	34.5	35.5	35.5	41.5	46.5	58.5	57.5	16.5
	SAMPLE SIZE	4566	4562	4529	4522	4525	4534	4530	4524	4531	4533	4531	4532	4512
10 YEAR MEAN	AT 90% CONFD.	52.5	60.5	38.5	33.5	38.5	30.5	32.5	34.5	39.5	43.5	47.5	52.5	15.5
	80% CONFD.	40.5	44.5	29.5	26.5	30.5	23.5	25.5	26.5	31.5	35.5	37.5	40.5	12.5
	SAMPLE SIZE	4191	4187	4154	4147	4150	4159	4155	4149	4156	4158	4156	4157	4137
15 YEAR MEAN	AT 90% CONFD.	41.5	39.5	29.5	27.5	28.5	23.5	24.5	29.5	34.5	34.5	33.5	41.5	13.5
	80% CONFD.	30.5	30.5	22.5	22.5	22.5	18.5	19.5	22.5	27.5	28.5	26.5	31.5	9.5
	SAMPLE SIZE	3816	3812	3779	3772	3775	3784	3780	3774	3781	3783	3781	3782	3762
20 YEAR MEAN	AT 90% CONFD.	32.5	30.5	23.5	24.5	22.5	20.5	18.5	26.5	30.5	28.5	30.5	34.5	10.5
	80% CONFD.	23.5	24.5	17.5	19.5	17.5	14.5	14.5	19.5	23.5	23.5	23.5	26.5	7.5
	SAMPLE SIZE	3441	3437	3404	3397	3400	3409	3405	3399	3406	3408	3406	3407	3387
25 YEAR MEAN	AT 90% CONFD.	29.5	25.5	20.5	21.5	20.5	18.5	14.5	22.5	25.5	24.5	24.5	31.5	8.5
	80% CONFD.	22.5	19.5	15.5	16.5	15.5	12.5	11.5	17.5	19.5	19.5	18.5	23.5	6.5
	SAMPLE SIZE	3066	3062	3029	3022	3025	3034	3030	3024	3031	3033	3031	3032	3012
30 YEAR MEAN	AT 90% CONFD.	26.5	23.5	16.5	19.5	17.5	14.5	11.5	18.5	21.5	21.5	19.5	28.5	7.5
	80% CONFD.	19.5	18.5	12.5	15.5	12.5	11.5	8.5	14.5	16.5	16.5	15.5	21.5	5.5
	SAMPLE SIZE	2691	2687	2654	2647	2650	2659	2655	2649	2656	2658	2656	2657	2637
35 YEAR MEAN	AT 90% CONFD.	23.5	19.5	14.5	17.5	14.5	13.5	9.5	15.5	17.5	17.5	17.5	25.5	6.5
	80% CONFD.	17.5	15.5	11.5	13.5	10.5	9.5	7.5	11.5	13.5	13.5	13.5	18.5	4.5
	SAMPLE SIZE	2316	2312	2279	2272	2275	2284	2280	2274	2281	2283	2281	2282	2262
40 YEAR MEAN	AT 90% CONFD.	21.5	17.5	12.5	15.5	11.5	11.5	8.5	13.5	13.5	14.5	14.5	21.5	5.5
	80% CONFD.	15.5	13.5	9.5	12.5	8.5	8.5	6.5	10.5	11.5	10.5	11.5	16.5	4.5
	SAMPLE SIZE	1941	1937	1904	1897	1900	1909	1905	1899	1906	1908	1906	1907	1887
WELDING & HAVENGA REGIONS 2		POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 115 STATIONS												
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN	AT 90% CONFD.	75.5	89.5	64.5	57.5	46.5	45.5	48.5	37.5	43.5	45.5	69.5	72.5	20.5
	80% CONFD.	58.5	72.5	50.5	43.5	35.5	33.5	36.5	27.5	34.5	35.5	53.5	54.5	16.5
	SAMPLE SIZE	7735	7732	7708	7662	7669	7674	7678	7677	7680	7682	7691	7685	7615
10 YEAR MEAN	AT 90% CONFD.	58.5	71.5	46.5	40.5	32.5	32.5	34.5	23.5	31.5	32.5	42.5	51.5	15.5
	80% CONFD.	44.5	53.5	35.5	31.5	24.5	23.5	26.5	18.5	24.5	25.5	33.5	39.5	11.5
	SAMPLE SIZE	7160	7157	7133	7087	7094	7099	7103	7102	7105	7107	7116	7110	7040
15 YEAR MEAN	AT 90% CONFD.	48.5	53.5	37.5	34.5	23.5	25.5	25.5	19.5	26.5	26.5	32.5	41.5	12.5
	80% CONFD.	37.5	41.5	27.5	27.5	18.5	18.5	19.5	15.5	20.5	19.5	26.5	32.5	10.5
	SAMPLE SIZE	6585	6582	6558	6512	6519	6524	6528	6527	6530	6532	6541	6535	6465
20 YEAR MEAN	AT 90% CONFD.	41.5	42.5	30.5	30.5	19.5	19.5	19.5	16.5	23.5	21.5	28.5	35.5	10.5
	80% CONFD.	31.5	33.5	23.5	24.5	15.5	14.5	15.5	12.5	17.5	17.5	21.5	27.5	8.5
	SAMPLE SIZE	6010	6007	5983	5937	5944	5949	5953	5952	5955	5957	5966	5960	5890
25 YEAR MEAN	AT 90% CONFD.	38.5	33.5	25.5	25.5	17.5	17.5	15.5	15.5	20.5	18.5	21.5	31.5	8.5
	80% CONFD.	28.5	26.5	19.5	20.5	13.5	12.5	12.5	11.5	15.5	15.5	17.5	24.5	6.5
	SAMPLE SIZE	5435	5432	5408	5362	5369	5374	5378	5377	5380	5382	5391	5385	5315
30 YEAR MEAN	AT 90% CONFD.	34.5	29.5	22.5	20.5	14.5	14.5	11.5	13.5	16.5	17.5	18.5	28.5	7.5
	80% CONFD.	26.5	22.5	16.5	16.5	11.5	11.5	9.5	9.5	12.5	12.5	14.5	22.5	5.5
	SAMPLE SIZE	4860	4857	4833	4787	4794	4799	4803	4802	4805	4807	4816	4810	4740
35 YEAR MEAN	AT 90% CONFD.	32.5	24.5	19.5	17.5	12.5	12.5	9.5	11.5	13.5	14.5	16.5	26.5	6.5
	80% CONFD.	24.5	18.5	15.5	13.5	9.5	8.5	7.5	8.5	10.5	11.5	13.5	20.5	4.5
	SAMPLE SIZE	4285	4282	4258	4212	4219	4224	4228	4227	4230	4232	4241	4235	4165
40 YEAR MEAN	AT 90% CONFD.	28.5	22.5	17.5	13.5	10.5	10.5	9.5	9.5	11.5	12.5	14.5	25.5	5.5
	80% CONFD.	22.5	17.5	13.5	10.5	7.5	7.5	7.5	7.5	8.5	9.5	11.5	19.5	4.5
	SAMPLE SIZE	3710	3707	3683	3637	3644	3649	3653	3652	3655	3657	3666	3660	3590

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 3 AND 10 POSSIBLE			PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 55 STATIONS											
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN	
5 YEAR MEAN AT 90% CONFD.	100.5	99.5	77.5	70.5	64.5	67.5	68.5	55.5	77.5	73.5	98.5	100.5	31.5	
80% CONFD.	84.5	82.5	61.5	56.5	50.5	50.5	52.5	43.5	61.5	56.5	78.5	77.5	24.5	
SAMPLE SIZE	3602	3600	3570	3574	3577	3570	3577	3579	3584	3583	3586	3586	3550	
10 YEAR MEAN AT 90% CONFD.	85.5	77.5	51.5	51.5	45.5	51.5	49.5	37.5	55.5	53.5	68.5	73.5	23.5	
80% CONFD.	68.5	61.5	40.5	39.5	34.5	39.5	38.5	28.5	45.5	39.5	55.5	58.5	17.5	
SAMPLE SIZE	3327	3325	3295	3299	3302	3295	3302	3304	3309	3308	3311	3311	3275	
15 YEAR MEAN AT 90% CONFD.	69.5	52.5	42.5	42.5	34.5	40.5	38.5	30.5	47.5	42.5	55.5	55.5	18.5	
80% CONFD.	56.5	41.5	31.5	33.5	26.5	31.5	29.5	21.5	37.5	32.5	42.5	42.5	13.5	
SAMPLE SIZE	3052	3050	3020	3024	3027	3020	3027	3029	3034	3033	3036	3036	3000	
20 YEAR MEAN AT 90% CONFD.	59.5	39.5	33.5	36.5	28.5	33.5	31.5	25.5	41.5	35.5	46.5	44.5	14.5	
80% CONFD.	46.5	30.5	25.5	29.5	21.5	25.5	24.5	17.5	33.5	27.5	36.5	35.5	10.5	
SAMPLE SIZE	2777	2775	2745	2749	2752	2745	2752	2754	2759	2758	2761	2761	2725	
25 YEAR MEAN AT 90% CONFD.	53.5	34.5	27.5	31.5	24.5	29.5	26.5	20.5	38.5	31.5	39.5	41.5	12.5	
80% CONFD.	40.5	27.5	20.5	25.5	17.5	22.5	20.5	15.5	30.5	23.5	30.5	30.5	8.5	
SAMPLE SIZE	2502	2500	2470	2474	2477	2470	2477	2479	2484	2483	2486	2486	2450	
30 YEAR MEAN AT 90% CONFD.	46.5	30.5	22.5	27.5	20.5	25.5	22.5	18.5	33.5	28.5	34.5	36.5	10.5	
80% CONFD.	37.5	23.5	17.5	22.5	14.5	19.5	16.5	13.5	23.5	21.5	26.5	28.5	7.5	
SAMPLE SIZE	2227	2225	2195	2199	2202	2195	2202	2204	2209	2208	2211	2211	2175	
35 YEAR MEAN AT 90% CONFD.	41.5	25.5	20.5	24.5	16.5	21.5	19.5	14.5	26.5	24.5	31.5	31.5	9.5	
80% CONFD.	32.5	19.5	15.5	19.5	12.5	16.5	14.5	10.5	19.5	17.5	23.5	23.5	7.5	
SAMPLE SIZE	1952	1950	1920	1924	1927	1920	1927	1929	1934	1933	1936	1936	1900	
40 YEAR MEAN AT 90% CONFD.	39.5	23.5	17.5	21.5	14.5	18.5	18.5	12.5	22.5	21.5	27.5	27.5	8.5	
80% CONFD.	31.5	17.5	13.5	16.5	10.5	14.5	13.5	9.5	16.5	15.5	20.5	19.5	6.5	
SAMPLE SIZE	1677	1675	1645	1649	1652	1645	1652	1654	1659	1658	1661	1661	1625	
WELDING & HAVENGA REGIONS 4			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 52 STATIONS											
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN	
5 YEAR MEAN AT 90% CONFD.	71.5	68.5	51.5	51.5	55.5	50.5	54.5	51.5	53.5	58.5	61.5	73.5	20.5	
80% CONFD.	55.5	53.5	39.5	39.5	43.5	39.5	42.5	38.5	40.5	43.5	47.5	56.5	15.5	
SAMPLE SIZE	3532	3535	3515	3506	3502	3502	3504	3507	3503	3506	3504	3507	3497	
10 YEAR MEAN AT 90% CONFD.	44.5	45.5	32.5	33.5	36.5	37.5	38.5	38.5	35.5	41.5	39.5	50.5	15.5	
80% CONFD.	34.5	33.5	25.5	26.5	29.5	28.5	29.5	27.5	28.5	30.5	30.5	39.5	11.5	
SAMPLE SIZE	3272	3275	3255	3246	3242	3242	3244	3247	3243	3246	3244	3247	3237	
15 YEAR MEAN AT 90% CONFD.	34.5	32.5	24.5	27.5	29.5	31.5	28.5	32.5	28.5	28.5	31.5	40.5	12.5	
80% CONFD.	25.5	24.5	18.5	20.5	23.5	25.5	21.5	24.5	22.5	22.5	24.5	31.5	9.5	
SAMPLE SIZE	3012	3015	2995	2986	2982	2982	2984	2987	2983	2986	2984	2987	2977	
20 YEAR MEAN AT 90% CONFD.	27.5	27.5	20.5	24.5	25.5	28.5	22.5	27.5	23.5	22.5	28.5	34.5	10.5	
80% CONFD.	20.5	20.5	14.5	18.5	19.5	21.5	16.5	22.5	18.5	17.5	20.5	25.5	7.5	
SAMPLE SIZE	2752	2755	2735	2726	2722	2722	2724	2727	2723	2726	2724	2727	2717	
25 YEAR MEAN AT 90% CONFD.	24.5	25.5	16.5	21.5	22.5	24.5	19.5	24.5	21.5	19.5	23.5	30.5	9.5	
80% CONFD.	17.5	19.5	12.5	16.5	17.5	18.5	14.5	19.5	16.5	14.5	17.5	22.5	6.5	
SAMPLE SIZE	2492	2495	2475	2466	2462	2462	2464	2467	2463	2466	2464	2467	2457	
30 YEAR MEAN AT 90% CONFD.	21.5	21.5	15.5	19.5	20.5	22.5	16.5	20.5	18.5	15.5	20.5	28.5	8.5	
80% CONFD.	15.5	17.5	11.5	15.5	15.5	16.5	12.5	16.5	14.5	12.5	15.5	21.5	6.5	
SAMPLE SIZE	2232	2235	2215	2206	2202	2202	2204	2207	2203	2206	2204	2207	2197	
35 YEAR MEAN AT 90% CONFD.	19.5	19.5	13.5	18.5	17.5	20.5	14.5	17.5	17.5	13.5	17.5	25.5	7.5	
80% CONFD.	14.5	15.5	9.5	14.5	12.5	15.5	10.5	13.5	12.5	10.5	13.5	19.5	5.5	
SAMPLE SIZE	1972	1975	1955	1946	1942	1942	1944	1947	1943	1946	1944	1947	1937	
40 YEAR MEAN AT 90% CONFD.	16.5	17.5	11.5	17.5	14.5	18.5	12.5	15.5	14.5	11.5	14.5	23.5	7.5	
80% CONFD.	11.5	13.5	8.5	12.5	11.5	13.5	9.5	11.5	10.5	9.5	11.5	16.5	5.5	
SAMPLE SIZE	1712	1715	1695	1686	1682	1682	1684	1687	1683	1686	1684	1687	1677	

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 5			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 42 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN	AT 90% CONFD.		60.5	58.5	48.5	52.5	57.5	53.5	51.5	58.5	49.5	50.5	50.5	61.5	18.5
	80% CONFD.		43.5	42.5	34.5	40.5	43.5	40.5	39.5	44.5	37.5	36.5	39.5	42.5	14.5
	SAMPLE SIZE		2956	2952	2955	2930	2926	2933	2935	2933	2933	2935	2940	2934	2922
10 YEAR MEAN	AT 90% CONFD.		37.5	40.5	34.5	35.5	33.5	36.5	37.5	42.5	34.5	34.5	33.5	45.5	13.5
	80% CONFD.		27.5	29.5	24.5	27.5	25.5	28.5	27.5	32.5	25.5	26.5	25.5	34.5	10.5
	SAMPLE SIZE		2746	2742	2745	2720	2716	2723	2725	2723	2723	2725	2730	2724	2712
15 YEAR MEAN	AT 90% CONFD.		29.5	29.5	24.5	30.5	26.5	31.5	28.5	37.5	27.5	25.5	26.5	35.5	11.5
	80% CONFD.		21.5	21.5	19.5	23.5	20.5	24.5	21.5	28.5	21.5	20.5	19.5	25.5	8.5
	SAMPLE SIZE		2536	2532	2535	2510	2506	2513	2515	2513	2513	2515	2520	2514	2502
20 YEAR MEAN	AT 90% CONFD.		23.5	24.5	18.5	26.5	21.5	27.5	22.5	31.5	23.5	20.5	23.5	29.5	10.5
	80% CONFD.		17.5	17.5	14.5	20.5	15.5	21.5	17.5	25.5	18.5	16.5	17.5	21.5	7.5
	SAMPLE SIZE		2326	2322	2325	2300	2296	2303	2305	2303	2303	2305	2310	2304	2292
25 YEAR MEAN	AT 90% CONFD.		20.5	22.5	16.5	22.5	19.5	25.5	21.5	27.5	21.5	17.5	20.5	25.5	9.5
	80% CONFD.		15.5	14.5	12.5	17.5	13.5	18.5	16.5	22.5	15.5	13.5	15.5	19.5	6.5
	SAMPLE SIZE		2116	2112	2115	2090	2086	2093	2095	2093	2093	2095	2100	2094	2082
30 YEAR MEAN	AT 90% CONFD.		17.5	19.5	14.5	20.5	16.5	23.5	19.5	24.5	18.5	15.5	18.5	24.5	8.5
	80% CONFD.		12.5	14.5	10.5	14.5	11.5	17.5	14.5	20.5	13.5	11.5	13.5	17.5	5.5
	SAMPLE SIZE		1906	1902	1905	1880	1876	1883	1885	1883	1883	1885	1890	1884	1872
35 YEAR MEAN	AT 90% CONFD.		15.5	16.5	12.5	17.5	14.5	21.5	17.5	21.5	15.5	13.5	16.5	21.5	7.5
	80% CONFD.		11.5	12.5	8.5	13.5	10.5	16.5	13.5	17.5	11.5	9.5	11.5	16.5	5.5
	SAMPLE SIZE		1696	1692	1695	1670	1666	1673	1675	1673	1673	1675	1680	1674	1662
40 YEAR MEAN	AT 90% CONFD.		13.5	15.5	10.5	16.5	13.5	19.5	15.5	18.5	14.5	11.5	14.5	19.5	6.5
	80% CONFD.		9.5	11.5	7.5	11.5	9.5	14.5	12.5	15.5	10.5	8.5	10.5	14.5	4.5
	SAMPLE SIZE		1486	1482	1485	1460	1456	1463	1465	1463	1463	1465	1470	1464	1452
WELDING & HAVENGA REGIONS 6			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 42 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN	AT 90% CONFD.		59.5	57.5	50.5	52.5	59.5	52.5	46.5	51.5	50.5	46.5	52.5	54.5	16.5
	80% CONFD.		45.5	42.5	36.5	41.5	45.5	38.5	35.5	39.5	38.5	34.5	40.5	40.5	12.5
	SAMPLE SIZE		2858	2853	2852	2840	2830	2836	2836	2837	2838	2844	2843	2841	2824
10 YEAR MEAN	AT 90% CONFD.		36.5	36.5	34.5	36.5	33.5	32.5	34.5	38.5	33.5	33.5	32.5	37.5	12.5
	80% CONFD.		27.5	27.5	27.5	29.5	25.5	25.5	25.5	29.5	26.5	25.5	24.5	28.5	8.5
	SAMPLE SIZE		2648	2643	2642	2630	2620	2626	2626	2627	2628	2634	2633	2631	2614
15 YEAR MEAN	AT 90% CONFD.		27.5	26.5	25.5	30.5	26.5	25.5	27.5	31.5	25.5	25.5	24.5	29.5	10.5
	80% CONFD.		20.5	20.5	20.5	25.5	19.5	19.5	21.5	24.5	19.5	19.5	19.5	22.5	7.5
	SAMPLE SIZE		2438	2433	2432	2420	2410	2416	2416	2417	2418	2424	2423	2421	2404
20 YEAR MEAN	AT 90% CONFD.		20.5	21.5	20.5	25.5	20.5	20.5	22.5	26.5	19.5	20.5	21.5	24.5	8.5
	80% CONFD.		15.5	16.5	16.5	21.5	15.5	15.5	17.5	21.5	14.5	15.5	16.5	19.5	5.5
	SAMPLE SIZE		2228	2223	2222	2210	2200	2206	2206	2207	2208	2214	2213	2211	2194
25 YEAR MEAN	AT 90% CONFD.		17.5	19.5	17.5	22.5	17.5	16.5	19.5	23.5	16.5	15.5	17.5	21.5	6.5
	80% CONFD.		13.5	15.5	14.5	18.5	12.5	12.5	14.5	18.5	12.5	12.5	14.5	17.5	4.5
	SAMPLE SIZE		2018	2013	2012	2000	1990	1996	1996	1997	1998	2004	2003	2001	1984
30 YEAR MEAN	AT 90% CONFD.		14.5	16.5	14.5	19.5	14.5	14.5	16.5	21.5	15.5	13.5	15.5	19.5	6.5
	80% CONFD.		11.5	12.5	11.5	15.5	10.5	11.5	12.5	17.5	11.5	10.5	11.5	16.5	4.5
	SAMPLE SIZE		1808	1803	1802	1790	1780	1786	1786	1787	1788	1794	1793	1791	1774
35 YEAR MEAN	AT 90% CONFD.		12.5	14.5	13.5	18.5	12.5	12.5	14.5	19.5	13.5	12.5	14.5	17.5	5.5
	80% CONFD.		9.5	11.5	10.5	13.5	9.5	10.5	11.5	15.5	10.5	9.5	10.5	13.5	4.5
	SAMPLE SIZE		1598	1593	1592	1580	1570	1576	1576	1577	1578	1584	1583	1581	1564
40 YEAR MEAN	AT 90% CONFD.		10.5	13.5	11.5	16.5	10.5	11.5	13.5	16.5	11.5	11.5	12.5	15.5	4.5
	80% CONFD.		8.5	10.5	8.5	12.5	8.5	8.5	10.5	13.5	9.5	8.5	9.5	12.5	3.5
	SAMPLE SIZE		1388	1383	1382	1370	1360	1366	1366	1367	1368	1374	1373	1371	1354

