SPATIAL PATTERNS OF SOCIAL DIFFERENTIATION AMONG THE COLOUREDS IN GREATER CAPE TOWN

ROSELYNNE LARAINE MORRIS

THESIS SUBMITTED FOR THE DEGREE OF MASTER OF ARTS
IN THE DEPARTMENT OF GEOGRAPHY
UNIVERSITY OF CAPE TOWN

OCTOBER 1980

ACKNOWLEDGEMENTS

My sincere thanks go to my supervisor, Dr G.P. Cook for her valuable insight and keen involvement with this research, as well as her constant encouragement. In addition, I owe a great deal to my husband for his patient support and his assistance with computing. Further, I wish to thank our families for their encouragement and for the financial assistance, without which this project could not have been completed. My gratitude goes also to the following people:

Ian Cook, for changes in the autocorrelation programme;
Gary Ward, for suggestions and ideas on the nature of "space";
Dr D.R. Matravers, for advice on the applications of the weighting factor in the Cliff & Ord statistic;
Dr T. Crowe for assistance with the cluster and information statistic programmes;
Mr K. Behr for the production of some of the diagrams;
Leonora Freeland for the typing;
and finally, the staff of the U.C.T. Georgraphy Department for their assistance and encouragement.

I wish to acknowledge financial assistance from the Human Sciences Research Council of South Africa, The Harry Oppenheimer Institute for African Studies and the University of Cape Town. Ideas and opinions expressed in this text, however, remain my responsibility alone.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Coloureds in South Africa are generally defined as any persons not belonging to the distinct racial groupings of white, black or Asiatic. They are the group known in other countries as mulattos, mestisos or half-castes. As a group they have always formed an appreciable percentage of the population of Cape Town and in 1975 were estimated to form the majority (57%) of the population (Bureau of Market Research, 1976). Very little, if any, geographical research on social organization has been done in Cape Town, and least of all on the coloured population in particular.[1] It therefore seems appropriate that this study should focus on the coloured population group.

Greater Cape Town consists of the four magisterial districts of Bellville, Cape Town, Wynberg and Simonstown which together comprise the 01 economic region. This region has been officially subdivided between the whites, coloureds and africans and the residential area selected for consideration is that defined by the Coloured Group Areas of Greater Cape Town. These cover 35 suburbs in which 476 742 people live in a total of 552 enumerator's sub-districts or ESD's. These are the districts into which the city is divided for the enumeration of the population for census purposes and each ESD is the area which one enumerator covers at the time of the census. As 91% of the coloureds in Greater Cape Town live in the delimited Group Areas (Ferrario, 1979) it is obvious that the study should concentrate on these areas rather than on all ESD's in which coloureds are resident. In addition, it is only in these officially designated coloured areas that coloureds have some freedom to select their place of residence and therefore if any