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 7			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 48 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN	AT 90% CONFD.		59.5	51.5	49.5	56.5	64.5	66.5	66.5	68.5	70.5	51.5	55.5	54.5	22.5
	80% CONFD.		47.5	40.5	35.5	45.5	51.5	51.5	51.5	50.5	51.5	40.5	43.5	42.5	16.5
	SAMPLE SIZE		3457	3451	3448	3431	3428	3427	3430	3429	3427	3433	3439	3437	3419
10 YEAR MEAN	AT 90% CONFD.		37.5	35.5	32.5	39.5	39.5	41.5	45.5	56.5	48.5	36.5	35.5	39.5	16.5
	80% CONFD.		29.5	27.5	25.5	29.5	30.5	33.5	35.5	40.5	37.5	28.5	27.5	29.5	11.5
	SAMPLE SIZE		3217	3211	3208	3191	3188	3187	3190	3189	3187	3193	3199	3197	3179
15 YEAR MEAN	AT 90% CONFD.		28.5	25.5	25.5	30.5	30.5	34.5	34.5	46.5	34.5	28.5	26.5	30.5	12.5
	80% CONFD.		21.5	19.5	18.5	23.5	23.5	25.5	26.5	35.5	26.5	22.5	20.5	23.5	9.5
	SAMPLE SIZE		2977	2971	2968	2951	2948	2947	2950	2949	2947	2953	2959	2957	2939
20 YEAR MEAN	AT 90% CONFD.		22.5	19.5	21.5	25.5	23.5	25.5	29.5	38.5	25.5	23.5	21.5	26.5	10.5
	80% CONFD.		17.5	15.5	15.5	19.5	17.5	19.5	21.5	30.5	19.5	18.5	16.5	20.5	7.5
	SAMPLE SIZE		2737	2731	2728	2711	2708	2707	2710	2709	2707	2713	2719	2717	2699
25 YEAR MEAN	AT 90% CONFD.		19.5	17.5	18.5	21.5	19.5	21.5	25.5	33.5	20.5	19.5	18.5	24.5	9.5
	80% CONFD.		14.5	13.5	13.5	16.5	14.5	16.5	19.5	27.5	16.5	15.5	14.5	18.5	6.5
	SAMPLE SIZE		2497	2491	2488	2471	2468	2467	2470	2469	2467	2473	2479	2477	2459
30 YEAR MEAN	AT 90% CONFD.		16.5	13.5	15.5	17.5	16.5	19.5	23.5	30.5	18.5	16.5	16.5	21.5	8.5
	80% CONFD.		12.5	10.5	11.5	13.5	12.5	14.5	17.5	25.5	14.5	12.5	11.5	16.5	6.5
	SAMPLE SIZE		2257	2251	2248	2231	2228	2227	2230	2229	2227	2233	2239	2237	2219
35 YEAR MEAN	AT 90% CONFD.		14.5	11.5	12.5	15.5	14.5	17.5	20.5	27.5	15.5	13.5	14.5	18.5	7.5
	80% CONFD.		11.5	9.5	8.5	12.5	11.5	13.5	15.5	23.5	12.5	11.5	10.5	14.5	5.5
	SAMPLE SIZE		2017	2011	2008	1991	1988	1987	1990	1989	1987	1993	1999	1997	1979
40 YEAR MEAN	AT 90% CONFD.		12.5	11.5	11.5	13.5	12.5	15.5	18.5	24.5	12.5	11.5	12.5	16.5	6.5
	80% CONFD.		8.5	8.5	8.5	10.5	9.5	11.5	13.5	21.5	10.5	9.5	9.5	13.5	4.5
	SAMPLE SIZE		1777	1771	1768	1751	1748	1747	1750	1749	1747	1753	1759	1757	1739
WELDING & HAVENGA REGIONS 8			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 68 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN	AT 90% CONFD.		42.5	42.5	45.5	54.5	64.5	79.5	91.5	83.5	59.5	43.5	45.5	43.5	19.5
	80% CONFD.		32.5	33.5	34.5	42.5	49.5	61.5	72.5	61.5	47.5	33.5	35.5	33.5	14.5
	SAMPLE SIZE		4805	4814	4797	4773	4771	4776	4767	4789	4769	4785	4789	4789	4720
10 YEAR MEAN	AT 90% CONFD.		28.5	29.5	31.5	37.5	43.5	48.5	79.5	78.5	37.5	29.5	32.5	28.5	14.5
	80% CONFD.		22.5	22.5	24.5	29.5	33.5	38.5	63.5	52.5	28.5	21.5	24.5	22.5	10.5
	SAMPLE SIZE		4465	4474	4457	4433	4431	4436	4427	4449	4429	4445	4449	4449	4380
15 YEAR MEAN	AT 90% CONFD.		21.5	21.5	25.5	30.5	33.5	39.5	69.5	69.5	28.5	22.5	26.5	21.5	11.5
	80% CONFD.		15.5	16.5	19.5	24.5	25.5	29.5	56.5	47.5	21.5	17.5	20.5	16.5	8.5
	SAMPLE SIZE		4125	4134	4117	4093	4091	4096	4087	4109	4089	4105	4109	4109	4040
20 YEAR MEAN	AT 90% CONFD.		16.5	17.5	20.5	25.5	26.5	31.5	58.5	53.5	21.5	18.5	22.5	17.5	10.5
	80% CONFD.		12.5	13.5	16.5	21.5	19.5	24.5	47.5	42.5	16.5	14.5	17.5	13.5	6.5
	SAMPLE SIZE		3785	3794	3777	3753	3751	3756	3747	3769	3749	3765	3769	3769	3700
25 YEAR MEAN	AT 90% CONFD.		14.5	15.5	17.5	23.5	19.5	27.5	47.5	45.5	16.5	15.5	19.5	16.5	9.5
	80% CONFD.		10.5	12.5	13.5	18.5	15.5	20.5	38.5	38.5	13.5	11.5	15.5	12.5	6.5
	SAMPLE SIZE		3445	3454	3437	3413	3411	3416	3407	3429	3409	3425	3429	3429	3360
30 YEAR MEAN	AT 90% CONFD.		13.5	13.5	14.5	20.5	17.5	24.5	37.5	39.5	15.5	12.5	17.5	14.5	8.5
	80% CONFD.		9.5	10.5	11.5	16.5	13.5	18.5	32.5	35.5	12.5	9.5	13.5	11.5	5.5
	SAMPLE SIZE		3105	3114	3097	3073	3071	3076	3067	3089	3069	3085	3089	3089	3020
35 YEAR MEAN	AT 90% CONFD.		12.5	11.5	13.5	18.5	16.5	21.5	33.5	37.5	13.5	9.5	14.5	12.5	7.5
	80% CONFD.		8.5	8.5	10.5	14.5	12.5	17.5	26.5	34.5	10.5	7.5	12.5	9.5	5.5
	SAMPLE SIZE		2765	2774	2757	2733	2731	2736	2727	2749	2729	2745	2749	2749	2680
40 YEAR MEAN	AT 90% CONFD.		10.5	11.5	12.5	15.5	14.5	20.5	25.5	35.5	12.5	8.5	12.5	11.5	7.5
	80% CONFD.		7.5	8.5	9.5	12.5	10.5	15.5	21.5	32.5	9.5	6.5	10.5	8.5	4.5
	SAMPLE SIZE		2425	2434	2417	2393	2391	2396	2387	2409	2389	2405	2409	2409	2340

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 9 AND 13 POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 81 STATIONS													
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN AT 90% CONFD.	85.5	80.5	63.5	63.5	82.5	83.5	81.5	92.5	88.5	79.5	85.5	81.5	32.5
80% CONFD.	66.5	65.5	49.5	50.5	64.5	63.5	64.5	72.5	70.5	61.5	68.5	65.5	24.5
SAMPLE SIZE	5259	5257	5229	5224	5212	5207	5206	5204	5196	5210	5203	5212	5172
10 YEAR MEAN AT 90% CONFD.	58.5	55.5	46.5	47.5	56.5	59.5	60.5	71.5	63.5	58.5	57.5	57.5	24.5
80% CONFD.	45.5	43.5	35.5	37.5	42.5	44.5	46.5	55.5	49.5	44.5	44.5	45.5	17.5
SAMPLE SIZE	4854	4852	4824	4819	4807	4802	4801	4799	4791	4805	4798	4807	4767
15 YEAR MEAN AT 90% CONFD.	46.5	40.5	37.5	38.5	45.5	48.5	49.5	57.5	49.5	46.5	46.5	41.5	19.5
80% CONFD.	36.5	31.5	28.5	29.5	34.5	36.5	38.5	43.5	38.5	36.5	36.5	32.5	13.5
SAMPLE SIZE	4449	4447	4419	4414	4402	4397	4396	4394	4386	4400	4393	4402	4362
20 YEAR MEAN AT 90% CONFD.	39.5	31.5	31.5	32.5	36.5	38.5	40.5	47.5	41.5	36.5	41.5	32.5	17.5
80% CONFD.	30.5	24.5	25.5	25.5	27.5	29.5	30.5	35.5	31.5	29.5	31.5	25.5	11.5
SAMPLE SIZE	4044	4042	4014	4009	3997	3992	3991	3989	3981	3995	3988	3997	3957
25 YEAR MEAN AT 90% CONFD.	36.5	27.5	26.5	28.5	31.5	33.5	33.5	38.5	35.5	31.5	35.5	27.5	14.5
80% CONFD.	27.5	20.5	20.5	21.5	23.5	25.5	26.5	28.5	26.5	23.5	26.5	21.5	10.5
SAMPLE SIZE	3639	3637	3609	3604	3592	3587	3586	3584	3576	3590	3583	3592	3552
30 YEAR MEAN AT 90% CONFD.	32.5	24.5	23.5	26.5	25.5	28.5	30.5	30.5	30.5	26.5	29.5	24.5	13.5
80% CONFD.	24.5	18.5	17.5	19.5	19.5	22.5	22.5	23.5	24.5	19.5	22.5	18.5	9.5
SAMPLE SIZE	3234	3232	3204	3199	3187	3182	3181	3179	3171	3185	3178	3187	3147
35 YEAR MEAN AT 90% CONFD.	29.5	20.5	19.5	24.5	22.5	24.5	27.5	27.5	26.5	21.5	26.5	21.5	11.5
80% CONFD.	21.5	15.5	15.5	17.5	17.5	18.5	19.5	20.5	21.5	16.5	19.5	16.5	8.5
SAMPLE SIZE	2829	2827	2799	2794	2782	2777	2776	2774	2766	2780	2773	2782	2742
40 YEAR MEAN AT 90% CONFD.	27.5	18.5	17.5	20.5	20.5	20.5	24.5	24.5	24.5	19.5	22.5	18.5	10.5
80% CONFD.	20.5	13.5	13.5	15.5	15.5	16.5	17.5	17.5	18.5	14.5	17.5	14.5	7.5
SAMPLE SIZE	2424	2422	2394	2389	2377	2372	2371	2369	2361	2375	2368	2377	2337
WELDING & HAVENGA REGIONS 11 POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 53 STATIONS													
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN AT 90% CONFD.	64.5	61.5	50.5	51.5	76.5	80.5	74.5	82.5	75.5	66.5	61.5	59.5	25.5
80% CONFD.	51.5	49.5	37.5	39.5	61.5	65.5	59.5	62.5	59.5	53.5	48.5	45.5	18.5
SAMPLE SIZE	3664	3665	3662	3637	3640	3645	3643	3642	3647	3646	3647	3651	3634
10 YEAR MEAN AT 90% CONFD.	40.5	43.5	34.5	35.5	49.5	58.5	57.5	63.5	48.5	47.5	37.5	42.5	17.5
80% CONFD.	31.5	33.5	26.5	27.5	38.5	46.5	45.5	47.5	39.5	36.5	29.5	31.5	12.5
SAMPLE SIZE	3399	3400	3397	3372	3375	3380	3378	3377	3382	3381	3382	3386	3369
15 YEAR MEAN AT 90% CONFD.	30.5	29.5	26.5	27.5	37.5	47.5	47.5	53.5	37.5	37.5	29.5	32.5	13.5
80% CONFD.	23.5	23.5	21.5	21.5	30.5	38.5	37.5	42.5	28.5	29.5	22.5	24.5	9.5
SAMPLE SIZE	3134	3135	3132	3107	3110	3115	3113	3112	3117	3116	3117	3121	3104
20 YEAR MEAN AT 90% CONFD.	24.5	22.5	23.5	23.5	29.5	41.5	39.5	43.5	30.5	30.5	24.5	26.5	11.5
80% CONFD.	18.5	17.5	17.5	17.5	22.5	32.5	31.5	35.5	23.5	24.5	18.5	20.5	8.5
SAMPLE SIZE	2869	2870	2867	2842	2845	2850	2848	2847	2852	2851	2852	2856	2839
25 YEAR MEAN AT 90% CONFD.	21.5	18.5	20.5	20.5	23.5	37.5	34.5	38.5	26.5	24.5	20.5	23.5	10.5
80% CONFD.	16.5	14.5	15.5	15.5	16.5	29.5	26.5	31.5	19.5	19.5	15.5	17.5	7.5
SAMPLE SIZE	2604	2605	2602	2577	2580	2585	2583	2582	2587	2586	2587	2591	2574
30 YEAR MEAN AT 90% CONFD.	19.5	16.5	16.5	18.5	19.5	33.5	29.5	32.5	23.5	19.5	17.5	20.5	9.5
80% CONFD.	14.5	12.5	12.5	13.5	14.5	26.5	22.5	27.5	17.5	15.5	13.5	15.5	6.5
SAMPLE SIZE	2339	2340	2337	2312	2315	2320	2318	2317	2322	2321	2322	2326	2309
35 YEAR MEAN AT 90% CONFD.	17.5	13.5	13.5	16.5	16.5	31.5	24.5	28.5	20.5	15.5	15.5	18.5	8.5
80% CONFD.	13.5	10.5	10.5	11.5	12.5	24.5	19.5	24.5	15.5	12.5	12.5	14.5	6.5
SAMPLE SIZE	2074	2075	2072	2047	2050	2055	2053	2052	2057	2056	2057	2061	2044
40 YEAR MEAN AT 90% CONFD.	15.5	12.5	11.5	14.5	14.5	27.5	20.5	25.5	19.5	13.5	13.5	16.5	7.5
80% CONFD.	12.5	9.5	8.5	10.5	10.5	22.5	15.5	22.5	14.5	10.5	10.5	12.5	5.5
SAMPLE SIZE	1809	1810	1807	1782	1785	1790	1788	1787	1792	1791	1792	1796	1779

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 12												POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 143 STATIONS											
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN										
5 YEAR MEAN AT 90% CONFD.	63.5	60.5	52.5	52.5	71.5	80.5	78.5	90.5	81.5	65.5	65.5	63.5	25.5										
80% CONFD.	49.5	46.5	39.5	40.5	56.5	63.5	62.5	70.5	64.5	51.5	51.5	49.5	18.5										
SAMPLE SIZE	10187	10183	10168	10118	10109	10098	10105	10098	10109	10127	10127	10131	10071										
10 YEAR MEAN AT 90% CONFD.	43.5	42.5	36.5	37.5	51.5	58.5	58.5	72.5	60.5	46.5	41.5	45.5	17.5										
80% CONFD.	34.5	33.5	28.5	29.5	39.5	46.5	46.5	54.5	47.5	34.5	32.5	35.5	13.5										
SAMPLE SIZE	9472	9468	9453	9403	9394	9383	9390	9383	9394	9412	9412	9416	9356										
15 YEAR MEAN AT 90% CONFD.	33.5	31.5	27.5	29.5	41.5	49.5	50.5	60.5	49.5	36.5	33.5	34.5	13.5										
80% CONFD.	25.5	24.5	21.5	23.5	32.5	39.5	39.5	47.5	38.5	28.5	26.5	26.5	10.5										
SAMPLE SIZE	8757	8753	8738	8688	8679	8668	8675	8668	8679	8697	8697	8701	8641										
20 YEAR MEAN AT 90% CONFD.	27.5	24.5	22.5	24.5	34.5	42.5	42.5	50.5	42.5	30.5	26.5	27.5	11.5										
80% CONFD.	20.5	18.5	18.5	19.5	27.5	34.5	33.5	40.5	31.5	23.5	20.5	20.5	7.5										
SAMPLE SIZE	8042	8038	8023	7973	7964	7953	7960	7953	7964	7982	7982	7986	7926										
25 YEAR MEAN AT 90% CONFD.	23.5	19.5	18.5	21.5	28.5	38.5	35.5	43.5	35.5	23.5	23.5	23.5	10.5										
80% CONFD.	17.5	15.5	14.5	16.5	22.5	30.5	26.5	35.5	26.5	18.5	18.5	17.5	7.5										
SAMPLE SIZE	7327	7323	7308	7258	7249	7238	7245	7238	7249	7267	7267	7271	7211										
30 YEAR MEAN AT 90% CONFD.	19.5	16.5	15.5	18.5	24.5	34.5	29.5	37.5	30.5	19.5	20.5	20.5	9.5										
80% CONFD.	15.5	12.5	11.5	14.5	18.5	27.5	22.5	29.5	23.5	15.5	15.5	15.5	7.5										
SAMPLE SIZE	6612	6608	6593	6543	6534	6523	6530	6523	6534	6552	6552	6556	6496										
35 YEAR MEAN AT 90% CONFD.	17.5	13.5	12.5	16.5	20.5	30.5	25.5	32.5	27.5	16.5	18.5	19.5	8.5										
80% CONFD.	13.5	9.5	10.5	12.5	15.5	25.5	18.5	26.5	21.5	12.5	14.5	14.5	6.5										
SAMPLE SIZE	5897	5893	5878	5828	5819	5808	5815	5808	5819	5837	5837	5841	5781										
40 YEAR MEAN AT 90% CONFD.	16.5	12.5	11.5	14.5	18.5	27.5	22.5	30.5	26.5	14.5	16.5	17.5	7.5										
80% CONFD.	12.5	9.5	8.5	10.5	13.5	22.5	16.5	24.5	20.5	10.5	12.5	12.5	5.5										
SAMPLE SIZE	5182	5178	5163	5113	5104	5093	5100	5093	5104	5122	5122	5126	5066										
WELDING & HAVENGA REGIONS 14												POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 58 STATIONS											
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN										
5 YEAR MEAN AT 90% CONFD.	100.5	87.5	66.5	64.5	76.5	82.5	92.5	100.5	94.5	85.5	94.5	86.5	35.5										
80% CONFD.	80.5	69.5	50.5	49.5	59.5	66.5	74.5	79.5	79.5	68.5	75.5	68.5	25.5										
SAMPLE SIZE	3274	3281	3268	3252	3251	3250	3250	3243	3237	3246	3259	3258	3221										
10 YEAR MEAN AT 90% CONFD.	83.5	66.5	47.5	44.5	47.5	61.5	59.5	80.5	70.5	66.5	63.5	62.5	26.5										
80% CONFD.	62.5	51.5	38.5	33.5	35.5	48.5	45.5	62.5	55.5	52.5	50.5	46.5	18.5										
SAMPLE SIZE	2984	2991	2978	2962	2961	2960	2960	2953	2947	2956	2969	2965	2931										
15 YEAR MEAN AT 90% CONFD.	62.5	47.5	38.5	35.5	36.5	50.5	48.5	66.5	54.5	54.5	50.5	43.5	19.5										
80% CONFD.	50.5	37.5	30.5	26.5	27.5	39.5	36.5	53.5	42.5	42.5	38.5	32.5	14.5										
SAMPLE SIZE	2694	2701	2688	2672	2671	2670	2670	2663	2657	2666	2679	2675	2641										
20 YEAR MEAN AT 90% CONFD.	51.5	35.5	30.5	29.5	29.5	40.5	37.5	53.5	43.5	44.5	39.5	32.5	15.5										
80% CONFD.	42.5	29.5	24.5	22.5	22.5	32.5	29.5	43.5	32.5	36.5	28.5	25.5	11.5										
SAMPLE SIZE	2404	2411	2398	2382	2381	2380	2380	2373	2367	2376	2389	2385	2351										
25 YEAR MEAN AT 90% CONFD.	43.5	30.5	26.5	24.5	24.5	33.5	31.5	46.5	36.5	39.5	31.5	28.5	13.5										
80% CONFD.	35.5	23.5	20.5	19.5	18.5	26.5	23.5	35.5	27.5	30.5	23.5	22.5	10.5										
SAMPLE SIZE	2114	2121	2108	2092	2091	2090	2090	2083	2077	2086	2099	2095	2061										
30 YEAR MEAN AT 90% CONFD.	36.5	28.5	24.5	21.5	21.5	27.5	26.5	35.5	30.5	34.5	24.5	23.5	11.5										
80% CONFD.	28.5	19.5	18.5	17.5	16.5	21.5	19.5	27.5	23.5	24.5	18.5	18.5	8.5										
SAMPLE SIZE	1824	1831	1818	1802	1801	1800	1800	1793	1787	1796	1809	1805	1771										
35 YEAR MEAN AT 90% CONFD.	28.5	20.5	20.5	18.5	17.5	23.5	20.5	29.5	25.5	27.5	23.5	20.5	9.5										
80% CONFD.	22.5	16.5	15.5	15.5	13.5	17.5	15.5	22.5	19.5	20.5	16.5	15.5	7.5										
SAMPLE SIZE	1534	1541	1528	1512	1511	1510	1510	1503	1497	1506	1519	1515	1481										
40 YEAR MEAN AT 90% CONFD.	25.5	19.5	15.5	16.5	15.5	19.5	17.5	19.5	22.5	24.5	21.5	17.5	8.5										
80% CONFD.	20.5	14.5	12.5	12.5	11.5	13.5	13.5	14.5	17.5	17.5	15.5	12.5	6.5										
SAMPLE SIZE	1244	1251	1238	1222	1221	1220	1220	1213	1207	1216	1229	1225	1191										

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 15													
POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 158 STATIONS													
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN AT 90% CONFD.	35.5	39.5	38.5	43.5	63.5	82.5	85.5	73.5	56.5	40.5	44.5	37.5	16.5
80% CONFD.	26.5	29.5	29.5	33.5	49.5	63.5	65.5	55.5	44.5	31.5	33.5	28.5	12.5
SAMPLE SIZE	11209	11197	11198	11150	11130	11127	11125	11138	11141	11156	11151	11172	11077
10 YEAR MEAN AT 90% CONFD.	23.5	27.5	26.5	30.5	43.5	57.5	67.5	52.5	35.5	27.5	30.5	25.5	11.5
80% CONFD.	17.5	21.5	20.5	23.5	33.5	45.5	52.5	39.5	27.5	20.5	23.5	19.5	8.5
SAMPLE SIZE	10419	10407	10408	10360	10340	10337	10335	10348	10351	10366	10361	10382	10287
15 YEAR MEAN AT 90% CONFD.	17.5	20.5	19.5	24.5	32.5	48.5	57.5	45.5	26.5	21.5	25.5	19.5	9.5
80% CONFD.	13.5	16.5	15.5	18.5	25.5	38.5	45.5	33.5	20.5	16.5	18.5	14.5	6.5
SAMPLE SIZE	9629	9617	9618	9570	9550	9547	9545	9558	9561	9576	9571	9592	9497
20 YEAR MEAN AT 90% CONFD.	14.5	16.5	16.5	20.5	26.5	37.5	48.5	37.5	22.5	17.5	20.5	15.5	7.5
80% CONFD.	10.5	12.5	12.5	16.5	20.5	29.5	38.5	28.5	17.5	13.5	14.5	11.5	5.5
SAMPLE SIZE	8839	8827	8828	8780	8760	8757	8755	8768	8771	8786	8781	8802	8707
25 YEAR MEAN AT 90% CONFD.	12.5	13.5	13.5	18.5	22.5	32.5	42.5	32.5	18.5	14.5	18.5	13.5	6.5
80% CONFD.	9.5	10.5	10.5	14.5	17.5	24.5	33.5	25.5	13.5	10.5	13.5	10.5	5.5
SAMPLE SIZE	8049	8037	8038	7990	7970	7967	7965	7978	7981	7996	7991	8012	7917
30 YEAR MEAN AT 90% CONFD.	11.5	11.5	11.5	16.5	20.5	29.5	36.5	29.5	15.5	12.5	16.5	11.5	6.5
80% CONFD.	8.5	8.5	8.5	12.5	15.5	22.5	28.5	22.5	11.5	9.5	12.5	9.5	4.5
SAMPLE SIZE	7259	7247	7248	7200	7180	7177	7175	7188	7191	7206	7201	7222	7127
35 YEAR MEAN AT 90% CONFD.	10.5	9.5	9.5	13.5	16.5	25.5	30.5	26.5	13.5	10.5	14.5	10.5	5.5
80% CONFD.	7.5	7.5	7.5	10.5	12.5	19.5	23.5	21.5	10.5	8.5	10.5	8.5	4.5
SAMPLE SIZE	6469	6457	6458	6410	6390	6387	6385	6398	6401	6416	6411	6432	6337
40 YEAR MEAN AT 90% CONFD.	9.5	9.5	8.5	11.5	15.5	23.5	24.5	24.5	12.5	9.5	12.5	9.5	5.5
80% CONFD.	7.5	7.5	6.5	9.5	11.5	17.5	18.5	19.5	9.5	7.5	9.5	7.5	3.5
SAMPLE SIZE	5679	5667	5668	5620	5600	5597	5595	5608	5611	5626	5621	5642	5547
WELDING & HAVENGA REGIONS 16													
POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 313 STATIONS													
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN AT 90% CONFD.	49.5	54.5	46.5	54.5	72.5	81.5	85.5	92.5	87.5	53.5	64.5	51.5	23.5
80% CONFD.	37.5	42.5	36.5	42.5	56.5	65.5	68.5	75.5	69.5	40.5	50.5	39.5	17.5
SAMPLE SIZE	20227	20223	20200	20102	20081	20026	20037	20024	20065	20124	20113	20142	19962
10 YEAR MEAN AT 90% CONFD.	33.5	40.5	32.5	40.5	45.5	63.5	59.5	66.5	62.5	37.5	43.5	36.5	17.5
80% CONFD.	25.5	30.5	25.5	31.5	34.5	49.5	45.5	49.5	47.5	28.5	34.5	27.5	12.5
SAMPLE SIZE	18662	18658	18635	18537	18516	18461	18492	18459	18500	18559	18548	18577	18397
15 YEAR MEAN AT 90% CONFD.	27.5	28.5	23.5	33.5	37.5	53.5	46.5	49.5	48.5	29.5	32.5	26.5	12.5
80% CONFD.	20.5	22.5	18.5	26.5	29.5	42.5	36.5	37.5	36.5	22.5	25.5	20.5	9.5
SAMPLE SIZE	17097	17093	17070	16972	16951	16896	16927	16894	16935	16994	16983	17012	16832
20 YEAR MEAN AT 90% CONFD.	23.5	21.5	18.5	28.5	32.5	46.5	38.5	39.5	40.5	24.5	23.5	20.5	10.5
80% CONFD.	17.5	16.5	14.5	22.5	25.5	37.5	29.5	31.5	30.5	18.5	17.5	15.5	7.5
SAMPLE SIZE	15532	15528	15505	15407	15386	15331	15362	15329	15370	15429	15418	15447	15267
25 YEAR MEAN AT 90% CONFD.	19.5	18.5	14.5	24.5	26.5	41.5	33.5	34.5	34.5	19.5	19.5	17.5	9.5
80% CONFD.	14.5	13.5	11.5	19.5	21.5	32.5	25.5	27.5	25.5	15.5	14.5	13.5	7.5
SAMPLE SIZE	13967	13963	13940	13842	13821	13766	13797	13764	13805	13864	13853	13882	13702
30 YEAR MEAN AT 90% CONFD.	16.5	15.5	12.5	22.5	23.5	37.5	28.5	28.5	29.5	15.5	17.5	15.5	8.5
80% CONFD.	12.5	12.5	9.5	18.5	18.5	29.5	21.5	22.5	21.5	12.5	13.5	11.5	6.5
SAMPLE SIZE	12402	12398	12375	12277	12256	12201	12232	12199	12240	12299	12288	12317	12137
35 YEAR MEAN AT 90% CONFD.	14.5	13.5	11.5	20.5	19.5	32.5	24.5	24.5	25.5	13.5	15.5	13.5	7.5
80% CONFD.	10.5	10.5	8.5	16.5	15.5	25.5	19.5	19.5	19.5	10.5	12.5	10.5	5.5
SAMPLE SIZE	10837	10833	10810	10712	10691	10636	10667	10634	10675	10734	10723	10752	10572
40 YEAR MEAN AT 90% CONFD.	12.5	13.5	9.5	17.5	17.5	29.5	21.5	22.5	21.5	12.5	13.5	11.5	6.5
80% CONFD.	9.5	10.5	7.5	14.5	13.5	22.5	16.5	17.5	16.5	9.5	10.5	8.5	4.5
SAMPLE SIZE	9272	9268	9245	9147	9126	9071	9102	9069	9110	9169	9158	9187	9007

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 17													POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 78 STATIONS												
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN											
5 YEAR MEAN AT 90% CONFD.		65.5	54.5	59.5	63.5	85.5	100.5	100.5	100.5	100.5	78.5	73.5	60.5	31.5											
80% CONFD.		46.5	42.5	44.5	49.5	67.5	86.5	90.5	92.5	88.5	61.5	57.5	45.5	22.5											
SAMPLE SIZE		4589	4583	4591	4544	4540	4520	4509	4518	4518	4541	4551	4553	4484											
10 YEAR MEAN AT 90% CONFD.		46.5	38.5	41.5	46.5	57.5	100.5	89.5	90.5	90.5	57.5	51.5	39.5	23.5											
80% CONFD.		33.5	29.5	33.5	36.5	43.5	83.5	72.5	76.5	73.5	44.5	40.5	31.5	17.5											
SAMPLE SIZE		4199	4193	4201	4154	4150	4130	4119	4128	4128	4151	4161	4163	4094											
15 YEAR MEAN AT 90% CONFD.		35.5	26.5	30.5	37.5	45.5	92.5	74.5	73.5	72.5	43.5	38.5	28.5	16.5											
80% CONFD.		26.5	19.5	23.5	28.5	34.5	76.5	59.5	60.5	55.5	33.5	30.5	22.5	12.5											
SAMPLE SIZE		3809	3803	3811	3764	3760	3740	3729	3738	3738	3761	3771	3773	3704											
20 YEAR MEAN AT 90% CONFD.		28.5	19.5	23.5	31.5	36.5	80.5	52.5	59.5	57.5	34.5	29.5	22.5	13.5											
80% CONFD.		22.5	15.5	17.5	24.5	28.5	66.5	41.5	47.5	47.5	27.5	23.5	16.5	10.5											
SAMPLE SIZE		3419	3413	3421	3374	3370	3350	3339	3348	3348	3371	3381	3383	3314											
25 YEAR MEAN AT 90% CONFD.		24.5	18.5	18.5	26.5	30.5	69.5	44.5	47.5	51.5	25.5	24.5	19.5	12.5											
80% CONFD.		18.5	14.5	14.5	19.5	22.5	56.5	34.5	36.5	41.5	20.5	18.5	14.5	9.5											
SAMPLE SIZE		3029	3023	3031	2984	2980	2960	2949	2958	2958	2981	2991	2993	2924											
30 YEAR MEAN AT 90% CONFD.		19.5	16.5	16.5	22.5	26.5	58.5	43.5	37.5	46.5	20.5	20.5	17.5	11.5											
80% CONFD.		16.5	12.5	13.5	18.5	20.5	48.5	31.5	29.5	36.5	16.5	15.5	13.5	8.5											
SAMPLE SIZE		2639	2633	2641	2594	2590	2570	2559	2568	2568	2591	2601	2603	2534											
35 YEAR MEAN AT 90% CONFD.		17.5	13.5	15.5	19.5	21.5	47.5	33.5	31.5	42.5	17.5	18.5	15.5	9.5											
80% CONFD.		14.5	10.5	12.5	15.5	17.5	39.5	27.5	24.5	32.5	14.5	13.5	12.5	7.5											
SAMPLE SIZE		2249	2243	2251	2204	2200	2180	2169	2178	2178	2201	2211	2213	2144											
40 YEAR MEAN AT 90% CONFD.		16.5	12.5	14.5	17.5	18.5	38.5	28.5	28.5	37.5	14.5	16.5	13.5	8.5											
80% CONFD.		13.5	9.5	10.5	13.5	14.5	30.5	22.5	22.5	28.5	11.5	12.5	10.5	6.5											
SAMPLE SIZE		1859	1853	1861	1814	1810	1790	1779	1788	1788	1811	1821	1823	1754											
WELDING & HAVENGA REGIONS 18													POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 118 STATIONS												
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN											
5 YEAR MEAN AT 90% CONFD.		85.5	74.5	58.5	65.5	83.5	92.5	94.5	100.5	100.5	74.5	78.5	79.5	31.5											
80% CONFD.		65.5	58.5	44.5	50.5	66.5	75.5	75.5	83.5	82.5	58.5	62.5	62.5	22.5											
SAMPLE SIZE		7606	7607	7612	7555	7555	7536	7531	7528	7534	7559	7563	7566	7504											
10 YEAR MEAN AT 90% CONFD.		64.5	55.5	39.5	47.5	50.5	76.5	73.5	79.5	83.5	51.5	52.5	54.5	24.5											
80% CONFD.		47.5	43.5	30.5	37.5	39.5	60.5	55.5	62.5	68.5	39.5	41.5	42.5	16.5											
SAMPLE SIZE		7016	7017	7022	6965	6965	6946	6941	6938	6944	6969	6973	6976	6914											
15 YEAR MEAN AT 90% CONFD.		47.5	38.5	29.5	37.5	41.5	59.5	62.5	67.5	70.5	40.5	40.5	40.5	17.5											
80% CONFD.		35.5	30.5	22.5	28.5	32.5	47.5	46.5	53.5	57.5	31.5	31.5	31.5	12.5											
SAMPLE SIZE		6426	6427	6432	6375	6375	6356	6351	6348	6354	6379	6383	6386	6324											
20 YEAR MEAN AT 90% CONFD.		39.5	28.5	22.5	29.5	35.5	49.5	51.5	53.5	62.5	34.5	30.5	32.5	13.5											
80% CONFD.		29.5	22.5	17.5	23.5	27.5	39.5	37.5	44.5	49.5	26.5	23.5	25.5	10.5											
SAMPLE SIZE		5836	5837	5842	5785	5785	5766	5761	5758	5764	5789	5793	5796	5734											
25 YEAR MEAN AT 90% CONFD.		33.5	23.5	18.5	24.5	28.5	42.5	43.5	48.5	52.5	28.5	25.5	29.5	12.5											
80% CONFD.		24.5	17.5	14.5	19.5	22.5	33.5	32.5	37.5	43.5	21.5	19.5	23.5	9.5											
SAMPLE SIZE		5246	5247	5252	5195	5195	5176	5171	5168	5174	5199	5203	5206	5144											
30 YEAR MEAN AT 90% CONFD.		25.5	20.5	15.5	22.5	25.5	36.5	36.5	40.5	44.5	24.5	22.5	26.5	10.5											
80% CONFD.		19.5	14.5	12.5	17.5	19.5	27.5	28.5	32.5	36.5	18.5	17.5	20.5	8.5											
SAMPLE SIZE		4656	4657	4662	4605	4605	4586	4581	4578	4584	4609	4613	4616	4554											
35 YEAR MEAN AT 90% CONFD.		20.5	16.5	14.5	20.5	20.5	30.5	32.5	36.5	38.5	21.5	20.5	23.5	9.5											
80% CONFD.		15.5	12.5	11.5	16.5	15.5	23.5	24.5	28.5	30.5	16.5	15.5	18.5	7.5											
SAMPLE SIZE		4066	4067	4072	4015	4015	3996	3991	3988	3994	4019	4023	4026	3964											
40 YEAR MEAN AT 90% CONFD.		19.5	15.5	12.5	17.5	17.5	26.5	28.5	31.5	33.5	19.5	17.5	20.5	8.5											
80% CONFD.		15.5	11.5	9.5	14.5	13.5	20.5	21.5	25.5	26.5	15.5	13.5	15.5	6.5											
SAMPLE SIZE		3476	3477	3482	3425	3425	3406	3401	3398	3404	3429	3433	3436	3374											

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 19													POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 119 STATIONS												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN												
5 YEAR MEAN AT 90% CONFD.	35.5	39.5	44.5	58.5	75.5	95.5	81.5	60.5	52.5	36.5	34.5	31.5	15.5												
80% CONFD.	27.5	31.5	35.5	44.5	58.5	69.5	66.5	46.5	41.5	27.5	26.5	25.5	11.5												
SAMPLE SIZE	7581	7572	7575	7554	7547	7534	7538	7555	7550	7547	7559	7560	7499												
10 YEAR MEAN AT 90% CONFD.	24.5	24.5	29.5	43.5	54.5	76.5	67.5	40.5	34.5	25.5	23.5	21.5	10.5												
80% CONFD.	18.5	19.5	22.5	33.5	44.5	57.5	54.5	31.5	27.5	19.5	17.5	16.5	5.5												
SAMPLE SIZE	6986	6977	6980	6959	6952	6939	6943	6960	6955	6952	6964	6965	6904												
15 YEAR MEAN AT 90% CONFD.	19.5	17.5	20.5	34.5	43.5	60.5	59.5	33.5	24.5	21.5	17.5	16.5	7.5												
80% CONFD.	14.5	13.5	16.5	28.5	35.5	46.5	45.5	25.5	19.5	15.5	13.5	12.5	3.5												
SAMPLE SIZE	6391	6382	6385	6364	6357	6344	6348	6365	6360	6357	6369	6370	6309												
20 YEAR MEAN AT 90% CONFD.	16.5	13.5	17.5	30.5	34.5	44.5	47.5	27.5	21.5	17.5	14.5	13.5	6.5												
80% CONFD.	12.5	10.5	13.5	24.5	27.5	35.5	39.5	21.5	16.5	13.5	10.5	10.5	4.5												
SAMPLE SIZE	5796	5787	5790	5769	5762	5749	5753	5770	5765	5762	5774	5775	5714												
25 YEAR MEAN AT 90% CONFD.	13.5	12.5	15.5	27.5	28.5	35.5	39.5	23.5	18.5	15.5	12.5	11.5	5.5												
80% CONFD.	10.5	9.5	11.5	21.5	21.5	28.5	32.5	18.5	14.5	11.5	9.5	8.5	4.5												
SAMPLE SIZE	5201	5192	5195	5174	5167	5154	5158	5175	5170	5167	5179	5180	5119												
30 YEAR MEAN AT 90% CONFD.	12.5	10.5	12.5	23.5	21.5	28.5	33.5	20.5	15.5	13.5	11.5	9.5	5.5												
80% CONFD.	9.5	7.5	9.5	18.5	16.5	23.5	27.5	16.5	12.5	9.5	8.5	7.5	3.5												
SAMPLE SIZE	4606	4597	4600	4579	4572	4559	4563	4580	4575	4572	4584	4585	4524												
35 YEAR MEAN AT 90% CONFD.	10.5	8.5	11.5	20.5	18.5	25.5	28.5	18.5	13.5	11.5	9.5	8.5	4.5												
80% CONFD.	8.5	6.5	8.5	16.5	13.5	20.5	23.5	14.5	10.5	8.5	7.5	6.5	3.5												
SAMPLE SIZE	4011	4002	4005	3984	3977	3964	3968	3985	3980	3977	3989	3990	3929												
40 YEAR MEAN AT 90% CONFD.	9.5	7.5	10.5	17.5	17.5	22.5	24.5	17.5	12.5	9.5	8.5	7.5	3.5												
80% CONFD.	7.5	5.5	7.5	14.5	13.5	18.5	19.5	13.5	9.5	6.5	6.5	5.5	2.5												
SAMPLE SIZE	3416	3407	3410	3389	3382	3369	3373	3390	3385	3382	3394	3395	3334												
WELDING & HAVENGA REGIONS 20A													POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 333 STATIONS												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN												
5 YEAR MEAN AT 90% CONFD.	31.5	37.5	39.5	51.5	76.5	100.5	86.5	81.5	64.5	36.5	31.5	31.5	16.5												
80% CONFD.	24.5	28.5	30.5	39.5	58.5	78.5	69.5	62.5	49.5	28.5	24.5	24.5	12.5												
SAMPLE SIZE	19489	19496	19485	19433	19394	19357	19360	19373	19411	19428	19416	19416	19311												
10 YEAR MEAN AT 90% CONFD.	21.5	24.5	28.5	36.5	51.5	79.5	63.5	53.5	41.5	26.5	20.5	21.5	11.5												
80% CONFD.	16.5	19.5	21.5	28.5	39.5	62.5	50.5	41.5	33.5	20.5	15.5	16.5	8.5												
SAMPLE SIZE	17824	17831	17820	17768	17729	17692	17695	17708	17746	17763	17751	17751	17646												
15 YEAR MEAN AT 90% CONFD.	17.5	17.5	21.5	30.5	37.5	60.5	51.5	41.5	27.5	20.5	14.5	16.5	8.5												
80% CONFD.	12.5	13.5	16.5	23.5	29.5	48.5	39.5	31.5	21.5	15.5	11.5	12.5	6.5												
SAMPLE SIZE	16159	16166	16155	16103	16064	16027	16030	16043	16081	16098	16086	16086	15981												
20 YEAR MEAN AT 90% CONFD.	14.5	14.5	18.5	27.5	30.5	45.5	40.5	32.5	22.5	16.5	11.5	13.5	6.5												
80% CONFD.	10.5	10.5	14.5	20.5	23.5	35.5	30.5	25.5	17.5	12.5	8.5	9.5	4.5												
SAMPLE SIZE	14494	14501	14490	14438	14399	14362	14365	14378	14416	14433	14421	14421	14316												
25 YEAR MEAN AT 90% CONFD.	12.5	12.5	15.5	23.5	25.5	39.5	34.5	26.5	20.5	13.5	10.5	11.5	5.5												
80% CONFD.	9.5	9.5	12.5	18.5	19.5	30.5	25.5	20.5	15.5	10.5	7.5	8.5	4.5												
SAMPLE SIZE	12829	12836	12825	12773	12734	12697	12700	12713	12751	12768	12756	12756	12651												
30 YEAR MEAN AT 90% CONFD.	10.5	9.5	13.5	20.5	21.5	35.5	30.5	22.5	16.5	11.5	8.5	9.5	4.5												
80% CONFD.	7.5	7.5	10.5	15.5	16.5	27.5	22.5	17.5	12.5	8.5	6.5	7.5	3.5												
SAMPLE SIZE	11164	11171	11160	11108	11069	11032	11035	11048	11086	11103	11091	11091	10986												
35 YEAR MEAN AT 90% CONFD.	9.5	8.5	12.5	17.5	19.5	31.5	26.5	20.5	14.5	9.5	7.5	8.5	4.5												
80% CONFD.	6.5	6.5	9.5	13.5	14.5	24.5	20.5	16.5	10.5	7.5	5.5	6.5	3.5												
SAMPLE SIZE	9499	9506	9495	9443	9404	9367	9370	9383	9421	9438	9426	9426	9321												
40 YEAR MEAN AT 90% CONFD.	8.5	7.5	10.5	15.5	17.5	27.5	22.5	18.5	13.5	7.5	6.5	7.5	3.5												
80% CONFD.	6.5	5.5	8.5	12.5	13.5	21.5	17.5	14.5	10.5	5.5	4.5	5.5	2.5												
SAMPLE SIZE	7834	7841	7830	7778	7739	7702	7705	7718	7756	7773	7761	7761	7656												

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 20B POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 412 STATIONS													
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN AT 90% CONFD.	36.5	39.5	42.5	59.5	74.5	100.5	95.5	100.5	88.5	42.5	41.5	36.5	17.5
80% CONFD.	27.5	30.5	32.5	45.5	59.5	79.5	76.5	71.5	33.5	31.5	27.5	13.5	
SAMPLE SIZE	24262	24247	24234	24096	24026	23974	23938	23972	24023	24097	24091	24115	23812
10 YEAR MEAN AT 90% CONFD.	24.5	25.5	28.5	46.5	51.5	80.5	75.5	80.5	58.5	29.5	27.5	24.5	12.5
80% CONFD.	18.5	19.5	22.5	35.5	39.5	64.5	58.5	61.5	46.5	23.5	20.5	18.5	9.5
SAMPLE SIZE	22202	22187	22174	22036	21966	21914	21878	21912	21963	22037	22031	22055	21752
15 YEAR MEAN AT 90% CONFD.	19.5	19.5	21.5	39.5	39.5	63.5	59.5	63.5	38.5	23.5	20.5	19.5	10.5
80% CONFD.	14.5	15.5	17.5	31.5	30.5	49.5	45.5	50.5	29.5	18.5	15.5	14.5	7.5
SAMPLE SIZE	20142	20127	20114	19976	19906	19854	19818	19852	19903	19977	19971	19995	19692
20 YEAR MEAN AT 90% CONFD.	16.5	16.5	17.5	35.5	34.5	51.5	50.5	51.5	31.5	19.5	16.5	15.5	8.5
80% CONFD.	12.5	12.5	13.5	27.5	27.5	38.5	38.5	42.5	24.5	14.5	12.5	11.5	6.5
SAMPLE SIZE	18082	18067	18054	17916	17846	17794	17758	17792	17843	17917	17911	17935	17632
25 YEAR MEAN AT 90% CONFD.	13.5	14.5	14.5	30.5	30.5	47.5	40.5	45.5	28.5	16.5	14.5	13.5	7.5
80% CONFD.	10.5	10.5	11.5	24.5	23.5	34.5	31.5	36.5	21.5	12.5	10.5	9.5	5.5
SAMPLE SIZE	16022	16007	15994	15856	15786	15734	15698	15732	15783	15857	15851	15875	15572
30 YEAR MEAN AT 90% CONFD.	11.5	11.5	13.5	27.5	27.5	41.5	35.5	41.5	23.5	13.5	12.5	11.5	6.5
80% CONFD.	9.5	8.5	10.5	21.5	21.5	31.5	28.5	33.5	17.5	10.5	9.5	8.5	4.5
SAMPLE SIZE	13962	13947	13934	13796	13726	13674	13638	13672	13723	13797	13791	13815	13512
35 YEAR MEAN AT 90% CONFD.	10.5	10.5	11.5	24.5	24.5	36.5	32.5	36.5	20.5	11.5	10.5	9.5	5.5
80% CONFD.	7.5	7.5	8.5	19.5	19.5	27.5	25.5	29.5	14.5	8.5	7.5	7.5	3.5
SAMPLE SIZE	11902	11887	11874	11736	11666	11614	11578	11612	11663	11737	11731	11755	11452
40 YEAR MEAN AT 90% CONFD.	9.5	8.5	10.5	21.5	22.5	31.5	28.5	33.5	19.5	10.5	8.5	8.5	4.5
80% CONFD.	7.5	6.5	7.5	17.5	17.5	23.5	21.5	27.5	14.5	7.5	6.5	6.5	3.5
SAMPLE SIZE	9842	9827	9814	9676	9606	9554	9518	9552	9603	9677	9671	9695	9392
WELDING & HAVENGA REGIONS 20C POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 213 STATIONS													
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR MEAN AT 90% CONFD.	42.5	47.5	48.5	57.5	81.5	95.5	93.5	100.5	94.5	52.5	57.5	44.5	21.5
80% CONFD.	32.5	36.5	37.5	45.5	62.5	75.5	73.5	81.5	80.5	40.5	45.5	34.5	15.5
SAMPLE SIZE	12507	12508	12499	12399	12377	12330	12340	12347	12354	12435	12426	12423	12259
10 YEAR MEAN AT 90% CONFD.	29.5	32.5	33.5	42.5	53.5	79.5	68.5	79.5	71.5	37.5	38.5	30.5	15.5
80% CONFD.	22.5	25.5	25.5	33.5	41.5	60.5	50.5	63.5	54.5	28.5	30.5	22.5	11.5
SAMPLE SIZE	11442	11443	11434	11334	11312	11265	11275	11282	11289	11370	11361	11358	11194
15 YEAR MEAN AT 90% CONFD.	23.5	23.5	24.5	35.5	39.5	64.5	55.5	65.5	48.5	29.5	28.5	23.5	11.5
80% CONFD.	17.5	17.5	19.5	28.5	30.5	48.5	39.5	51.5	37.5	22.5	22.5	17.5	8.5
SAMPLE SIZE	10377	10378	10369	10269	10247	10200	10210	10217	10224	10305	10296	10293	10129
20 YEAR MEAN AT 90% CONFD.	18.5	17.5	18.5	30.5	32.5	52.5	41.5	53.5	39.5	23.5	21.5	19.5	9.5
80% CONFD.	14.5	13.5	14.5	24.5	25.5	39.5	28.5	40.5	30.5	18.5	16.5	14.5	7.5
SAMPLE SIZE	9312	9313	9304	9204	9182	9135	9145	9152	9159	9240	9231	9228	9064
25 YEAR MEAN AT 90% CONFD.	16.5	16.5	14.5	27.5	26.5	44.5	33.5	44.5	34.5	18.5	17.5	16.5	8.5
80% CONFD.	12.5	12.5	11.5	22.5	20.5	32.5	24.5	32.5	27.5	14.5	13.5	12.5	6.5
SAMPLE SIZE	8247	8248	8239	8139	8117	8070	8080	8087	8094	8175	8166	8163	7999
30 YEAR MEAN AT 90% CONFD.	13.5	13.5	12.5	24.5	23.5	36.5	30.5	35.5	29.5	15.5	14.5	14.5	7.5
80% CONFD.	10.5	10.5	9.5	19.5	18.5	27.5	21.5	25.5	22.5	11.5	11.5	11.5	5.5
SAMPLE SIZE	7182	7183	7174	7074	7052	7005	7015	7022	7029	7110	7101	7098	6934
35 YEAR MEAN AT 90% CONFD.	11.5	11.5	11.5	21.5	20.5	30.5	26.5	28.5	26.5	12.5	13.5	12.5	6.5
80% CONFD.	8.5	8.5	8.5	17.5	15.5	22.5	19.5	20.5	20.5	9.5	10.5	9.5	4.5
SAMPLE SIZE	6117	6118	6109	6009	5987	5940	5950	5957	5964	6045	6036	6033	5869
40 YEAR MEAN AT 90% CONFD.	10.5	10.5	10.5	18.5	18.5	25.5	22.5	25.5	25.5	11.5	11.5	10.5	5.5
80% CONFD.	7.5	7.5	7.5	14.5	14.5	18.5	16.5	18.5	19.5	8.5	8.5	8.5	4.5
SAMPLE SIZE	5052	5053	5044	4944	4922	4875	4885	4892	4899	4980	4971	4968	4804