1 With the exception of Van der Merwe & Zietsman(1977).

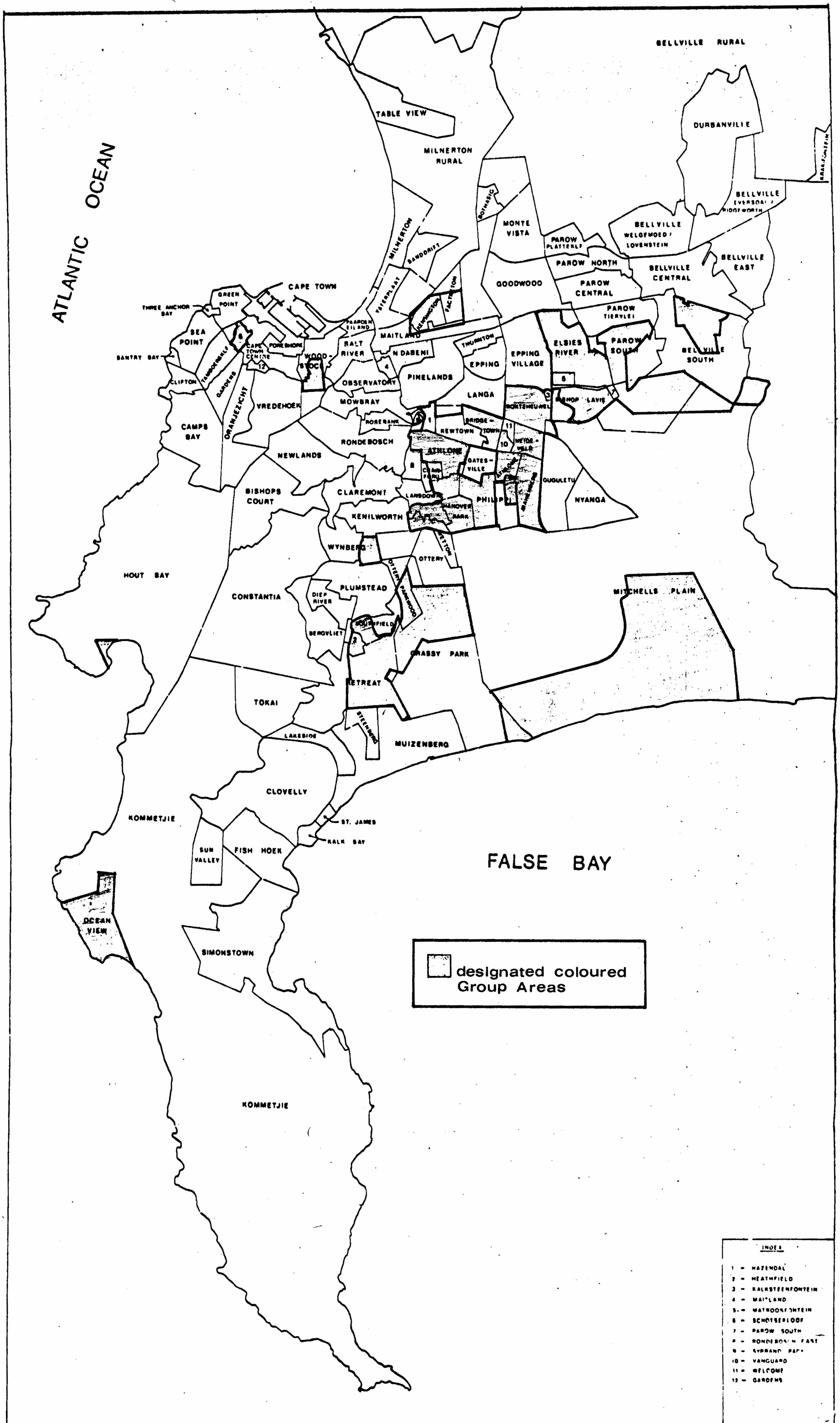
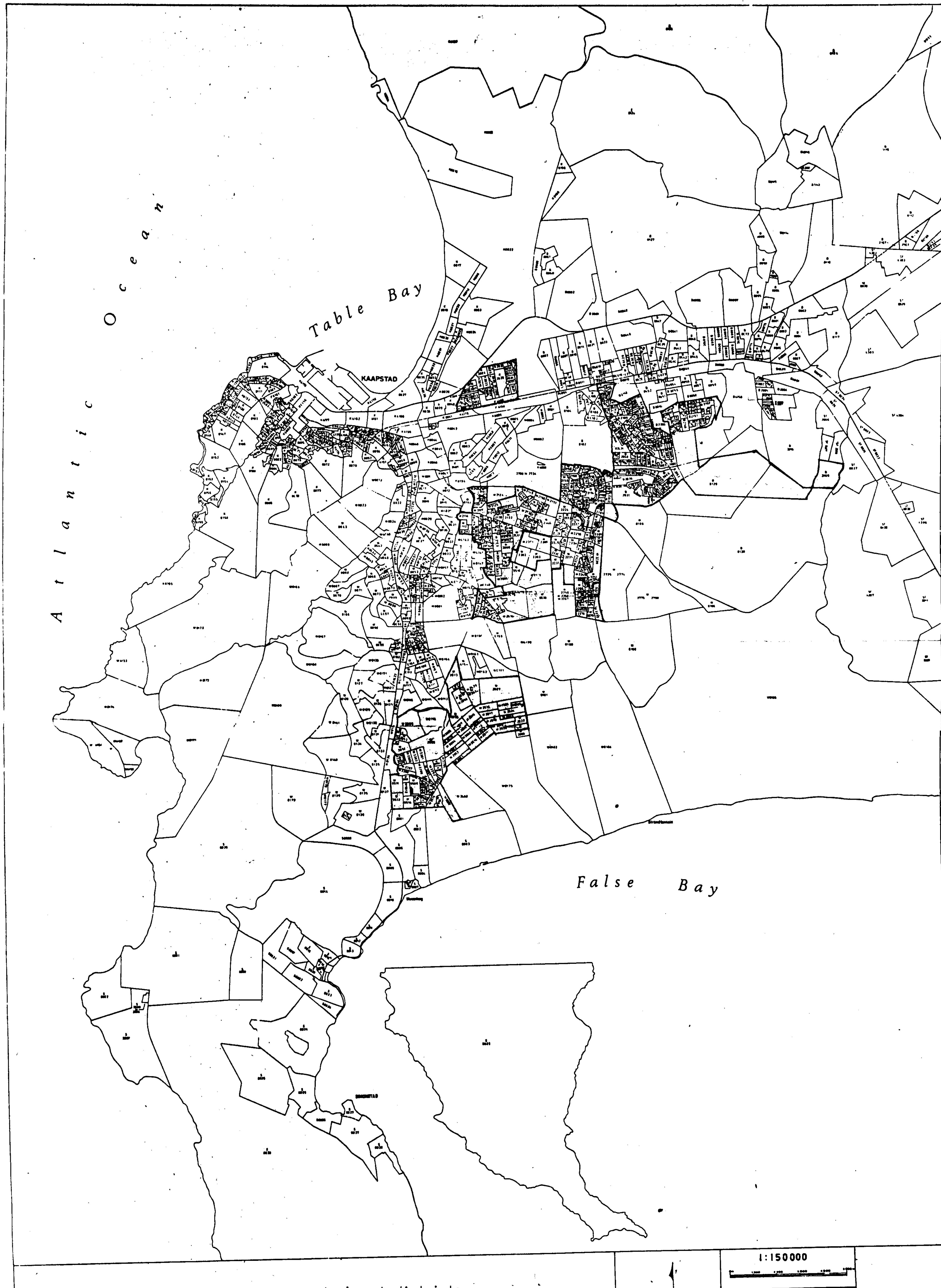