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 21A POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 37 STATIONS												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
5 YEAR MEAN AT 90% CONFD.	46.5	49.5	47.5	50.5	72.5	100.5	86.5	84.5	72.5	44.5	39.5	45.5
80% CONFD.	34.5	38.5	36.5	40.5	55.5	77.5	70.5	67.5	55.5	35.5	30.5	34.5
SAMPLE SIZE	2234	2235	2228	2231	2230	2222	2212	2218	2218	2228	2232	2221
10 YEAR MEAN AT 90% CONFD.	30.5	30.5	35.5	37.5	45.5	87.5	62.5	53.5	48.5	33.5	26.5	27.5
80% CONFD.	23.5	23.5	27.5	29.5	34.5	68.5	49.5	42.5	37.5	25.5	20.5	21.5
SAMPLE SIZE	2049	2050	2043	2046	2045	2037	2027	2033	2033	2043	2047	2036
15 YEAR MEAN AT 90% CONFD.	25.5	22.5	28.5	29.5	33.5	72.5	48.5	42.5	29.5	27.5	19.5	20.5
80% CONFD.	19.5	17.5	22.5	23.5	25.5	55.5	36.5	32.5	23.5	21.5	14.5	15.5
SAMPLE SIZE	1864	1865	1858	1861	1860	1852	1842	1848	1848	1858	1862	1851
20 YEAR MEAN AT 90% CONFD.	21.5	15.5	25.5	24.5	27.5	58.5	37.5	35.5	23.5	23.5	15.5	17.5
80% CONFD.	15.5	11.5	19.5	19.5	20.5	43.5	27.5	26.5	17.5	18.5	11.5	13.5
SAMPLE SIZE	1679	1680	1673	1676	1675	1667	1657	1663	1663	1673	1677	1666
25 YEAR MEAN AT 90% CONFD.	17.5	13.5	21.5	20.5	23.5	51.5	30.5	28.5	20.5	20.5	12.5	15.5
80% CONFD.	13.5	10.5	16.5	16.5	17.5	39.5	23.5	23.5	15.5	14.5	9.5	11.5
SAMPLE SIZE	1494	1495	1488	1491	1490	1482	1472	1478	1478	1488	1492	1481
30 YEAR MEAN AT 90% CONFD.	15.5	11.5	18.5	16.5	19.5	46.5	26.5	24.5	17.5	17.5	10.5	13.5
80% CONFD.	11.5	8.5	14.5	13.5	15.5	36.5	21.5	18.5	13.5	12.5	7.5	10.5
SAMPLE SIZE	1309	1310	1303	1306	1305	1297	1287	1293	1293	1303	1307	1296
35 YEAR MEAN AT 90% CONFD.	14.5	9.5	16.5	15.5	17.5	41.5	24.5	19.5	15.5	13.5	7.5	11.5
80% CONFD.	10.5	7.5	12.5	12.5	13.5	33.5	19.5	15.5	11.5	10.5	6.5	8.5
SAMPLE SIZE	1124	1125	1118	1121	1120	1112	1102	1108	1108	1118	1122	1111
40 YEAR MEAN AT 90% CONFD.	12.5	9.5	14.5	14.5	16.5	35.5	22.5	16.5	13.5	10.5	6.5	9.5
80% CONFD.	9.5	7.5	10.5	11.5	12.5	29.5	16.5	12.5	10.5	7.5	5.5	6.5
SAMPLE SIZE	939	940	933	936	935	927	917	923	923	933	937	926
WELDING & HAVENGA REGIONS 21B POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 392 STATIONS												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
5 YEAR MEAN AT 90% CONFD.	38.5	44.5	47.5	65.5	87.5	100.5	100.5	100.5	96.5	46.5	35.5	39.5
80% CONFD.	29.5	33.5	35.5	52.5	70.5	93.5	92.5	84.5	78.5	35.5	27.5	29.5
SAMPLE SIZE	22575	22579	22566	22410	22307	22259	22247	22247	22368	22420	22421	22440
10 YEAR MEAN AT 90% CONFD.	27.5	29.5	34.5	47.5	58.5	100.5	95.5	90.5	69.5	32.5	23.5	25.5
80% CONFD.	21.5	22.5	26.5	36.5	44.5	85.5	78.5	69.5	55.5	25.5	18.5	19.5
SAMPLE SIZE	20615	20619	20606	20450	20347	20299	20287	20287	20408	20460	20461	20480
15 YEAR MEAN AT 90% CONFD.	22.5	22.5	26.5	40.5	43.5	85.5	83.5	77.5	46.5	25.5	18.5	18.5
80% CONFD.	16.5	17.5	20.5	32.5	33.5	69.5	65.5	60.5	36.5	19.5	13.5	14.5
SAMPLE SIZE	18655	18659	18646	18490	18387	18339	18327	18327	18448	18500	18501	18520
20 YEAR MEAN AT 90% CONFD.	17.5	18.5	21.5	34.5	37.5	69.5	73.5	64.5	39.5	21.5	14.5	14.5
80% CONFD.	13.5	14.5	16.5	27.5	28.5	57.5	55.5	51.5	30.5	16.5	11.5	11.5
SAMPLE SIZE	16695	16699	16686	16530	16427	16379	16367	16367	16488	16540	16541	16560
25 YEAR MEAN AT 90% CONFD.	15.5	15.5	18.5	30.5	31.5	63.5	59.5	54.5	34.5	17.5	12.5	12.5
80% CONFD.	11.5	12.5	14.5	24.5	23.5	50.5	44.5	44.5	26.5	13.5	9.5	9.5
SAMPLE SIZE	14735	14739	14726	14570	14467	14419	14407	14407	14528	14580	14581	14600
30 YEAR MEAN AT 90% CONFD.	13.5	12.5	15.5	26.5	26.5	58.5	50.5	48.5	27.5	14.5	10.5	11.5
80% CONFD.	10.5	9.5	12.5	20.5	20.5	44.5	38.5	39.5	20.5	11.5	7.5	8.5
SAMPLE SIZE	12775	12779	12766	12610	12507	12459	12447	12447	12568	12620	12621	12640
35 YEAR MEAN AT 90% CONFD.	12.5	11.5	13.5	23.5	23.5	49.5	45.5	43.5	21.5	12.5	8.5	9.5
80% CONFD.	9.5	8.5	10.5	18.5	18.5	39.5	34.5	35.5	16.5	9.5	6.5	7.5
SAMPLE SIZE	10815	10819	10806	10650	10547	10499	10487	10487	10608	10660	10661	10680
40 YEAR MEAN AT 90% CONFD.	10.5	9.5	11.5	20.5	21.5	41.5	37.5	40.5	19.5	10.5	7.5	8.5
80% CONFD.	8.5	6.5	8.5	16.5	16.5	33.5	29.5	32.5	14.5	7.5	5.5	6.5
SAMPLE SIZE	8855	8859	8846	8690	8587	8539	8527	8527	8648	8700	8701	8720

SHORT TERM MEAN

WELDING & HAVENGA REGIONS 21C													POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 194 STATIONS												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN												
5 YEAR MEAN AT 90% CONFD.	48.5	50.5	57.5	57.5	78.5	99.5	93.5	94.5	90.5	50.5	40.5	43.5	24.5												
80% CONFD.	37.5	39.5	43.5	43.5	60.5	77.5	72.5	76.5	69.5	38.5	31.5	33.5	18.5												
SAMPLE SIZE	10927	10934	10931	10848	10831	10817	10790	10792	10848	10873	10876	10868	10707												
10 YEAR MEAN AT 90% CONFD.	32.5	32.5	42.5	38.5	52.5	82.5	74.5	74.5	66.5	33.5	26.5	26.5	17.5												
80% CONFD.	24.5	25.5	32.5	29.5	40.5	65.5	56.5	57.5	50.5	25.5	20.5	20.5	12.5												
SAMPLE SIZE	9957	9964	9961	9878	9861	9847	9820	9822	9878	9903	9906	9898	9737												
15 YEAR MEAN AT 90% CONFD.	23.5	23.5	33.5	31.5	39.5	64.5	62.5	61.5	45.5	24.5	20.5	19.5	11.5												
80% CONFD.	18.5	17.5	25.5	23.5	30.5	51.5	47.5	45.5	34.5	19.5	15.5	15.5	8.5												
SAMPLE SIZE	8987	8994	8991	8908	8891	8877	8850	8852	8908	8933	8936	8928	8767												
20 YEAR MEAN AT 90% CONFD.	17.5	17.5	25.5	25.5	32.5	53.5	54.5	50.5	36.5	19.5	17.5	15.5	9.5												
80% CONFD.	13.5	13.5	20.5	19.5	24.5	42.5	41.5	38.5	27.5	15.5	12.5	11.5	6.5												
SAMPLE SIZE	8017	8024	8021	7938	7921	7907	7880	7882	7938	7963	7966	7958	777												
25 YEAR MEAN AT 90% CONFD.	15.5	15.5	20.5	21.5	27.5	47.5	45.5	41.5	32.5	16.5	14.5	13.5	8.5												
80% CONFD.	11.5	12.5	16.5	16.5	20.5	36.5	34.5	31.5	24.5	12.5	10.5	9.5	6.5												
SAMPLE SIZE	7047	7054	7051	6968	6951	6937	6910	6912	6968	6993	6996	6988	6827												
30 YEAR MEAN AT 90% CONFD.	13.5	12.5	17.5	18.5	23.5	43.5	39.5	35.5	25.5	14.5	12.5	11.5	7.5												
80% CONFD.	10.5	9.5	14.5	13.5	17.5	33.5	29.5	26.5	18.5	11.5	9.5	8.5	5.5												
SAMPLE SIZE	6077	6084	6081	5998	5981	5967	5940	5942	5998	6023	6026	6018	5857												
35 YEAR MEAN AT 90% CONFD.	12.5	11.5	15.5	15.5	20.5	40.5	34.5	29.5	18.5	13.5	10.5	9.5	6.5												
80% CONFD.	9.5	8.5	12.5	12.5	15.5	31.5	26.5	22.5	14.5	9.5	7.5	6.5	4.5												
SAMPLE SIZE	5107	5114	5111	5028	5011	4997	4970	4972	5028	5053	5056	5048	4887												
40 YEAR MEAN AT 90% CONFD.	11.5	9.5	13.5	14.5	18.5	35.5	29.5	27.5	17.5	11.5	8.5	8.5	5.5												
80% CONFD.	8.5	7.5	11.5	11.5	13.5	28.5	21.5	18.5	12.5	8.5	6.5	6.5	4.5												
SAMPLE SIZE	4137	4144	4141	4058	4041	4027	4000	4002	4058	4083	4086	4078	3917												
WELDING & HAVENGA REGIONS 22													POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 51 STATIONS												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN												
5 YEAR MEAN AT 90% CONFD.	42.5	48.5	49.5	48.5	73.5	79.5	69.5	51.5	59.5	45.5	40.5	45.5	18.5												
80% CONFD.	32.5	39.5	38.5	38.5	58.5	59.5	51.5	39.5	44.5	37.5	31.5	36.5	14.5												
SAMPLE SIZE	2828	2828	2817	2821	2821	2818	2823	2823	2826	2827	2824	2829	2796												
10 YEAR MEAN AT 90% CONFD.	29.5	32.5	36.5	34.5	52.5	60.5	49.5	32.5	38.5	37.5	25.5	34.5	12.5												
80% CONFD.	23.5	24.5	27.5	26.5	44.5	48.5	38.5	24.5	30.5	32.5	19.5	26.5	9.5												
SAMPLE SIZE	2573	2573	2562	2566	2566	2563	2568	2568	2571	2572	2569	2574	2541												
15 YEAR MEAN AT 90% CONFD.	22.5	20.5	25.5	27.5	40.5	49.5	37.5	23.5	25.5	34.5	17.5	25.5	8.5												
80% CONFD.	18.5	16.5	19.5	21.5	34.5	39.5	29.5	18.5	20.5	29.5	13.5	19.5	6.5												
SAMPLE SIZE	2318	2318	2307	2311	2311	2308	2313	2313	2316	2317	2314	2319	2286												
20 YEAR MEAN AT 90% CONFD.	19.5	14.5	20.5	23.5	34.5	38.5	27.5	20.5	21.5	31.5	13.5	21.5	6.5												
80% CONFD.	16.5	11.5	15.5	19.5	26.5	30.5	21.5	15.5	17.5	26.5	10.5	17.5	4.5												
SAMPLE SIZE	2063	2063	2052	2056	2056	2053	2058	2058	2061	2062	2059	2064	2031												
25 YEAR MEAN AT 90% CONFD.	17.5	13.5	17.5	19.5	31.5	30.5	22.5	16.5	19.5	27.5	11.5	17.5	6.5												
80% CONFD.	15.5	10.5	13.5	16.5	21.5	24.5	16.5	13.5	15.5	22.5	8.5	13.5	4.5												
SAMPLE SIZE	1808	1808	1797	1801	1801	1798	1803	1803	1806	1807	1804	1809	1776												
30 YEAR MEAN AT 90% CONFD.	16.5	11.5	14.5	16.5	24.5	26.5	19.5	14.5	17.5	22.5	9.5	14.5	5.5												
80% CONFD.	14.5	8.5	12.5	13.5	17.5	21.5	14.5	11.5	13.5	18.5	7.5	11.5	3.5												
SAMPLE SIZE	1553	1553	1542	1546	1546	1543	1548	1548	1551	1552	1549	1554	1521												
35 YEAR MEAN AT 90% CONFD.	15.5	9.5	13.5	14.5	19.5	24.5	17.5	13.5	14.5	16.5	8.5	12.5	4.5												
80% CONFD.	13.5	7.5	10.5	11.5	15.5	19.5	12.5	9.5	11.5	13.5	6.5	9.5	3.5												
SAMPLE SIZE	1298	1298	1287	1291	1291	1288	1293	1293	1296	1297	1294	1299	1266												
40 YEAR MEAN AT 90% CONFD.	13.5	8.5	10.5	12.5	17.5	20.5	15.5	12.5	11.5	11.5	6.5	10.5	4.5												
80% CONFD.	12.5	6.5	8.5	10.5	13.5	17.5	10.5	8.5	9.5	9.5	5.5	8.5	3.5												
SAMPLE SIZE	1043	1043	1032	1036	1036	1033	1038	1038	1041	1042	1039	1044	1011												

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 1			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 75 STATIONS												
5 YEAR STD	AT 90% CONFD.	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN	
	80% CONFD.	94.5	87.5	68.5	74.5	70.5	60.5	59.5	94.5	68.5	71.5	80.5	80.5	62.5	
	SAMPLE SIZE	77.5	75.5	55.5	61.5	58.5	49.5	47.5	58.5	57.5	59.5	67.5	67.5	50.5	
		4566	4562	4529	4522	4525	4534	4530	4524	4531	4533	4531	4532	4512	
10 YEAR STD	AT 90% CONFD.	74.5	74.5	48.5	52.5	51.5	39.5	40.5	65.5	51.5	56.5	62.5	63.5	46.5	
	80% CONFD.	59.5	61.5	37.5	40.5	41.5	31.5	32.5	46.5	42.5	44.5	49.5	50.5	38.5	
	SAMPLE SIZE	4191	4187	4154	4147	4150	4159	4155	4149	4156	4158	4156	4157	4137	
15 YEAR STD	AT 90% CONFD.	62.5	59.5	39.5	42.5	42.5	31.5	31.5	52.5	42.5	43.5	51.5	50.5	40.5	
	80% CONFD.	52.5	51.5	30.5	34.5	34.5	23.5	24.5	40.5	34.5	33.5	40.5	39.5	31.5	
	SAMPLE SIZE	3816	3812	3779	3772	3775	3784	3780	3774	3781	3783	3781	3782	3762	
20 YEAR STD	AT 90% CONFD.	54.5	50.5	32.5	36.5	35.5	26.5	26.5	44.5	36.5	32.5	44.5	44.5	35.5	
	80% CONFD.	42.5	41.5	23.5	29.5	27.5	18.5	19.5	35.5	28.5	27.5	35.5	34.5	28.5	
	SAMPLE SIZE	3441	3437	3404	3397	3400	3409	3405	3399	3406	3408	3406	3407	3387	
25 YEAR STD	AT 90% CONFD.	48.5	42.5	25.5	32.5	28.5	23.5	20.5	39.5	30.5	27.5	39.5	39.5	31.5	
	80% CONFD.	37.5	33.5	19.5	25.5	22.5	16.5	16.5	30.5	23.5	22.5	29.5	30.5	24.5	
	SAMPLE SIZE	3066	3062	3029	3022	3025	3034	3030	3024	3031	3033	3031	3032	3012	
30 YEAR STD	AT 90% CONFD.	42.5	36.5	21.5	30.5	24.5	19.5	17.5	34.5	23.5	24.5	34.5	38.5	27.5	
	80% CONFD.	32.5	29.5	16.5	23.5	18.5	13.5	12.5	26.5	18.5	19.5	26.5	27.5	20.5	
	SAMPLE SIZE	2691	2687	2654	2647	2650	2659	2655	2649	2656	2658	2656	2657	2637	
35 YEAR STD	AT 90% CONFD.	38.5	30.5	18.5	27.5	20.5	16.5	14.5	29.5	19.5	21.5	28.5	34.5	24.5	
	80% CONFD.	27.5	23.5	14.5	20.5	15.5	12.5	10.5	23.5	15.5	16.5	22.5	23.5	18.5	
	SAMPLE SIZE	2316	2312	2279	2272	2275	2284	2280	2274	2281	2283	2281	2282	2262	
40 YEAR STD	AT 90% CONFD.	36.5	24.5	17.5	24.5	18.5	15.5	12.5	27.5	16.5	19.5	24.5	29.5	20.5	
	80% CONFD.	24.5	19.5	12.5	20.5	13.5	10.5	9.5	20.5	13.5	14.5	18.5	22.5	15.5	
	SAMPLE SIZE	1941	1937	1904	1897	1900	1909	1905	1899	1906	1908	1906	1907	1887	
WELDING & HAVENGA REGIONS 2			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 115 STATIONS												
5 YEAR STD	AT 90% CONFD.	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN	
	80% CONFD.	85.5	92.5	78.5	68.5	70.5	67.5	58.5	72.5	67.5	67.5	81.5	77.5	64.5	
	SAMPLE SIZE	70.5	77.5	64.5	55.5	56.5	52.5	46.5	55.5	55.5	54.5	63.5	62.5	55.5	
		7735	7732	7708	7662	7669	7674	7678	7677	7680	7682	7691	7685	7615	
10 YEAR STD	AT 90% CONFD.	71.5	74.5	61.5	49.5	51.5	47.5	42.5	53.5	46.5	51.5	56.5	58.5	50.5	
	80% CONFD.	56.5	63.5	50.5	40.5	40.5	36.5	33.5	40.5	37.5	40.5	44.5	45.5	40.5	
	SAMPLE SIZE	7160	7157	7133	7087	7094	7099	7103	7102	7105	7107	7116	7110	7040	
15 YEAR STD	AT 90% CONFD.	64.5	61.5	50.5	40.5	42.5	35.5	36.5	41.5	38.5	42.5	45.5	48.5	39.5	
	80% CONFD.	50.5	51.5	40.5	32.5	33.5	27.5	29.5	32.5	30.5	34.5	36.5	38.5	30.5	
	SAMPLE SIZE	6585	6582	6558	6512	6519	6524	6528	6527	6530	6532	6541	6535	6465	
20 YEAR STD	AT 90% CONFD.	55.5	54.5	41.5	33.5	35.5	29.5	31.5	35.5	32.5	35.5	39.5	43.5	32.5	
	80% CONFD.	43.5	44.5	32.5	26.5	28.5	21.5	24.5	27.5	25.5	28.5	30.5	33.5	25.5	
	SAMPLE SIZE	6010	6007	5983	5937	5944	5949	5953	5952	5955	5957	5966	5960	5890	
25 YEAR STD	AT 90% CONFD.	49.5	47.5	34.5	28.5	30.5	25.5	25.5	32.5	27.5	30.5	32.5	39.5	29.5	
	80% CONFD.	38.5	38.5	27.5	23.5	24.5	18.5	20.5	25.5	20.5	24.5	24.5	30.5	22.5	
	SAMPLE SIZE	5435	5432	5408	5362	5369	5374	5378	5377	5380	5382	5391	5385	5315	
30 YEAR STD	AT 90% CONFD.	45.5	40.5	29.5	24.5	27.5	22.5	20.5	29.5	22.5	25.5	29.5	36.5	26.5	
	80% CONFD.	34.5	33.5	24.5	19.5	21.5	16.5	15.5	23.5	16.5	21.5	21.5	27.5	20.5	
	SAMPLE SIZE	4860	4857	4833	4787	4794	4799	4803	4802	4805	4807	4816	4810	4740	
35 YEAR STD	AT 90% CONFD.	41.5	34.5	26.5	20.5	23.5	19.5	16.5	25.5	17.5	23.5	26.5	34.5	22.5	
	80% CONFD.	31.5	28.5	20.5	16.5	18.5	14.5	13.5	20.5	13.5	17.5	19.5	26.5	17.5	
	SAMPLE SIZE	4285	4282	4258	4212	4219	4224	4228	4227	4230	4232	4241	4235	4165	
40 YEAR STD	AT 90% CONFD.	37.5	29.5	23.5	17.5	20.5	17.5	13.5	23.5	14.5	21.5	23.5	32.5	19.5	
	80% CONFD.	28.5	24.5	18.5	14.5	16.5	12.5	10.5	18.5	11.5	15.5	17.5	24.5	14.5	
	SAMPLE SIZE	3710	3707	3683	3637	3644	3649	3653	3652	3655	3657	3666	3660	3590	

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 3 AND 10 POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 55 STATIONS														
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.	100.5	90.5	75.5	80.5	71.5	74.5	66.5	76.5	79.5	73.5	88.5	89.5	64.5
	80% CONFD.	87.5	78.5	63.5	62.5	59.5	60.5	54.5	61.5	65.5	61.5	76.5	75.5	55.5
	SAMPLE SIZE	3602	3600	3570	3574	3577	3570	3577	3579	3584	3583	3586	3586	3550
10 YEAR STD	AT 90% CONFD.	87.5	73.5	55.5	60.5	54.5	59.5	51.5	55.5	59.5	55.5	73.5	74.5	49.5
	80% CONFD.	74.5	61.5	46.5	48.5	44.5	46.5	40.5	43.5	49.5	42.5	62.5	61.5	40.5
	SAMPLE SIZE	3327	3325	3295	3299	3302	3295	3302	3304	3309	3308	3311	3311	3275
15 YEAR STD	AT 90% CONFD.	74.5	55.5	45.5	52.5	44.5	50.5	44.5	46.5	51.5	45.5	64.5	60.5	41.5
	80% CONFD.	65.5	45.5	35.5	41.5	36.5	40.5	33.5	36.5	40.5	32.5	53.5	48.5	32.5
	SAMPLE SIZE	3052	3050	3020	3024	3027	3020	3027	3029	3034	3033	3036	3036	3000
20 YEAR STD	AT 90% CONFD.	67.5	44.5	36.5	45.5	38.5	44.5	38.5	41.5	44.5	37.5	54.5	50.5	35.5
	80% CONFD.	55.5	34.5	28.5	35.5	31.5	34.5	28.5	30.5	34.5	26.5	45.5	42.5	29.5
	SAMPLE SIZE	2777	2775	2745	2749	2752	2745	2752	2754	2759	2758	2761	2761	2725
25 YEAR STD	AT 90% CONFD.	61.5	37.5	31.5	40.5	34.5	38.5	31.5	36.5	37.5	30.5	47.5	46.5	31.5
	80% CONFD.	51.5	28.5	24.5	31.5	29.5	30.5	23.5	28.5	29.5	22.5	38.5	34.5	25.5
	SAMPLE SIZE	2502	2500	2470	2474	2477	2470	2477	2479	2484	2483	2486	2486	2450
30 YEAR STD	AT 90% CONFD.	56.5	31.5	26.5	35.5	32.5	36.5	25.5	31.5	30.5	26.5	43.5	39.5	27.5
	80% CONFD.	48.5	24.5	20.5	28.5	25.5	27.5	19.5	24.5	24.5	19.5	32.5	29.5	22.5
	SAMPLE SIZE	2227	2225	2195	2199	2202	2195	2202	2204	2209	2208	2211	2211	2175
35 YEAR STD	AT 90% CONFD.	52.5	26.5	23.5	31.5	28.5	35.5	21.5	27.5	24.5	23.5	36.5	34.5	24.5
	80% CONFD.	44.5	21.5	17.5	24.5	22.5	23.5	15.5	21.5	19.5	16.5	28.5	25.5	18.5
	SAMPLE SIZE	1952	1950	1920	1924	1927	1920	1927	1929	1934	1933	1936	1936	1900
40 YEAR STD	AT 90% CONFD.	52.5	24.5	21.5	28.5	24.5	32.5	21.5	24.5	21.5	20.5	31.5	28.5	23.5
	80% CONFD.	41.5	19.5	15.5	21.5	20.5	20.5	14.5	17.5	15.5	13.5	24.5	23.5	16.5
	SAMPLE SIZE	1677	1675	1645	1649	1652	1645	1652	1654	1659	1658	1661	1661	1625
WELDING & HAVENGA REGIONS 4 POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 52 STATIONS														
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.	86.5	78.5	62.5	82.5	73.5	68.5	70.5	100.5	77.5	76.5	78.5	86.5	62.5
	80% CONFD.	68.5	67.5	49.5	64.5	58.5	57.5	58.5	68.5	61.5	59.5	62.5	70.5	50.5
	SAMPLE SIZE	3532	3535	3515	3506	3502	3502	3504	3507	3503	3506	3504	3507	3497
10 YEAR STD	AT 90% CONFD.	62.5	59.5	43.5	59.5	53.5	49.5	52.5	81.5	56.5	57.5	59.5	67.5	45.5
	80% CONFD.	50.5	48.5	32.5	46.5	42.5	39.5	41.5	55.5	45.5	41.5	46.5	54.5	36.5
	SAMPLE SIZE	3272	3275	3255	3246	3242	3242	3244	3247	3243	3246	3244	3247	3237
15 YEAR STD	AT 90% CONFD.	54.5	47.5	32.5	45.5	42.5	38.5	39.5	65.5	43.5	44.5	50.5	56.5	38.5
	80% CONFD.	42.5	35.5	26.5	36.5	32.5	29.5	32.5	47.5	35.5	32.5	38.5	43.5	32.5
	SAMPLE SIZE	3012	3015	2995	2986	2982	2982	2984	2987	2983	2986	2984	2987	2977
20 YEAR STD	AT 90% CONFD.	47.5	38.5	27.5	39.5	35.5	32.5	32.5	53.5	36.5	34.5	41.5	46.5	34.5
	80% CONFD.	36.5	29.5	20.5	30.5	26.5	26.5	25.5	41.5	28.5	25.5	33.5	37.5	27.5
	SAMPLE SIZE	2752	2755	2735	2726	2722	2722	2724	2727	2723	2726	2724	2727	2717
25 YEAR STD	AT 90% CONFD.	42.5	33.5	22.5	33.5	28.5	29.5	29.5	46.5	32.5	29.5	35.5	41.5	30.5
	80% CONFD.	32.5	26.5	17.5	27.5	22.5	21.5	23.5	34.5	25.5	20.5	28.5	30.5	23.5
	SAMPLE SIZE	2492	2495	2475	2466	2462	2462	2464	2467	2463	2466	2464	2467	2457
30 YEAR STD	AT 90% CONFD.	35.5	30.5	19.5	31.5	25.5	25.5	26.5	38.5	28.5	23.5	30.5	36.5	26.5
	80% CONFD.	27.5	23.5	14.5	24.5	19.5	19.5	20.5	28.5	22.5	17.5	24.5	27.5	21.5
	SAMPLE SIZE	2232	2235	2215	2206	2202	2202	2204	2207	2203	2206	2204	2207	2197
35 YEAR STD	AT 90% CONFD.	32.5	26.5	15.5	28.5	21.5	24.5	23.5	32.5	25.5	19.5	25.5	33.5	23.5
	80% CONFD.	24.5	20.5	12.5	22.5	16.5	18.5	17.5	25.5	20.5	14.5	20.5	24.5	19.5
	SAMPLE SIZE	1972	1975	1955	1946	1942	1942	1944	1947	1943	1946	1944	1947	1937
40 YEAR STD	AT 90% CONFD.	28.5	23.5	13.5	27.5	18.5	22.5	18.5	28.5	21.5	17.5	21.5	30.5	21.5
	80% CONFD.	21.5	18.5	10.5	21.5	14.5	14.5	14.5	22.5	17.5	13.5	18.5	21.5	17.5
	SAMPLE SIZE	1712	1715	1695	1686	1682	1682	1684	1687	1683	1686	1684	1687	1677

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 5			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 42 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.		72.5	76.5	67.5	75.5	100.5	90.5	72.5	80.5	80.5	63.5	76.5	78.5	59.5
	80% CONFD.		59.5	65.5	53.5	59.5	72.5	67.5	56.5	65.5	66.5	54.5	62.5	62.5	47.5
	SAMPLE SIZE		2956	2952	2955	2930	2926	2933	2935	2933	2933	2935	2940	2934	2922
10 YEAR STD	AT 90% CONFD.		55.5	54.5	49.5	54.5	68.5	65.5	55.5	64.5	55.5	46.5	56.5	67.5	43.5
	80% CONFD.		44.5	44.5	38.5	42.5	56.5	53.5	43.5	53.5	46.5	36.5	43.5	52.5	34.5
	SAMPLE SIZE		2746	2742	2745	2720	2716	2723	2725	2723	2723	2725	2730	2724	2712
15 YEAR STD	AT 90% CONFD.		47.5	40.5	38.5	44.5	54.5	53.5	43.5	54.5	46.5	38.5	43.5	56.5	37.5
	80% CONFD.		36.5	29.5	30.5	35.5	45.5	44.5	36.5	47.5	37.5	29.5	35.5	47.5	29.5
	SAMPLE SIZE		2536	2532	2535	2510	2506	2513	2515	2513	2513	2515	2520	2514	2502
20 YEAR STD	AT 90% CONFD.		39.5	32.5	32.5	37.5	46.5	50.5	37.5	48.5	39.5	32.5	35.5	50.5	32.5
	80% CONFD.		30.5	24.5	26.5	29.5	36.5	39.5	30.5	42.5	31.5	24.5	30.5	40.5	24.5
	SAMPLE SIZE		2326	2322	2325	2300	2296	2303	2305	2303	2303	2305	2310	2304	2292
25 YEAR STD	AT 90% CONFD.		34.5	27.5	27.5	34.5	38.5	44.5	34.5	43.5	36.5	27.5	31.5	43.5	28.5
	80% CONFD.		26.5	19.5	23.5	25.5	31.5	36.5	28.5	38.5	29.5	21.5	25.5	37.5	21.5
	SAMPLE SIZE		2116	2112	2115	2090	2086	2093	2095	2093	2093	2095	2100	2094	2082
30 YEAR STD	AT 90% CONFD.		29.5	20.5	23.5	31.5	35.5	40.5	31.5	39.5	33.5	24.5	26.5	39.5	25.5
	80% CONFD.		21.5	15.5	20.5	23.5	27.5	31.5	26.5	34.5	25.5	18.5	22.5	33.5	18.5
	SAMPLE SIZE		1906	1902	1905	1880	1876	1883	1885	1883	1883	1885	1890	1884	1872
35 YEAR STD	AT 90% CONFD.		23.5	18.5	20.5	29.5	29.5	38.5	30.5	37.5	31.5	21.5	21.5	36.5	23.5
	80% CONFD.		17.5	14.5	18.5	20.5	22.5	28.5	22.5	32.5	23.5	16.5	18.5	30.5	17.5
	SAMPLE SIZE		1696	1692	1695	1670	1666	1673	1675	1673	1673	1675	1680	1674	1662
40 YEAR STD	AT 90% CONFD.		21.5	17.5	18.5	25.5	24.5	34.5	26.5	35.5	28.5	18.5	19.5	33.5	22.5
	80% CONFD.		16.5	12.5	15.5	19.5	18.5	26.5	19.5	29.5	21.5	13.5	16.5	25.5	15.5
	SAMPLE SIZE		1486	1482	1485	1460	1456	1463	1465	1463	1463	1465	1470	1464	1452
WELDING & HAVENGA REGIONS 6			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 42 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.		78.5	73.5	84.5	66.5	87.5	81.5	69.5	71.5	78.5	75.5	72.5	76.5	61.5
	80% CONFD.		64.5	63.5	64.5	54.5	69.5	61.5	54.5	57.5	63.5	61.5	59.5	64.5	51.5
	SAMPLE SIZE		2858	2853	2852	2840	2830	2836	2836	2837	2838	2844	2843	2841	2824
10 YEAR STD	AT 90% CONFD.		57.5	57.5	60.5	49.5	62.5	70.5	56.5	52.5	59.5	53.5	51.5	61.5	47.5
	80% CONFD.		43.5	46.5	47.5	39.5	52.5	51.5	43.5	41.5	46.5	43.5	39.5	50.5	37.5
	SAMPLE SIZE		2648	2643	2642	2630	2620	2626	2626	2627	2628	2634	2633	2631	2614
15 YEAR STD	AT 90% CONFD.		45.5	44.5	50.5	39.5	55.5	54.5	44.5	43.5	47.5	42.5	40.5	53.5	38.5
	80% CONFD.		35.5	32.5	38.5	33.5	43.5	44.5	36.5	35.5	38.5	33.5	33.5	41.5	30.5
	SAMPLE SIZE		2438	2433	2432	2420	2410	2416	2416	2417	2418	2424	2423	2421	2404
20 YEAR STD	AT 90% CONFD.		37.5	34.5	40.5	34.5	46.5	43.5	38.5	36.5	38.5	34.5	33.5	46.5	33.5
	80% CONFD.		28.5	26.5	32.5	29.5	37.5	36.5	31.5	29.5	29.5	26.5	26.5	35.5	25.5
	SAMPLE SIZE		2228	2223	2222	2210	2200	2206	2206	2207	2208	2214	2213	2211	2194
25 YEAR STD	AT 90% CONFD.		31.5	30.5	36.5	32.5	38.5	35.5	34.5	31.5	33.5	28.5	27.5	39.5	29.5
	80% CONFD.		25.5	22.5	28.5	25.5	30.5	31.5	28.5	25.5	25.5	22.5	22.5	28.5	22.5
	SAMPLE SIZE		2018	2013	2012	2000	1990	1996	1996	1997	1998	2004	2003	2001	1984
30 YEAR STD	AT 90% CONFD.		26.5	25.5	32.5	28.5	31.5	33.5	33.5	27.5	30.5	25.5	22.5	33.5	25.5
	80% CONFD.		20.5	19.5	24.5	23.5	25.5	27.5	26.5	22.5	21.5	19.5	18.5	24.5	20.5
	SAMPLE SIZE		1808	1803	1802	1790	1780	1786	1786	1787	1788	1794	1793	1791	1774
35 YEAR STD	AT 90% CONFD.		23.5	21.5	28.5	26.5	27.5	30.5	31.5	25.5	26.5	23.5	18.5	29.5	22.5
	80% CONFD.		17.5	16.5	22.5	21.5	21.5	25.5	24.5	21.5	21.5	17.5	14.5	22.5	18.5
	SAMPLE SIZE		1598	1593	1592	1580	1570	1576	1576	1577	1578	1584	1583	1581	1564
40 YEAR STD	AT 90% CONFD.		19.5	17.5	26.5	24.5	23.5	28.5	28.5	24.5	24.5	22.5	16.5	25.5	21.5
	80% CONFD.		14.5	15.5	19.5	19.5	19.5	23.5	21.5	20.5	18.5	15.5	12.5	19.5	16.5
	SAMPLE SIZE		1388	1383	1382	1370	1360	1366	1366	1367	1368	1374	1373	1371	1354

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 7			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 48 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.		71.5	70.5	72.5	75.5	92.5	79.5	85.5	90.5	85.5	70.5	72.5	80.5	59.5
	80% CONFD.		57.5	55.5	59.5	58.5	70.5	65.5	72.5	71.5	70.5	55.5	59.5	64.5	47.5
	SAMPLE SIZE		3457	3451	3448	3431	3428	3427	3430	3429	3427	3433	3439	3437	3419
10 YEAR STD	AT 90% CONFD.		50.5	50.5	54.5	55.5	70.5	62.5	71.5	72.5	66.5	54.5	53.5	61.5	41.5
	80% CONFD.		39.5	40.5	42.5	40.5	56.5	50.5	59.5	62.5	54.5	40.5	43.5	48.5	33.5
	SAMPLE SIZE		3217	3211	3208	3191	3188	3187	3190	3189	3187	3193	3199	3197	3179
15 YEAR STD	AT 90% CONFD.		41.5	40.5	44.5	41.5	57.5	50.5	61.5	64.5	57.5	44.5	47.5	52.5	35.5
	80% CONFD.		30.5	31.5	35.5	31.5	44.5	41.5	49.5	55.5	43.5	34.5	36.5	39.5	27.5
	SAMPLE SIZE		2977	2971	2968	2951	2948	2947	2950	2949	2947	2953	2959	2957	2939
20 YEAR STD	AT 90% CONFD.		34.5	33.5	38.5	33.5	44.5	39.5	53.5	55.5	44.5	36.5	40.5	45.5	29.5
	80% CONFD.		25.5	25.5	30.5	27.5	35.5	31.5	44.5	46.5	35.5	28.5	32.5	33.5	22.5
	SAMPLE SIZE		2737	2731	2728	2711	2708	2707	2710	2709	2707	2713	2719	2717	2699
25 YEAR STD	AT 90% CONFD.		29.5	27.5	33.5	30.5	34.5	32.5	50.5	50.5	39.5	30.5	33.5	38.5	25.5
	80% CONFD.		23.5	21.5	27.5	24.5	27.5	26.5	43.5	42.5	31.5	24.5	28.5	30.5	20.5
	SAMPLE SIZE		2497	2491	2488	2471	2468	2467	2470	2469	2467	2473	2479	2477	2459
30 YEAR STD	AT 90% CONFD.		24.5	23.5	31.5	26.5	30.5	28.5	47.5	45.5	33.5	27.5	29.5	33.5	22.5
	80% CONFD.		20.5	17.5	24.5	21.5	24.5	24.5	40.5	38.5	28.5	21.5	24.5	27.5	18.5
	SAMPLE SIZE		2257	2251	2248	2231	2228	2227	2230	2229	2227	2233	2239	2237	2219
35 YEAR STD	AT 90% CONFD.		22.5	18.5	29.5	22.5	26.5	26.5	46.5	42.5	27.5	22.5	25.5	30.5	19.5
	80% CONFD.		17.5	15.5	21.5	18.5	21.5	21.5	38.5	33.5	23.5	18.5	21.5	24.5	15.5
	SAMPLE SIZE		2017	2011	2008	1991	1988	1987	1990	1989	1987	1993	1999	1997	1979
40 YEAR STD	AT 90% CONFD.		18.5	16.5	25.5	19.5	22.5	22.5	43.5	39.5	22.5	19.5	22.5	28.5	17.5
	80% CONFD.		14.5	13.5	19.5	16.5	18.5	17.5	35.5	32.5	18.5	15.5	18.5	19.5	13.5
	SAMPLE SIZE		1777	1771	1768	1751	1748	1747	1750	1749	1747	1753	1759	1757	1739
WELDING & HAVENGA REGIONS 8			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 68 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.		62.5	62.5	71.5	71.5	87.5	84.5	100.5	89.5	73.5	74.5	68.5	65.5	55.5
	80% CONFD.		51.5	49.5	57.5	57.5	67.5	69.5	84.5	79.5	61.5	61.5	54.5	52.5	44.5
	SAMPLE SIZE		4805	4814	4797	4773	4771	4776	4767	4789	4769	4785	4789	4789	4728
10 YEAR STD	AT 90% CONFD.		43.5	40.5	55.5	51.5	67.5	57.5	87.5	84.5	48.5	56.5	53.5	43.5	39.5
	80% CONFD.		34.5	32.5	43.5	40.5	53.5	46.5	78.5	73.5	37.5	45.5	44.5	33.5	31.5
	SAMPLE SIZE		4465	4474	4457	4433	4431	4436	4427	4449	4429	4445	4449	4449	4380
15 YEAR STD	AT 90% CONFD.		34.5	31.5	46.5	40.5	51.5	46.5	80.5	83.5	34.5	48.5	45.5	32.5	30.5
	80% CONFD.		26.5	24.5	36.5	31.5	41.5	37.5	72.5	72.5	27.5	39.5	37.5	25.5	25.5
	SAMPLE SIZE		4125	4134	4117	4093	4091	4096	4087	4109	4089	4105	4109	4109	4040
20 YEAR STD	AT 90% CONFD.		29.5	25.5	39.5	32.5	43.5	38.5	69.5	73.5	26.5	41.5	37.5	25.5	25.5
	80% CONFD.		22.5	19.5	32.5	25.5	35.5	30.5	61.5	67.5	21.5	33.5	32.5	18.5	19.5
	SAMPLE SIZE		3785	3794	3777	3753	3751	3756	3747	3769	3749	3765	3769	3769	3700
25 YEAR STD	AT 90% CONFD.		25.5	20.5	34.5	27.5	37.5	31.5	59.5	68.5	24.5	36.5	32.5	21.5	21.5
	80% CONFD.		19.5	16.5	27.5	22.5	30.5	25.5	50.5	64.5	18.5	27.5	28.5	16.5	17.5
	SAMPLE SIZE		3445	3454	3437	3413	3411	3416	3407	3429	3409	3425	3429	3429	3360
30 YEAR STD	AT 90% CONFD.		23.5	18.5	29.5	24.5	32.5	27.5	49.5	66.5	21.5	29.5	28.5	19.5	19.5
	80% CONFD.		18.5	14.5	23.5	19.5	27.5	21.5	41.5	58.5	17.5	23.5	24.5	15.5	15.5
	SAMPLE SIZE		3105	3114	3097	3073	3071	3076	3067	3089	3069	3085	3089	3089	3020
35 YEAR STD	AT 90% CONFD.		21.5	16.5	24.5	20.5	30.5	24.5	41.5	64.5	18.5	24.5	24.5	17.5	18.5
	80% CONFD.		16.5	12.5	20.5	16.5	24.5	20.5	36.5	53.5	14.5	19.5	21.5	13.5	14.5
	SAMPLE SIZE		2765	2774	2757	2733	2731	2736	2727	2749	2729	2745	2749	2749	2680
40 YEAR STD	AT 90% CONFD.		19.5	14.5	22.5	18.5	25.5	23.5	36.5	58.5	16.5	21.5	21.5	16.5	16.5
	80% CONFD.		14.5	11.5	16.5	14.5	22.5	18.5	30.5	52.5	13.5	16.5	18.5	12.5	12.5
	SAMPLE SIZE		2425	2434	2417	2393	2391	2396	2387	2409	2389	2405	2409	2409	2340

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS		9	AND	13	POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD,										81	STATIONS
5	YEAR STD	AT 90% CONFD.	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN	
		80% CONFD.	77.5	76.5	69.5	68.5	86.5	82.5	84.5	100.5	89.5	77.5	80.5	82.5	63.5	
		SAMPLE SIZE	64.5	66.5	57.5	56.5	70.5	67.5	71.5	82.5	72.5	64.5	67.5	68.5	52.5	
			5259	5257	5229	5224	5212	5207	5206	5204	5196	5210	5203	5212	5172	
10	YEAR STD	AT 90% CONFD.	62.5	58.5	53.5	50.5	66.5	64.5	70.5	82.5	68.5	61.5	60.5	65.5	51.5	
		80% CONFD.	50.5	47.5	42.5	40.5	52.5	50.5	58.5	70.5	55.5	48.5	49.5	52.5	41.5	
		SAMPLE SIZE	4854	4852	4824	4819	4807	4802	4801	4799	4791	4805	4798	4807	4767	
15	YEAR STD	AT 90% CONFD.	51.5	47.5	43.5	39.5	58.5	53.5	58.5	70.5	54.5	51.5	52.5	51.5	44.5	
		80% CONFD.	42.5	35.5	34.5	32.5	44.5	40.5	47.5	60.5	45.5	40.5	41.5	39.5	34.5	
		SAMPLE SIZE	4449	4447	4419	4414	4402	4397	4396	4394	4386	4400	4393	4402	4362	
20	YEAR STD	AT 90% CONFD.	45.5	38.5	36.5	34.5	48.5	45.5	49.5	58.5	46.5	44.5	43.5	43.5	38.5	
		80% CONFD.	36.5	28.5	28.5	25.5	37.5	34.5	38.5	49.5	38.5	33.5	35.5	34.5	29.5	
		SAMPLE SIZE	4044	4042	4014	4009	3997	3992	3991	3989	3981	3995	3988	3997	3957	
25	YEAR STD	AT 90% CONFD.	41.5	34.5	30.5	31.5	39.5	40.5	43.5	49.5	40.5	37.5	36.5	37.5	32.5	
		80% CONFD.	32.5	23.5	24.5	22.5	30.5	29.5	34.5	40.5	31.5	28.5	30.5	30.5	25.5	
		SAMPLE SIZE	3639	3637	3609	3604	3592	3587	3586	3584	3576	3590	3583	3592	3552	
30	YEAR STD	AT 90% CONFD.	37.5	28.5	25.5	28.5	33.5	35.5	40.5	37.5	33.5	31.5	31.5	33.5	28.5	
		80% CONFD.	29.5	21.5	20.5	19.5	26.5	26.5	30.5	30.5	26.5	23.5	25.5	26.5	21.5	
		SAMPLE SIZE	3234	3232	3204	3199	3187	3182	3181	3179	3171	3185	3178	3187	3147	
35	YEAR STD	AT 90% CONFD.	35.5	25.5	22.5	27.5	29.5	31.5	37.5	31.5	28.5	27.5	27.5	28.5	25.5	
		80% CONFD.	27.5	18.5	17.5	17.5	22.5	23.5	26.5	24.5	23.5	18.5	22.5	23.5	18.5	
		SAMPLE SIZE	2829	2827	2799	2794	2782	2777	2776	2774	2766	2780	2773	2782	2742	
40	YEAR STD	AT 90% CONFD.	32.5	24.5	19.5	25.5	27.5	27.5	33.5	26.5	26.5	23.5	23.5	24.5	23.5	
		80% CONFD.	24.5	16.5	14.5	15.5	21.5	20.5	24.5	20.5	20.5	16.5	18.5	20.5	17.5	
		SAMPLE SIZE	2424	2422	2394	2389	2377	2372	2371	2369	2361	2375	2368	2377	2337	
WELDING & HAVENGA REGIONS		11	POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD,										53	STATIONS		
5	YEAR STD	AT 90% CONFD.	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN	
		80% CONFD.	72.5	64.5	67.5	68.5	83.5	81.5	83.5	100.5	80.5	76.5	74.5	70.5	60.5	
		SAMPLE SIZE	60.5	53.5	54.5	54.5	71.5	68.5	71.5	77.5	66.5	62.5	59.5	55.5	50.5	
			3664	3665	3662	3637	3640	3645	3643	3642	3647	3646	3647	3651	3634	
10	YEAR STD	AT 90% CONFD.	54.5	46.5	48.5	52.5	67.5	65.5	69.5	80.5	55.5	57.5	51.5	53.5	46.5	
		80% CONFD.	42.5	35.5	38.5	40.5	53.5	53.5	59.5	65.5	45.5	45.5	41.5	40.5	37.5	
		SAMPLE SIZE	3399	3400	3397	3372	3375	3380	3378	3377	3382	3381	3382	3386	3369	
15	YEAR STD	AT 90% CONFD.	46.5	33.5	38.5	41.5	55.5	53.5	59.5	68.5	43.5	46.5	44.5	41.5	40.5	
		80% CONFD.	36.5	27.5	29.5	32.5	38.5	42.5	50.5	59.5	33.5	38.5	34.5	32.5	32.5	
		SAMPLE SIZE	3134	3135	3132	3107	3110	3115	3113	3112	3117	3116	3117	3121	3104	
20	YEAR STD	AT 90% CONFD.	38.5	25.5	31.5	35.5	42.5	43.5	52.5	59.5	34.5	41.5	36.5	35.5	34.5	
		80% CONFD.	31.5	20.5	25.5	28.5	33.5	34.5	44.5	51.5	27.5	32.5	29.5	26.5	26.5	
		SAMPLE SIZE	2869	2870	2867	2842	2845	2850	2848	2847	2852	2851	2852	2856	2839	
25	YEAR STD	AT 90% CONFD.	34.5	22.5	28.5	31.5	34.5	38.5	49.5	51.5	32.5	33.5	31.5	31.5	29.5	
		80% CONFD.	28.5	17.5	22.5	25.5	27.5	29.5	38.5	44.5	24.5	28.5	26.5	22.5	23.5	
		SAMPLE SIZE	2604	2605	2602	2577	2580	2585	2583	2582	2587	2586	2587	2591	2574	
30	YEAR STD	AT 90% CONFD.	31.5	20.5	25.5	27.5	28.5	34.5	42.5	44.5	30.5	28.5	27.5	26.5	27.5	
		80% CONFD.	25.5	15.5	20.5	21.5	23.5	24.5	35.5	37.5	22.5	24.5	23.5	19.5	21.5	
		SAMPLE SIZE	2339	2340	2337	2312	2315	2320	2318	2317	2322	2321	2322	2326	2309	
35	YEAR STD	AT 90% CONFD.	28.5	18.5	22.5	23.5	24.5	28.5	38.5	40.5	27.5	24.5	23.5	22.5	26.5	
		80% CONFD.	23.5	13.5	17.5	18.5	18.5	21.5	33.5	33.5	20.5	20.5	19.5	17.5	20.5	
		SAMPLE SIZE	2074	2075	2072	2047	2050	2055	2053	2052	2057	2056	2057	2061	2044	
40	YEAR STD	AT 90% CONFD.	25.5	15.5	21.5	21.5	20.5	24.5	37.5	38.5	25.5	20.5	20.5	19.5	24.5	
		80% CONFD.	20.5	11.5	16.5	15.5	17.5	19.5	32.5	32.5	18.5	17.5	17.5	14.5	18.5	
		SAMPLE SIZE	1809	1810	1807	1782	1785	1790	1788	1787	1792	1791	1792	1796	1779	

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 12				POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 143 STATIONS												
				JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.			74.5	63.5	62.5	64.5	83.5	73.5	81.5	100.5	83.5	70.5	73.5	67.5	62.5
	80% CONFD.			60.5	51.5	49.5	51.5	65.5	61.5	68.5	80.5	68.5	55.5	58.5	54.5	50.5
	SAMPLE SIZE			10187	10183	10168	10118	10109	10098	10105	10098	10109	10127	10127	10131	10071
10 YEAR STD	AT 90% CONFD.			59.5	45.5	45.5	45.5	60.5	55.5	65.5	89.5	64.5	48.5	50.5	49.5	45.5
	80% CONFD.			48.5	35.5	36.5	35.5	48.5	42.5	54.5	70.5	52.5	37.5	39.5	39.5	36.5
	SAMPLE SIZE			9472	9468	9453	9403	9394	9383	9390	9383	9394	9412	9412	9416	9356
15 YEAR STD	AT 90% CONFD.			50.5	36.5	36.5	35.5	51.5	46.5	56.5	74.5	53.5	38.5	41.5	40.5	36.5
	80% CONFD.			41.5	27.5	29.5	27.5	38.5	35.5	47.5	63.5	42.5	30.5	32.5	31.5	29.5
	SAMPLE SIZE			8757	8753	8738	8688	8679	8668	8675	8668	8679	8697	8697	8701	8641
20 YEAR STD	AT 90% CONFD.			44.5	28.5	31.5	28.5	42.5	38.5	49.5	64.5	46.5	32.5	34.5	34.5	31.5
	80% CONFD.			36.5	22.5	24.5	22.5	32.5	30.5	41.5	56.5	36.5	26.5	27.5	26.5	25.5
	SAMPLE SIZE			8042	8038	8023	7973	7964	7953	7960	7953	7964	7982	7982	7986	7926
25 YEAR STD	AT 90% CONFD.			39.5	24.5	26.5	23.5	36.5	35.5	43.5	56.5	40.5	27.5	29.5	29.5	25.5
	80% CONFD.			32.5	18.5	21.5	18.5	28.5	26.5	35.5	49.5	31.5	22.5	23.5	22.5	20.5
	SAMPLE SIZE			7327	7323	7308	7258	7249	7238	7245	7238	7249	7267	7267	7271	7211
30 YEAR STD	AT 90% CONFD.			35.5	21.5	23.5	20.5	30.5	31.5	37.5	49.5	36.5	24.5	26.5	25.5	23.5
	80% CONFD.			28.5	16.5	18.5	15.5	24.5	22.5	29.5	41.5	27.5	19.5	20.5	19.5	18.5
	SAMPLE SIZE			6612	6608	6593	6543	6534	6523	6530	6523	6534	6552	6552	6556	6496
35 YEAR STD	AT 90% CONFD.			32.5	18.5	19.5	17.5	26.5	27.5	33.5	41.5	32.5	20.5	23.5	23.5	22.5
	80% CONFD.			26.5	14.5	15.5	13.5	21.5	20.5	27.5	34.5	24.5	16.5	19.5	17.5	17.5
	SAMPLE SIZE			5897	5893	5878	5828	5819	5808	5815	5808	5819	5837	5837	5841	5781
40 YEAR STD	AT 90% CONFD.			32.5	16.5	17.5	15.5	23.5	24.5	32.5	36.5	29.5	18.5	19.5	21.5	19.5
	80% CONFD.			24.5	12.5	14.5	12.5	18.5	18.5	26.5	31.5	22.5	14.5	16.5	15.5	15.5
	SAMPLE SIZE			5182	5178	5163	5113	5104	5093	5100	5093	5104	5122	5122	5126	5066
WELDING & HAVENGA REGIONS 14				POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 58 STATIONS												
				JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.			90.5	82.5	64.5	71.5	75.5	77.5	96.5	98.5	98.5	82.5	79.5	79.5	66.5
	80% CONFD.			76.5	69.5	52.5	59.5	60.5	64.5	78.5	82.5	82.5	68.5	67.5	64.5	55.5
	SAMPLE SIZE			3274	3281	3268	3252	3251	3250	3250	3243	3237	3246	3259	3255	3221
10 YEAR STD	AT 90% CONFD.			76.5	66.5	49.5	51.5	53.5	61.5	74.5	85.5	76.5	66.5	56.5	61.5	49.5
	80% CONFD.			63.5	54.5	39.5	42.5	43.5	50.5	59.5	72.5	63.5	54.5	45.5	47.5	39.5
	SAMPLE SIZE			2984	2991	2978	2962	2961	2960	2960	2953	2947	2956	2969	2965	2931
15 YEAR STD	AT 90% CONFD.			60.5	55.5	39.5	39.5	43.5	48.5	61.5	72.5	60.5	55.5	44.5	48.5	39.5
	80% CONFD.			53.5	45.5	32.5	32.5	32.5	39.5	49.5	63.5	50.5	45.5	35.5	36.5	32.5
	SAMPLE SIZE			2694	2701	2688	2672	2671	2670	2670	2663	2657	2666	2679	2675	2641
20 YEAR STD	AT 90% CONFD.			52.5	47.5	33.5	36.5	34.5	36.5	51.5	63.5	51.5	49.5	35.5	38.5	33.5
	80% CONFD.			44.5	38.5	26.5	28.5	26.5	31.5	41.5	52.5	38.5	36.5	27.5	28.5	28.5
	SAMPLE SIZE			2404	2411	2398	2382	2381	2380	2380	2373	2367	2376	2389	2385	2351
25 YEAR STD	AT 90% CONFD.			45.5	40.5	27.5	33.5	28.5	30.5	41.5	54.5	41.5	41.5	29.5	31.5	29.5
	80% CONFD.			37.5	32.5	22.5	25.5	22.5	24.5	34.5	41.5	32.5	30.5	22.5	23.5	24.5
	SAMPLE SIZE			2114	2121	2108	2092	2091	2090	2090	2083	2077	2086	2099	2095	2061
30 YEAR STD	AT 90% CONFD.			39.5	34.5	23.5	28.5	23.5	24.5	33.5	45.5	33.5	36.5	23.5	25.5	24.5
	80% CONFD.			31.5	26.5	19.5	22.5	18.5	20.5	28.5	34.5	27.5	24.5	18.5	19.5	20.5
	SAMPLE SIZE			1824	1831	1818	1802	1801	1800	1800	1793	1787	1796	1809	1805	1771
35 YEAR STD	AT 90% CONFD.			34.5	28.5	19.5	25.5	18.5	20.5	27.5	39.5	27.5	30.5	20.5	22.5	21.5
	80% CONFD.			28.5	23.5	15.5	18.5	14.5	15.5	22.5	26.5	22.5	19.5	16.5	17.5	18.5
	SAMPLE SIZE			1534	1541	1528	1512	1511	1510	1510	1503	1497	1506	1519	1515	1481
40 YEAR STD	AT 90% CONFD.			32.5	26.5	16.5	24.5	17.5	17.5	21.5	28.5	22.5	24.5	19.5	21.5	19.5
	80% CONFD.			25.5	20.5	12.5	16.5	13.5	12.5	17.5	19.5	18.5	16.5	14.5	14.5	16.5
	SAMPLE SIZE			1244	1251	1238	1222	1221	1220	1220	1213	1207	1216	1229	1225	1191

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 15		POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 158 STATIONS												
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.	63.5	60.5	70.5	65.5	81.5	96.5	100.5	87.5	76.5	67.5	72.5	60.5	59.5
	80% CONFD.	50.5	49.5	56.5	52.5	67.5	76.5	81.5	68.5	60.5	53.5	57.5	48.5	48.5
	SAMPLE SIZE	11209	11197	11198	11150	11130	11127	11125	11138	11141	11156	11151	11172	11077
10 YEAR STD	AT 90% CONFD.	43.5	42.5	49.5	45.5	64.5	77.5	81.5	68.5	54.5	46.5	52.5	41.5	43.5
	80% CONFD.	35.5	34.5	39.5	36.5	51.5	58.5	70.5	55.5	42.5	36.5	42.5	31.5	35.5
	SAMPLE SIZE	10419	10407	10408	10360	10340	10337	10335	10348	10351	10366	10361	10382	10287
15 YEAR STD	AT 90% CONFD.	35.5	35.5	40.5	35.5	55.5	68.5	73.5	63.5	45.5	38.5	44.5	31.5	36.5
	80% CONFD.	28.5	27.5	32.5	27.5	42.5	49.5	62.5	50.5	33.5	30.5	35.5	24.5	28.5
	SAMPLE SIZE	9629	9617	9618	9570	9550	9547	9545	9558	9561	9576	9571	9592	9497
20 YEAR STD	AT 90% CONFD.	30.5	29.5	33.5	29.5	46.5	57.5	63.5	53.5	39.5	34.5	39.5	25.5	29.5
	80% CONFD.	23.5	22.5	27.5	22.5	36.5	41.5	55.5	44.5	29.5	26.5	30.5	20.5	23.5
	SAMPLE SIZE	8839	8827	8828	8780	8760	8757	8755	8768	8771	8786	8781	8802	8707
25 YEAR STD	AT 90% CONFD.	26.5	25.5	28.5	25.5	38.5	47.5	55.5	46.5	33.5	29.5	33.5	21.5	24.5
	80% CONFD.	20.5	19.5	22.5	19.5	30.5	35.5	47.5	39.5	25.5	23.5	26.5	17.5	19.5
	SAMPLE SIZE	8049	8037	8038	7990	7970	7967	7965	7978	7981	7996	7991	8012	7917
30 YEAR STD	AT 90% CONFD.	23.5	22.5	25.5	21.5	31.5	40.5	48.5	41.5	28.5	26.5	29.5	19.5	21.5
	80% CONFD.	18.5	16.5	18.5	16.5	25.5	31.5	39.5	34.5	21.5	20.5	23.5	15.5	17.5
	SAMPLE SIZE	7259	7247	7248	7200	7180	7177	7175	7188	7191	7206	7201	7222	7127
35 YEAR STD	AT 90% CONFD.	20.5	19.5	21.5	18.5	27.5	35.5	42.5	37.5	25.5	23.5	24.5	16.5	19.5
	80% CONFD.	15.5	14.5	15.5	14.5	22.5	28.5	33.5	30.5	19.5	17.5	20.5	13.5	15.5
	SAMPLE SIZE	6469	6457	6458	6410	6390	6387	6385	6398	6401	6416	6411	6432	6337
40 YEAR STD	AT 90% CONFD.	18.5	17.5	17.5	16.5	24.5	31.5	36.5	34.5	22.5	20.5	20.5	15.5	17.5
	80% CONFD.	14.5	12.5	13.5	12.5	18.5	25.5	27.5	27.5	17.5	16.5	17.5	11.5	14.5
	SAMPLE SIZE	5679	5667	5668	5620	5600	5597	5595	5608	5611	5626	5621	5642	5547
WELDING & HAVENGA REGIONS 16		POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 313 STATIONS												
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.	66.5	63.5	63.5	66.5	72.5	92.5	89.5	96.5	91.5	60.5	68.5	60.5	60.5
	80% CONFD.	53.5	51.5	50.5	53.5	59.5	74.5	73.5	80.5	75.5	47.5	55.5	48.5	47.5
	SAMPLE SIZE	20227	20223	20200	20102	20081	20026	20057	20024	20065	20124	20113	20142	19962
10 YEAR STD	AT 90% CONFD.	50.5	45.5	44.5	46.5	48.5	73.5	68.5	76.5	73.5	42.5	44.5	39.5	41.5
	80% CONFD.	40.5	35.5	35.5	37.5	37.5	57.5	56.5	63.5	56.5	33.5	35.5	31.5	32.5
	SAMPLE SIZE	18662	18658	18635	18537	18516	18461	18492	18459	18500	18559	18548	18577	18397
15 YEAR STD	AT 90% CONFD.	41.5	36.5	33.5	36.5	39.5	63.5	56.5	62.5	55.5	33.5	35.5	30.5	32.5
	80% CONFD.	33.5	28.5	26.5	28.5	31.5	46.5	46.5	49.5	42.5	25.5	27.5	23.5	25.5
	SAMPLE SIZE	17097	17093	17070	16972	16951	16896	16927	16894	16935	16994	16983	17012	16832
20 YEAR STD	AT 90% CONFD.	36.5	30.5	27.5	29.5	33.5	54.5	46.5	52.5	43.5	27.5	28.5	24.5	26.5
	80% CONFD.	28.5	23.5	21.5	22.5	26.5	42.5	38.5	41.5	34.5	21.5	22.5	19.5	20.5
	SAMPLE SIZE	15532	15528	15505	15407	15386	15331	15362	15329	15370	15429	15418	15447	15267
25 YEAR STD	AT 90% CONFD.	31.5	25.5	22.5	24.5	28.5	48.5	41.5	44.5	36.5	23.5	23.5	21.5	21.5
	80% CONFD.	24.5	20.5	17.5	18.5	22.5	37.5	32.5	35.5	28.5	17.5	19.5	16.5	16.5
	SAMPLE SIZE	13967	13963	13940	13842	13821	13766	13797	13764	13805	13864	13853	13882	13702
30 YEAR STD	AT 90% CONFD.	27.5	21.5	19.5	21.5	24.5	41.5	36.5	38.5	31.5	19.5	20.5	18.5	19.5
	80% CONFD.	21.5	17.5	15.5	15.5	19.5	32.5	27.5	30.5	25.5	15.5	16.5	13.5	15.5
	SAMPLE SIZE	12402	12398	12375	12277	12256	12201	12232	12199	12240	12299	12288	12317	12137
35 YEAR STD	AT 90% CONFD.	24.5	19.5	17.5	18.5	20.5	35.5	31.5	33.5	28.5	16.5	17.5	15.5	18.5
	80% CONFD.	19.5	15.5	13.5	13.5	16.5	28.5	23.5	26.5	22.5	12.5	14.5	12.5	14.5
	SAMPLE SIZE	10837	10833	10810	10712	10691	10636	10667	10634	10675	10734	10723	10752	10572
40 YEAR STD	AT 90% CONFD.	23.5	17.5	15.5	15.5	17.5	31.5	26.5	30.5	24.5	14.5	15.5	13.5	16.5
	80% CONFD.	17.5	13.5	12.5	11.5	14.5	25.5	19.5	24.5	19.5	11.5	12.5	10.5	13.5
	SAMPLE SIZE	9272	9268	9245	9147	9126	9071	9102	9069	9110	9169	9158	9187	9007

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 17			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 78 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.		77.5	62.5	61.5	67.5	74.5	97.5	100.5	100.5	96.5	77.5	75.5	70.5	62.5
	80% CONFD.		62.5	49.5	50.5	55.5	61.5	85.5	86.5	91.5	82.5	64.5	64.5	56.5	50.5
	SAMPLE SIZE		4589	4583	4591	4544	4540	4520	4509	4518	4518	4541	4551	4553	4484
10 YEAR STD	AT 90% CONFD.		60.5	44.5	45.5	50.5	54.5	89.5	84.5	90.5	82.5	63.5	57.5	49.5	45.5
	80% CONFD.		48.5	36.5	35.5	39.5	40.5	81.5	73.5	81.5	71.5	50.5	47.5	39.5	34.5
	SAMPLE SIZE		4199	4193	4201	4154	4150	4130	4119	4128	4128	4151	4161	4163	4094
15 YEAR STD	AT 90% CONFD.		49.5	36.5	35.5	40.5	42.5	81.5	68.5	76.5	68.5	52.5	47.5	38.5	38.5
	80% CONFD.		40.5	28.5	27.5	30.5	32.5	73.5	55.5	66.5	56.5	42.5	36.5	29.5	29.5
	SAMPLE SIZE		3809	3803	3811	3764	3760	3740	3729	3738	3738	3761	3771	3773	3704
20 YEAR STD	AT 90% CONFD.		41.5	30.5	28.5	31.5	34.5	69.5	49.5	65.5	53.5	41.5	38.5	29.5	32.5
	80% CONFD.		35.5	24.5	21.5	24.5	27.5	59.5	39.5	50.5	44.5	34.5	29.5	23.5	25.5
	SAMPLE SIZE		3419	3413	3421	3374	3370	3350	3339	3348	3348	3371	3381	3383	3314
25 YEAR STD	AT 90% CONFD.		38.5	27.5	23.5	24.5	28.5	59.5	39.5	50.5	43.5	33.5	33.5	26.5	26.5
	80% CONFD.		33.5	21.5	19.5	20.5	22.5	51.5	32.5	40.5	36.5	28.5	24.5	19.5	22.5
	SAMPLE SIZE		3029	3023	3031	2984	2980	2960	2949	2958	2958	2981	2991	2993	2924
30 YEAR STD	AT 90% CONFD.		35.5	24.5	21.5	21.5	23.5	51.5	34.5	41.5	39.5	27.5	27.5	22.5	24.5
	80% CONFD.		30.5	18.5	17.5	17.5	18.5	42.5	27.5	33.5	31.5	23.5	21.5	16.5	20.5
	SAMPLE SIZE		2639	2633	2641	2594	2590	2570	2559	2568	2568	2591	2601	2603	2534
35 YEAR STD	AT 90% CONFD.		32.5	21.5	19.5	17.5	20.5	41.5	28.5	34.5	37.5	23.5	25.5	20.5	22.5
	80% CONFD.		26.5	16.5	14.5	14.5	15.5	33.5	22.5	28.5	27.5	19.5	19.5	16.5	19.5
	SAMPLE SIZE		2249	2243	2251	2204	2200	2180	2169	2178	2178	2201	2211	2213	2144
40 YEAR STD	AT 90% CONFD.		28.5	19.5	17.5	14.5	16.5	33.5	24.5	30.5	32.5	18.5	22.5	18.5	20.5
	80% CONFD.		24.5	14.5	13.5	11.5	13.5	26.5	17.5	23.5	23.5	15.5	17.5	15.5	17.5
	SAMPLE SIZE		1859	1853	1861	1814	1810	1790	1779	1788	1788	1811	1821	1823	1754
WELDING & HAVENGA REGIONS 18			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 118 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.		83.5	74.5	64.5	69.5	77.5	84.5	96.5	100.5	92.5	72.5	78.5	75.5	62.5
	80% CONFD.		69.5	61.5	52.5	57.5	62.5	72.5	80.5	88.5	79.5	58.5	63.5	61.5	51.5
	SAMPLE SIZE		7606	7607	7612	7555	7555	7536	7531	7528	7534	7559	7563	7566	7504
10 YEAR STD	AT 90% CONFD.		69.5	57.5	46.5	49.5	49.5	69.5	78.5	86.5	81.5	52.5	58.5	56.5	43.5
	80% CONFD.		55.5	46.5	37.5	40.5	38.5	55.5	63.5	73.5	69.5	42.5	46.5	43.5	34.5
	SAMPLE SIZE		7016	7017	7022	6965	6965	6946	6941	6938	6944	6969	6973	6976	6914
15 YEAR STD	AT 90% CONFD.		57.5	44.5	36.5	40.5	39.5	57.5	64.5	73.5	69.5	41.5	47.5	45.5	33.5
	80% CONFD.		45.5	36.5	29.5	30.5	30.5	42.5	53.5	62.5	58.5	34.5	38.5	35.5	26.5
	SAMPLE SIZE		6426	6427	6432	6375	6375	6356	6351	6348	6354	6379	6383	6386	6324
20 YEAR STD	AT 90% CONFD.		47.5	35.5	31.5	33.5	33.5	45.5	53.5	62.5	60.5	34.5	40.5	38.5	28.5
	80% CONFD.		39.5	28.5	23.5	25.5	25.5	34.5	45.5	52.5	48.5	27.5	33.5	28.5	23.5
	SAMPLE SIZE		5836	5837	5842	5785	5785	5766	5761	5758	5764	5789	5793	5796	5734
25 YEAR STD	AT 90% CONFD.		41.5	30.5	25.5	28.5	27.5	38.5	47.5	54.5	50.5	28.5	35.5	31.5	22.5
	80% CONFD.		32.5	23.5	20.5	21.5	21.5	30.5	38.5	44.5	40.5	22.5	27.5	25.5	18.5
	SAMPLE SIZE		5246	5247	5252	5195	5195	5176	5171	5168	5174	5199	5203	5206	5144
30 YEAR STD	AT 90% CONFD.		34.5	26.5	22.5	24.5	23.5	33.5	40.5	46.5	41.5	25.5	29.5	28.5	20.5
	80% CONFD.		27.5	20.5	17.5	18.5	18.5	25.5	33.5	37.5	34.5	19.5	23.5	21.5	16.5
	SAMPLE SIZE		4656	4657	4662	4605	4605	4586	4581	4578	4584	4609	4613	4616	4554
35 YEAR STD	AT 90% CONFD.		29.5	22.5	19.5	21.5	19.5	28.5	36.5	37.5	34.5	21.5	24.5	25.5	19.5
	80% CONFD.		25.5	18.5	15.5	16.5	15.5	21.5	29.5	30.5	28.5	16.5	19.5	19.5	15.5
	SAMPLE SIZE		4066	4067	4072	4015	4015	3996	3991	3988	3994	4019	4023	4026	3964
40 YEAR STD	AT 90% CONFD.		28.5	22.5	15.5	18.5	16.5	23.5	31.5	30.5	29.5	18.5	20.5	22.5	16.5
	80% CONFD.		22.5	16.5	13.5	14.5	12.5	18.5	24.5	25.5	25.5	14.5	16.5	17.5	13.5
	SAMPLE SIZE		3476	3477	3482	3425	3425	3406	3401	3398	3404	3429	3433	3436	3374

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 19			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 119 STATIONS												
5 YEAR STD	AT 90% CONFD.	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN	
	80% CONFD.	65.5	65.5	77.5	77.5	85.5	100.5	93.5	72.5	73.5	73.5	68.5	69.5	57.5	
	SAMPLE SIZE	52.5	53.5	63.5	64.5	72.5	85.5	79.5	59.5	59.5	59.5	57.5	55.5	46.5	
		7581	7572	7575	7554	7547	7534	7538	7555	7550	7547	7559	7560	7499	
10 YEAR STD	AT 90% CONFD.	46.5	45.5	56.5	60.5	70.5	94.5	80.5	53.5	53.5	58.5	50.5	48.5	40.5	
	80% CONFD.	37.5	34.5	45.5	50.5	57.5	76.5	70.5	43.5	43.5	46.5	42.5	39.5	31.5	
	SAMPLE SIZE	6986	6977	6980	6959	6952	6939	6943	6960	6955	6952	6964	6965	6904	
15 YEAR STD	AT 90% CONFD.	37.5	32.5	46.5	49.5	58.5	76.5	75.5	42.5	41.5	50.5	42.5	38.5	32.5	
	80% CONFD.	30.5	24.5	37.5	40.5	48.5	66.5	61.5	32.5	34.5	39.5	34.5	29.5	25.5	
	SAMPLE SIZE	6391	6382	6385	6364	6357	6344	6348	6365	6360	6357	6369	6370	6309	
20 YEAR STD	AT 90% CONFD.	31.5	25.5	39.5	43.5	48.5	60.5	63.5	35.5	36.5	43.5	37.5	31.5	27.5	
	80% CONFD.	24.5	19.5	33.5	35.5	39.5	52.5	52.5	25.5	28.5	34.5	30.5	24.5	22.5	
	SAMPLE SIZE	5796	5787	5790	5769	5762	5749	5753	5770	5765	5762	5774	5775	5714	
25 YEAR STD	AT 90% CONFD.	27.5	22.5	34.5	36.5	41.5	50.5	54.5	29.5	30.5	38.5	33.5	26.5	24.5	
	80% CONFD.	19.5	17.5	29.5	29.5	33.5	43.5	47.5	21.5	24.5	30.5	27.5	20.5	18.5	
	SAMPLE SIZE	5201	5192	5195	5174	5167	5154	5158	5175	5170	5167	5179	5180	5119	
30 YEAR STD	AT 90% CONFD.	22.5	20.5	30.5	32.5	34.5	42.5	49.5	26.5	26.5	34.5	28.5	22.5	21.5	
	80% CONFD.	16.5	15.5	24.5	26.5	27.5	36.5	39.5	19.5	20.5	26.5	23.5	17.5	15.5	
	SAMPLE SIZE	4606	4597	4600	4579	4572	4559	4563	4580	4575	4572	4584	4585	4524	
35 YEAR STD	AT 90% CONFD.	19.5	17.5	27.5	30.5	29.5	37.5	42.5	25.5	24.5	32.5	24.5	19.5	17.5	
	80% CONFD.	14.5	13.5	21.5	23.5	23.5	30.5	34.5	17.5	17.5	23.5	20.5	15.5	13.5	
	SAMPLE SIZE	4011	4002	4005	3984	3977	3964	3968	3985	3980	3977	3989	3990	3929	
40 YEAR STD	AT 90% CONFD.	17.5	14.5	23.5	28.5	24.5	32.5	37.5	22.5	21.5	30.5	20.5	17.5	15.5	
	80% CONFD.	12.5	11.5	18.5	22.5	19.5	25.5	31.5	15.5	15.5	20.5	17.5	13.5	11.5	
	SAMPLE SIZE	3416	3407	3410	3389	3382	3369	3373	3390	3385	3382	3394	3395	3334	
WELDING & HAVENGA REGIONS 20A			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 333 STATIONS												
5 YEAR STD	AT 90% CONFD.	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN	
	80% CONFD.	63.5	64.5	70.5	71.5	86.5	100.5	100.5	92.5	80.5	66.5	61.5	57.5	59.5	
	SAMPLE SIZE	51.5	52.5	56.5	58.5	71.5	83.5	86.5	74.5	63.5	53.5	49.5	46.5	47.5	
		19489	19496	19485	19433	19394	19357	19360	19373	19411	19428	19416	19416	19311	
10 YEAR STD	AT 90% CONFD.	45.5	45.5	52.5	56.5	69.5	80.5	84.5	68.5	59.5	47.5	42.5	39.5	41.5	
	80% CONFD.	36.5	36.5	42.5	44.5	56.5	68.5	68.5	53.5	47.5	38.5	33.5	30.5	33.5	
	SAMPLE SIZE	17824	17831	17820	17768	17729	17692	17695	17708	17746	17763	17751	17751	17646	
15 YEAR STD	AT 90% CONFD.	36.5	35.5	42.5	45.5	57.5	67.5	73.5	55.5	46.5	37.5	34.5	30.5	34.5	
	80% CONFD.	28.5	27.5	33.5	36.5	46.5	56.5	55.5	42.5	35.5	29.5	27.5	23.5	27.5	
	SAMPLE SIZE	16159	16166	16155	16103	16064	16027	16030	16043	16081	16098	16086	16086	15981	
20 YEAR STD	AT 90% CONFD.	31.5	28.5	34.5	38.5	48.5	53.5	59.5	45.5	37.5	30.5	29.5	24.5	29.5	
	80% CONFD.	24.5	22.5	27.5	30.5	37.5	41.5	45.5	34.5	29.5	23.5	23.5	19.5	22.5	
	SAMPLE SIZE	14494	14501	14490	14438	14399	14362	14365	14378	14416	14433	14421	14421	14316	
25 YEAR STD	AT 90% CONFD.	27.5	24.5	29.5	30.5	38.5	43.5	47.5	35.5	31.5	26.5	25.5	20.5	24.5	
	80% CONFD.	20.5	18.5	22.5	24.5	30.5	33.5	38.5	27.5	25.5	19.5	20.5	16.5	19.5	
	SAMPLE SIZE	12829	12836	12825	12773	12734	12697	12700	12713	12751	12768	12756	12756	12651	
30 YEAR STD	AT 90% CONFD.	23.5	19.5	25.5	24.5	32.5	35.5	39.5	28.5	25.5	22.5	21.5	17.5	21.5	
	80% CONFD.	18.5	15.5	19.5	20.5	25.5	28.5	32.5	22.5	20.5	16.5	17.5	13.5	16.5	
	SAMPLE SIZE	11164	11171	11160	11108	11069	11032	11035	11048	11086	11103	11091	11091	10986	
35 YEAR STD	AT 90% CONFD.	21.5	16.5	22.5	20.5	27.5	29.5	33.5	24.5	21.5	18.5	18.5	15.5	17.5	
	80% CONFD.	15.5	12.5	16.5	16.5	21.5	24.5	26.5	18.5	17.5	13.5	14.5	11.5	13.5	
	SAMPLE SIZE	9499	9506	9495	9443	9404	9367	9370	9383	9421	9438	9426	9426	9321	
40 YEAR STD	AT 90% CONFD.	19.5	14.5	19.5	16.5	22.5	24.5	27.5	20.5	18.5	15.5	15.5	13.5	14.5	
	80% CONFD.	13.5	11.5	13.5	13.5	18.5	20.5	21.5	16.5	14.5	11.5	12.5	10.5	11.5	
	SAMPLE SIZE	7834	7841	7830	7778	7739	7702	7705	7718	7756	7773	7761	7761	7656	

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 20B		POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 412 STATIONS												
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.	64.5	65.5	63.5	60.5	76.5	100.5	99.5	100.5	88.5	59.5	70.5	58.5	61.5
	80% CONFD.	51.5	52.5	51.5	48.5	63.5	82.5	81.5	79.5	74.5	47.5	57.5	46.5	49.5
	SAMPLE SIZE	24262	24247	24234	24096	24026	23974	23938	23972	24023	24097	24091	24115	23812
10 YEAR STD	AT 90% CONFD.	47.5	45.5	43.5	41.5	57.5	80.5	75.5	79.5	70.5	40.5	52.5	40.5	39.5
	80% CONFD.	37.5	36.5	34.5	32.5	45.5	70.5	62.5	66.5	58.5	32.5	42.5	32.5	31.5
	SAMPLE SIZE	22202	22187	22174	22036	21966	21914	21878	21912	21963	22037	22031	22055	21752
15 YEAR STD	AT 90% CONFD.	37.5	35.5	34.5	31.5	45.5	67.5	61.5	66.5	52.5	31.5	43.5	32.5	32.5
	80% CONFD.	29.5	28.5	26.5	24.5	35.5	55.5	49.5	55.5	42.5	24.5	34.5	25.5	25.5
	SAMPLE SIZE	20142	20127	20114	19976	19906	19854	19818	19852	19903	19977	19971	19995	19692
20 YEAR STD	AT 90% CONFD.	30.5	29.5	27.5	26.5	36.5	54.5	51.5	56.5	42.5	24.5	35.5	26.5	26.5
	80% CONFD.	23.5	22.5	21.5	20.5	28.5	44.5	41.5	48.5	34.5	19.5	30.5	20.5	20.5
	SAMPLE SIZE	18082	18067	18054	17916	17846	17794	17758	17792	17843	17917	17911	17935	17632
25 YEAR STD	AT 90% CONFD.	25.5	23.5	23.5	22.5	29.5	46.5	41.5	51.5	35.5	20.5	31.5	22.5	22.5
	80% CONFD.	19.5	18.5	18.5	16.5	23.5	37.5	34.5	43.5	27.5	15.5	26.5	17.5	17.5
	SAMPLE SIZE	16022	16007	15994	15856	15786	15734	15698	15732	15783	15857	15851	15875	15572
30 YEAR STD	AT 90% CONFD.	22.5	20.5	20.5	18.5	24.5	39.5	35.5	47.5	28.5	17.5	27.5	19.5	18.5
	80% CONFD.	17.5	15.5	15.5	13.5	19.5	32.5	29.5	38.5	22.5	13.5	22.5	15.5	14.5
	SAMPLE SIZE	13962	13947	13934	13796	13726	13674	13638	13672	13723	13797	13791	13815	13512
35 YEAR STD	AT 90% CONFD.	19.5	17.5	17.5	15.5	20.5	33.5	30.5	43.5	24.5	14.5	24.5	16.5	16.5
	80% CONFD.	14.5	13.5	13.5	11.5	17.5	27.5	25.5	34.5	20.5	11.5	19.5	13.5	12.5
	SAMPLE SIZE	11902	11887	11874	11736	11666	11614	11578	11612	11663	11737	11731	11755	11452
40 YEAR STD	AT 90% CONFD.	17.5	14.5	15.5	13.5	18.5	27.5	25.5	41.5	22.5	12.5	20.5	14.5	14.5
	80% CONFD.	13.5	11.5	11.5	10.5	14.5	23.5	22.5	32.5	17.5	10.5	15.5	11.5	11.5
	SAMPLE SIZE	9842	9827	9814	9676	9606	9554	9518	9552	9603	9677	9671	9695	9392
WELDING & HAVENGA REGIONS 20C		POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 213 STATIONS												
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.	63.5	62.5	61.5	66.5	76.5	93.5	93.5	97.5	92.5	62.5	64.5	58.5	57.5
	80% CONFD.	52.5	51.5	48.5	53.5	64.5	78.5	76.5	83.5	80.5	49.5	53.5	46.5	45.5
	SAMPLE SIZE	12507	12508	12499	12399	12377	12330	12340	12347	12354	12435	12426	12423	12259
10 YEAR STD	AT 90% CONFD.	47.5	44.5	41.5	47.5	56.5	77.5	74.5	80.5	71.5	44.5	46.5	39.5	38.5
	80% CONFD.	38.5	35.5	33.5	37.5	44.5	64.5	59.5	68.5	56.5	35.5	36.5	31.5	29.5
	SAMPLE SIZE	11442	11443	11434	11334	11312	11265	11275	11282	11289	11370	11361	11358	11194
15 YEAR STD	AT 90% CONFD.	39.5	33.5	33.5	37.5	41.5	62.5	59.5	70.5	50.5	36.5	36.5	31.5	30.5
	80% CONFD.	31.5	26.5	26.5	29.5	32.5	50.5	45.5	58.5	39.5	28.5	28.5	25.5	24.5
	SAMPLE SIZE	10377	10378	10369	10269	10247	10200	10210	10217	10224	10305	10296	10293	10129
20 YEAR STD	AT 90% CONFD.	33.5	27.5	26.5	29.5	33.5	49.5	46.5	59.5	40.5	29.5	29.5	26.5	25.5
	80% CONFD.	27.5	22.5	20.5	23.5	24.5	39.5	36.5	47.5	31.5	23.5	22.5	20.5	20.5
	SAMPLE SIZE	9312	9313	9304	9204	9182	9135	9145	9152	9159	9240	9231	9228	9064
25 YEAR STD	AT 90% CONFD.	28.5	24.5	21.5	24.5	27.5	40.5	39.5	50.5	33.5	24.5	23.5	21.5	22.5
	80% CONFD.	22.5	18.5	16.5	19.5	20.5	32.5	29.5	37.5	25.5	18.5	18.5	17.5	17.5
	SAMPLE SIZE	8247	8248	8239	8139	8117	8070	8080	8087	8094	8175	8166	8163	7999
30 YEAR STD	AT 90% CONFD.	24.5	19.5	18.5	20.5	23.5	33.5	34.5	40.5	26.5	20.5	19.5	18.5	18.5
	80% CONFD.	19.5	15.5	14.5	15.5	18.5	26.5	25.5	30.5	20.5	15.5	15.5	14.5	14.5
	SAMPLE SIZE	7182	7183	7174	7074	7052	7005	7015	7022	7029	7110	7101	7098	6934
35 YEAR STD	AT 90% CONFD.	21.5	16.5	16.5	16.5	19.5	27.5	29.5	32.5	23.5	16.5	16.5	15.5	15.5
	80% CONFD.	17.5	12.5	12.5	12.5	15.5	21.5	21.5	23.5	17.5	13.5	13.5	12.5	12.5
	SAMPLE SIZE	6117	6118	6109	6009	5987	5940	5950	5957	5964	6045	6036	6033	5869
40 YEAR STD	AT 90% CONFD.	19.5	14.5	14.5	12.5	16.5	23.5	23.5	26.5	21.5	14.5	13.5	13.5	13.5
	80% CONFD.	15.5	11.5	11.5	10.5	13.5	19.5	18.5	19.5	16.5	12.5	11.5	11.5	10.5
	SAMPLE SIZE	5052	5053	5044	4944	4922	4875	4885	4892	4899	4980	4971	4968	4804

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 21A			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 37 STATIONS											
5 YEAR STD	AT 90% CONFD.	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
	80% CONFD.	68.5	70.5	78.5	73.5	87.5	100.5	100.5	91.5	100.5	65.5	68.5	66.5	55.5
	SAMPLE SIZE	56.5	59.5	61.5	58.5	69.5	85.5	93.5	77.5	87.5	52.5	54.5	54.5	44.5
		2234	2235	2228	2231	2230	2222	2212	2218	2218	2228	2232	2221	2187
10 YEAR STD	AT 90% CONFD.	51.5	52.5	60.5	58.5	60.5	82.5	100.5	71.5	100.5	47.5	52.5	46.5	38.5
	80% CONFD.	40.5	43.5	48.5	48.5	49.5	70.5	77.5	57.5	71.5	39.5	38.5	36.5	29.5
	SAMPLE SIZE	2049	2050	2043	2046	2045	2037	2027	2033	2033	2043	2047	2036	2002
15 YEAR STD	AT 90% CONFD.	42.5	40.5	53.5	50.5	50.5	67.5	85.5	57.5	77.5	38.5	39.5	37.5	30.5
	80% CONFD.	35.5	33.5	42.5	40.5	38.5	58.5	63.5	47.5	54.5	30.5	30.5	28.5	23.5
	SAMPLE SIZE	1864	1865	1858	1861	1860	1852	1842	1848	1848	1858	1862	1851	1817
20 YEAR STD	AT 90% CONFD.	36.5	31.5	47.5	42.5	42.5	52.5	76.5	49.5	66.5	32.5	33.5	32.5	23.5
	80% CONFD.	31.5	25.5	39.5	33.5	32.5	43.5	51.5	40.5	45.5	25.5	27.5	25.5	18.5
	SAMPLE SIZE	1679	1680	1673	1676	1675	1667	1657	1663	1663	1673	1677	1666	1632
25 YEAR STD	AT 90% CONFD.	33.5	25.5	42.5	33.5	35.5	44.5	64.5	42.5	62.5	27.5	29.5	28.5	19.5
	80% CONFD.	28.5	20.5	33.5	28.5	29.5	36.5	41.5	34.5	39.5	23.5	23.5	21.5	15.5
	SAMPLE SIZE	1494	1495	1488	1491	1490	1482	1472	1478	1478	1488	1492	1481	1447
30 YEAR STD	AT 90% CONFD.	31.5	20.5	38.5	26.5	30.5	38.5	52.5	36.5	53.5	23.5	25.5	24.5	16.5
	80% CONFD.	26.5	15.5	30.5	22.5	26.5	30.5	33.5	27.5	36.5	18.5	18.5	18.5	13.5
	SAMPLE SIZE	1309	1310	1303	1306	1305	1297	1287	1293	1293	1303	1307	1296	1262
35 YEAR STD	AT 90% CONFD.	29.5	16.5	34.5	22.5	28.5	32.5	39.5	28.5	48.5	19.5	20.5	21.5	13.5
	80% CONFD.	25.5	12.5	27.5	18.5	22.5	28.5	28.5	21.5	30.5	14.5	15.5	14.5	11.5
	SAMPLE SIZE	1124	1125	1118	1121	1120	1112	1102	1108	1108	1118	1122	1111	1077
40 YEAR STD	AT 90% CONFD.	27.5	14.5	30.5	18.5	24.5	27.5	31.5	23.5	38.5	15.5	16.5	17.5	11.5
	80% CONFD.	23.5	11.5	22.5	15.5	19.5	23.5	25.5	16.5	27.5	11.5	13.5	13.5	9.5
	SAMPLE SIZE	939	940	933	936	935	927	917	923	923	933	937	926	892
WELDING & HAVENGA REGIONS 21B			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 392 STATIONS											
5 YEAR STD	AT 90% CONFD.	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
	80% CONFD.	63.5	65.5	65.5	70.5	85.5	100.5	100.5	83.5	83.5	64.5	63.5	59.5	57.5
	SAMPLE SIZE	51.5	53.5	53.5	58.5	71.5	94.5	93.5	86.5	72.5	52.5	51.5	48.5	46.5
		22575	22579	22566	22410	22307	22259	22247	22247	22368	22420	22421	22440	22107
10 YEAR STD	AT 90% CONFD.	45.5	48.5	47.5	50.5	64.5	95.5	93.5	86.5	64.5	45.5	43.5	43.5	38.5
	80% CONFD.	36.5	38.5	37.5	40.5	50.5	83.5	80.5	73.5	54.5	36.5	34.5	33.5	29.5
	SAMPLE SIZE	20615	20619	20606	20450	20347	20299	20287	20287	20408	20460	20461	20480	20147
15 YEAR STD	AT 90% CONFD.	37.5	39.5	36.5	39.5	51.5	81.5	81.5	75.5	45.5	35.5	33.5	33.5	29.5
	80% CONFD.	29.5	32.5	28.5	30.5	40.5	70.5	68.5	60.5	36.5	28.5	26.5	26.5	23.5
	SAMPLE SIZE	18655	18659	18646	18490	18387	18339	18327	18327	18448	18500	18501	18520	18187
20 YEAR STD	AT 90% CONFD.	32.5	33.5	29.5	32.5	42.5	69.5	71.5	61.5	37.5	29.5	26.5	27.5	24.5
	80% CONFD.	25.5	26.5	23.5	25.5	33.5	58.5	54.5	52.5	28.5	22.5	21.5	21.5	18.5
	SAMPLE SIZE	16695	16699	16686	16530	16427	16379	16367	16367	16488	16540	16541	16560	16227
25 YEAR STD	AT 90% CONFD.	28.5	27.5	24.5	27.5	35.5	60.5	55.5	54.5	31.5	24.5	23.5	22.5	19.5
	80% CONFD.	22.5	22.5	18.5	21.5	27.5	48.5	42.5	44.5	24.5	19.5	18.5	17.5	15.5
	SAMPLE SIZE	14735	14739	14726	14570	14467	14419	14407	14407	14528	14580	14581	14600	14267
30 YEAR STD	AT 90% CONFD.	26.5	23.5	20.5	23.5	28.5	49.5	45.5	49.5	24.5	20.5	20.5	18.5	16.5
	80% CONFD.	20.5	18.5	15.5	17.5	23.5	40.5	35.5	39.5	19.5	15.5	16.5	14.5	13.5
	SAMPLE SIZE	12775	12779	12766	12610	12507	12459	12447	12447	12568	12620	12621	12640	12307
35 YEAR STD	AT 90% CONFD.	24.5	19.5	17.5	19.5	23.5	38.5	37.5	46.5	21.5	16.5	17.5	15.5	15.5
	80% CONFD.	17.5	15.5	13.5	15.5	19.5	32.5	30.5	35.5	16.5	12.5	13.5	12.5	11.5
	SAMPLE SIZE	10815	10819	10806	10650	10547	10499	10487	10487	10608	10660	10661	10680	10347
40 YEAR STD	AT 90% CONFD.	22.5	16.5	14.5	16.5	20.5	31.5	30.5	44.5	19.5	14.5	14.5	13.5	13.5
	80% CONFD.	16.5	13.5	11.5	13.5	16.5	27.5	24.5	31.5	14.5	10.5	11.5	10.5	9.5
	SAMPLE SIZE	8855	8859	8846	8690	8587	8539	8527	8527	8648	8700	8701	8720	8387

SHORT TERM STANDARD DEVIATION

WELDING & HAVENGA REGIONS 21C			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 194 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.		73.5	66.5	72.5	73.5	80.5	97.5	96.5	100.5	92.5	64.5	61.5	63.5	56.5
	80% CONFD.		61.5	53.5	60.5	60.5	65.5	81.5	80.5	83.5	74.5	51.5	49.5	52.5	45.5
	SAMPLE SIZE		10927	10934	10931	10848	10831	10817	10790	10792	10848	10873	10876	10868	10707
10 YEAR STD	AT 90% CONFD.		53.5	48.5	56.5	55.5	58.5	78.5	78.5	84.5	73.5	45.5	46.5	44.5	38.5
	80% CONFD.		43.5	37.5	45.5	43.5	46.5	66.5	64.5	71.5	60.5	34.5	35.5	35.5	29.5
	SAMPLE SIZE		9957	9964	9961	9878	9861	9847	9820	9822	9878	9903	9906	9898	9737
15 YEAR STD	AT 90% CONFD.		44.5	37.5	46.5	44.5	46.5	63.5	66.5	73.5	57.5	34.5	36.5	34.5	29.5
	80% CONFD.		35.5	28.5	34.5	35.5	35.5	54.5	53.5	60.5	45.5	25.5	28.5	27.5	23.5
	SAMPLE SIZE		8987	8994	8991	8908	8891	8877	8850	8852	8908	8933	8936	8928	8767
20 YEAR STD	AT 90% CONFD.		37.5	30.5	35.5	36.5	37.5	52.5	57.5	63.5	47.5	26.5	30.5	28.5	24.5
	80% CONFD.		28.5	23.5	27.5	28.5	29.5	41.5	46.5	49.5	38.5	20.5	23.5	22.5	19.5
	SAMPLE SIZE		8017	8024	8021	7938	7921	7907	7880	7882	7938	7963	7966	7958	7797
25 YEAR STD	AT 90% CONFD.		31.5	26.5	28.5	29.5	32.5	42.5	48.5	52.5	42.5	23.5	25.5	24.5	21.5
	80% CONFD.		23.5	19.5	22.5	23.5	25.5	33.5	38.5	41.5	33.5	17.5	20.5	18.5	16.5
	SAMPLE SIZE		7047	7054	7051	6968	6951	6937	6910	6912	6968	6993	6996	6988	6827
30 YEAR STD	AT 90% CONFD.		26.5	21.5	24.5	24.5	29.5	35.5	41.5	44.5	35.5	19.5	22.5	21.5	19.5
	80% CONFD.		19.5	16.5	19.5	19.5	21.5	28.5	33.5	35.5	26.5	14.5	17.5	15.5	14.5
	SAMPLE SIZE		6077	6084	6081	5998	5981	5967	5940	5942	5998	6023	6026	6018	5857
35 YEAR STD	AT 90% CONFD.		21.5	18.5	21.5	20.5	27.5	29.5	35.5	39.5	27.5	16.5	19.5	17.5	17.5
	80% CONFD.		15.5	14.5	16.5	16.5	18.5	24.5	28.5	28.5	21.5	12.5	14.5	13.5	13.5
	SAMPLE SIZE		5107	5114	5111	5028	5011	4997	4970	4972	5028	5053	5056	5048	4887
40 YEAR STD	AT 90% CONFD.		19.5	15.5	17.5	17.5	23.5	24.5	28.5	35.5	25.5	14.5	16.5	14.5	16.5
	80% CONFD.		13.5	11.5	13.5	14.5	16.5	21.5	23.5	23.5	19.5	10.5	12.5	11.5	11.5
	SAMPLE SIZE		4137	4144	4141	4058	4041	4027	4000	4002	4058	4083	4086	4078	3917
WELDING & HAVENGA REGIONS 22			POSSIBLE PERCENTAGE ERROR MADE BY USING A SHORT RECORD, 51 STATIONS												
			JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
5 YEAR STD	AT 90% CONFD.		63.5	70.5	80.5	76.5	87.5	85.5	97.5	78.5	75.5	65.5	67.5	84.5	53.5
	80% CONFD.		51.5	59.5	68.5	63.5	73.5	74.5	75.5	63.5	61.5	53.5	56.5	67.5	42.5
	SAMPLE SIZE		2828	2828	2817	2821	2821	2818	2823	2823	2826	2827	2824	2829	2796
10 YEAR STD	AT 90% CONFD.		45.5	50.5	62.5	57.5	71.5	71.5	75.5	51.5	52.5	52.5	47.5	67.5	33.5
	80% CONFD.		35.5	42.5	51.5	45.5	61.5	62.5	66.5	42.5	44.5	44.5	38.5	54.5	26.5
	SAMPLE SIZE		2573	2573	2562	2566	2566	2563	2568	2568	2571	2572	2569	2574	2541
15 YEAR STD	AT 90% CONFD.		39.5	38.5	53.5	45.5	58.5	59.5	66.5	40.5	40.5	47.5	38.5	58.5	25.5
	80% CONFD.		31.5	30.5	43.5	36.5	51.5	51.5	56.5	32.5	31.5	36.5	31.5	47.5	20.5
	SAMPLE SIZE		2318	2318	2307	2311	2311	2308	2313	2313	2316	2317	2314	2319	2286
20 YEAR STD	AT 90% CONFD.		35.5	29.5	47.5	38.5	54.5	44.5	53.5	35.5	34.5	42.5	29.5	49.5	20.5
	80% CONFD.		28.5	24.5	37.5	30.5	44.5	35.5	45.5	29.5	26.5	33.5	23.5	39.5	16.5
	SAMPLE SIZE		2063	2063	2052	2056	2056	2053	2058	2058	2061	2062	2059	2064	2031
25 YEAR STD	AT 90% CONFD.		32.5	27.5	40.5	33.5	48.5	34.5	44.5	30.5	30.5	37.5	22.5	39.5	18.5
	80% CONFD.		27.5	22.5	34.5	26.5	36.5	26.5	37.5	24.5	23.5	26.5	18.5	31.5	14.5
	SAMPLE SIZE		1808	1808	1797	1801	1801	1798	1803	1803	1806	1807	1804	1809	1776
30 YEAR STD	AT 90% CONFD.		31.5	24.5	38.5	27.5	37.5	28.5	43.5	25.5	26.5	30.5	19.5	31.5	15.5
	80% CONFD.		26.5	20.5	31.5	21.5	26.5	24.5	32.5	20.5	20.5	20.5	15.5	28.5	11.5
	SAMPLE SIZE		1553	1553	1542	1546	1546	1543	1548	1548	1551	1552	1549	1554	1521
35 YEAR STD	AT 90% CONFD.		29.5	22.5	36.5	20.5	30.5	24.5	40.5	23.5	22.5	23.5	16.5	26.5	13.5
	80% CONFD.		25.5	18.5	30.5	18.5	23.5	20.5	22.5	18.5	16.5	16.5	13.5	22.5	10.5
	SAMPLE SIZE		1298	1298	1287	1291	1291	1288	1293	1293	1296	1297	1294	1299	1266
40 YEAR STD	AT 90% CONFD.		29.5	21.5	36.5	17.5	27.5	21.5	25.5	21.5	17.5	20.5	13.5	20.5	12.5
	80% CONFD.		24.5	16.5	27.5	15.5	19.5	17.5	15.5	15.5	13.5	12.5	11.5	18.5	10.5
	SAMPLE SIZE		1043	1043	1032	1036	1036	1033	1033	1038	1038	1041	1042	1039	1044

CHAPTER 11

APPENDIX B

B.1 UNIVERSAL KRIGING

A basic assumption in the Kriging technique is that the process being modelled is a random surface which is generated by some probability density function. In order to make statistical inferences, some assumptions need to be introduced into the model. The random surface is assumed to satisfy the intrinsic hypothesis, namely that the first and second moments of the difference

$$D(x,h) = Z(x+h) - Z(x)$$

where

$Z(x)$ and $Z(x+h)$ = the value assumed by the surface at the positions x and $x+h$ respectively,

depend only on the distance between the points, h , and not on the relative positions of the points on the surface.

These assumptions may be written as:

$$\begin{aligned} E(D(x,h)) &= m(h) \text{ and} \\ E(D(x,h) - m(h))^2 &= 2\gamma(h) \end{aligned}$$

where

the operator E = the expected value and

$\gamma(h)$ = the semivariogram (Olea, 1974).

Davis (1973) showed that the semivariogram may be expressed as

$$\gamma(h) = K(o) - K(h)$$

where

$K(o)$ = the variance at the point of interpolation, and

$K(h)$ = the covariance between points a distance h apart.

The semivariogram tends to zero when h approaches zero, as the covariance becomes more nearly equal to the variance at that point. When h becomes large and two points become more nearly independent from one another and $K(h)$ tends towards $K(o)$. The semivariogram may be used to determine the distance over which the values are interdependent.

In practice $K(h)$ is computed using various directions for the vector h . If different semivariograms are found for the different directions and the distribution of values is anisotropic then the design of the Kriging equation becomes more difficult (Davis, 1973).

The simplest form of Kriging assumes that the mean is constant over the region and that the semivariogram is isotropic. The system of equations then becomes:

$$\sum_{j=1}^n a_j \gamma(x_i - x_j) = \gamma(x_o - x_i) - \lambda$$

where

$$\begin{matrix} i = 1, n \\ \sum_{j=1}^k a_j = 1 \end{matrix}$$

where

x_o = the point of interpolation

λ = a Lagrange multiplier and

a_i = the weights of the k neighbours such that

$$Z(x_0) = \sum_{i=1}^k a_i Z(x_i)$$

The coefficient λ can be used to evaluate the variance of the Kriged estimate ($\sigma_k^2(Z)$)

$$\sigma_k^2(Z) = \lambda + \sum_{j=1}^n a_j (x_j - x_0)$$

This estimate of the probable standard error in the interpolation value provides the major advantage of Kriging and enables mapping the spatial distribution of this error (Davis, 1973).

When the mean changes in a slow, systematic manner over the surface, the assumption of stationarity no longer holds, and the regionalised variable may then be split into the expected value, (or drift) and the residual,

$$W(x) = Z(X) - m(x)$$

where

$W(x)$ = the residual which then satisfies the stationarity assumptions.

In a local area, the drift may be assumed to be

$$m(x) = \sum_{i=1}^n b_i f^i(x)$$

where

f^i = some arbitrary function of x and
 b_i = the n unknown coefficients.

The system of equations then becomes

$$\sum_{j=1}^k \lambda_j \gamma(x_i - x_j) + \sum_{m=1}^n b_m f^m(x_i) + \lambda = \gamma(x_i - x_0)$$

where

$$i = 1, k$$

$$\sum_{i=1}^k \lambda_i f^m(x_i) = f^m(x_0)$$

where

$$m = 1, n$$

$$\sum_{i=1}^k \lambda_i = 1 \quad (\text{Olea, 1974}) \quad (\text{B.1})$$

The solution of the system of equations B.1 implies that the semivariogram is known. However, this implies that the drift at all points on the surface is known. Therefore, several arbitrary drifts must be examined and the drift that produces the most convenient solution is selected (Olea, 1974). Olea (1974) showed that the form of the semivariogram of the estimated residuals is no longer increasing monotonically, but has a parabolic form. The estimation of the drift is biased as the correlations between the residuals are unknown, which in turn causes a bias in the estimation of the semivariogram.

B.2 RIPLEY'S METHOD

Ripley's (1980) method assumes that there are many realisations of the same random surface. The covariances and means may be calculated at the sample points directly and no assumptions of stationarity need to be made, although it is assumed that the same process generated all the

surfaces. As with Kriging, in the presence of a drift, the estimated mean is first removed from the surface.

Suppose that there are n stations with records in year 1, then let

$$\underline{W}_n = (W(x_1), W(x_2), \dots, W(x_n))^T$$

Ripley (1980) shows that there is an optimum linear interpolator, such that

$$W(x) = \lambda^T \underline{W}_n$$

The value of λ at any point x on the surface is given by

$$\lambda = K^{-1} (\underline{k}(x))$$

where

K^{-1} = the covariance matrix of the n predictors

and

$\underline{k}(x)$ = the vector of covariances between the point and the n predictors.

The model then becomes

$$\hat{W}(x) = (\underline{W}_n^T K^{-1}) \underline{k}(x)$$

As K is non-singular and strictly positive definite, there exists a unique lower triangular matrix such that

$$LL^T = K$$

Ripley (1980) put $\underline{y} = K^{-1} \underline{W}_n$.

Then

$$LL^T(\underline{y}) = \underline{W}_n$$

which is solved for $\underline{y}$ since both L and L^T are triangular.

The variance of the interpolation error is

$$\sigma_E^2(x) = \sigma^2(x) - \underline{k}(x)^T K^{-1} \underline{k}(x)$$

where

$$\sigma^2(x) = \text{variance at the point } x.$$

This may be conveniently evaluated by solving for $\underline{e}$ in

$$L\underline{e} = \underline{k}(x)$$

then

$$\sigma_E^2(x) = \sigma^2(x) - \underline{e}^T \underline{e}$$

An implicit assumption in this method is that all the measured points are useful as predictors at every point on the surface. The term $(\underline{W}_n^T K^{-1})$ of B.2 is therefore calculated once per realisation of the surface, which greatly reduces the computational load.

However, this assumption causes the model to carry a number of redundant predictors when estimating the value at a point. Furthermore, K is likely to be ill-conditioned, which makes the solution of B.2 sensitive to small perturbations in K . This in turn is estimated with a degree of uncertainty, and therefore this method will give unreliable results if K is ill-conditioned (Seed et al, 1983)

B.3 GANDIN'S METHOD

The problem of having redundant predictors included in the model may be overcome by the use of the interpolation technique proposed by Gandin (1965) and cited by Creutin and Obled (1982). Three fundamental assumptions are made:

- (a) The events are assumed to be realisations of a unique random process or to have come from the same population.
- (b) Weak stationarity or homogeneity is assumed when choosing the correlation function.
- (c) The mean $m(x)$ and standard deviation $\sigma(x)$ were available separately (Creutin and Obled, 1982).

The Gandin (1965) analysis cited by Creutin and Obled (1982) began with the equation

$$\hat{W}(x) = \frac{W^T}{k} (K^{-1} \underline{k}(x))$$

where

$$\frac{W^T}{k} = k \text{ predictors in the neighbourhood of the point}$$

where

$$k = \text{number of predictors}$$

$(K^{-1} \underline{k}(x))$ was solved for each point on the surface.

The variance of the interpolation error for the standardised deviations is computed by:

$$\sigma_E^2(x) = \sigma^2(x) - \lambda^T \underline{k}(x)$$

(Creutin and Obled, 1982).

Two well known techniques for reducing ill-conditioning in least squares regression are the use of normalised variables and of subset selection (Neter and Wasserman, 1974). Gandin's method is far less susceptible to ill-conditioning than Ripley's (1980) mode. Gandin's (1965) model has an advantage over the classical least squares regression in that all the data are used to calculate the means and covariances. In addition a positive, definite function may be fitted to the correlation-distance relationship between gauges to estimate the correlations not available directly from the data.

CHAPTER 11

APPENDIX C

Personnel from a number of institutions have requested processed monthly rainfall data and/or physiographic information from this project. These institutions are inter alia:

- Agricultural Production Department, University of Natal
- Biochemistry Department, University of Natal
- Botany Department, University of Natal
- Bradford, Conning and Partners, Consulting Engineers
- Bruinette, Kruger and Stofberg, Consulting Engineers
- Cansas International Corporation, Nelspruit
- Citrus and Subtropical Fruit Research Institute, Nelspruit
- Department of Agriculture and Water Supply Soil and Irrigation Research Institute (Land type memoirs)
- Department of Geography, University of Stellenbosch
- Department of Geological Survey
- Department of Nature and Environmental Conservation, Cape Provincial Administration
- Department of Water Affairs, Hydrological Research Institute
- Department of Water Affairs, Planning Division
- Development Bank of Southern Africa
- Division of Agricultural Engineering, Department of Agriculture and Water Supply
- Drennan, Maud and Partners, Consulting Engineers
- Eksteen van der Walt and Nissen, Consulting Engineers
- Forestry Branch, Department of the Environment
- Geography Department, University of Durban/Westville
- Grassland Science Department, University of Natal

- Hawkins, Hawkins and Osborne, Civil Engineering Consultants
- Hunt, Leuchars and Hepburn
- Institute for Commercial Forestry Research, University of Natal
- Institute of Natural Resources, University of Natal
- Kwa Zulu Department of Agriculture and Forestry
- Kwa Zulu Finance Corporation
- Lotzaba Forest (Pty) Ltd, Barberton
- Murray, Biesenbach and Badenhorst, Consulting Engineers
- Natal Parks Board
- Office of the Chief Co-ordinator, Drought Relief, Bophutatswana
- Soil Science Department, University of Natal
- South African Sugar Association Experimental Station
- South West Africa Department of Water Affairs
- South West African Parks Board
- Summer Grain Centre, Department of Agriculture and Water Supply
- Theron, Prinsloo and van Tonder, Civil Engineering Consultants
- University of the North
- Welgedacht Colliery
- Zoology Department, University of Natal

CHAPTER 11

APPENDIX D

PATCHER1

summarises the data available on all located stations in terms of which years were missing and how many complete years of data were available. Also notes latitude, longitude and elevation.

PATCHER2

synthesises data according to Zucchini's Model 5. Finding and using the best of 5 predictor stations until whole period of target stations and maximum range of predictor stations is infilled.

MERGE12

Output of PATCHER2 is 12 monthly patched files. MERGE12 picks up this output and collates it to produce a January to December and ANNUAL picture of a now fully patched station. Statistics eg. MEAN, MEDIAN etc. are also calculated.

EXTRACT-T

extracts the mean annual precipitation figures for the regression area in question.

SAME

finds all the stations with the same coordinates, the station with the shorter length of record is ignored or used to enhance the longer term station.

ADJUST

normalizes any short term stations that may have been included to improve the spatial or altitudinal distribution of the stations.

SAGOUT

produces a file containing the latitude, longitude, altitude and distance from the sea values in raster form for the area in question.

PARAMS

determines the exposure, roughness and rain direction variables that are to be used in the regression analysis .

REGSAS

performs a forced forward regression on the data using the SAS stastical package.

RESID

interpolates the residuals, that were calculated in the regression program, onto a rectangular grid.

MULT

multiplies the residual grid by the mean annual precipitation grid that was found by using the regression program.

CONTOURG

produces isohyetal maps of mean annual precipitation according to the Gauss Conformal Projection at a scale of 1:250 000.