Figure 1 a) Coloured Group Areas and suburbs of Greater Cape Town



b) Enumerator's sub-districts

1:150000

N

differentiation occurs evidence will be found in these areas.

The designated coloured group area does not form one contiguous whole, but is clearly divided into four major residential concentrations (Fig. 1). These are:

a) The northern suburbs, which include Kensington, Factreton and a portion of Maitland are subdivided into 33 ESD'S and occupy an area of 4,5 square kilometers. Of the coloured population 7% live in the northern suburbs in both rented and owner occupied housing.

b) The eastern suburbs consist of five suburbs, viz. Elsie's River, Bellville South, Parow South, Matroosfontein and Bishop Lavis, subdivided into 146 ESD's. These occupy an area of approximately 17,5 square kilometers and house 26% of the coloured population. All dwellings in the eastern suburbs are rented, with the exception of a small enclave of home ownership in Matroosfontein.

c) Separated from the eastern suburbs by a major road and a day hospital so that there is less than half a kilometer of direct contact between them, lie the 15 suburbs which make up the central area. This area falls into the Wynberg rather than the Bellville magisterial district in which the eastern suburbs are located, and according to residents+ is considered to be socially distinct. The central suburbs occupy an area of approximately 36 square kilometers and are divided into 128 ESD's. Here 40% of the coloured population live in housing types ranging from subeconomic rented dwellings through the entire spectrum of housing to high income, owner occupied houses.

d) The southern suburbs forms an autonomous group of 93 ESD's comprising the suburbs of Heathfield, Southfield, Retreat, Parkwood and Grassy Park. These cover an area of approximately

---

\+ Personal communication

14 square kilometers and house 18% of the coloured population in both rented and owner occupied housing.

There are also four minor concentrations of ESD's which are designated as part of the coloured Group Areas viz. Walmer Estate, Schotsekloof, Wynberg and Bellville South. In no case do these comprise more than 17 ESD's, and together they contain only 8% of the population. Finally, some coloureds live in Ocean View, Hout Bay and Mitchell's Plain. In 1970, however, this was less than 1% of the total coloured population. Since then 474 rented and 121 selling scheme dwellings have been erected at Ocean View and 1 439 letting and selling scheme houses and flats have been built in Mitchell's Plain (Divisional Council Housing Report, 1978). The developments at Mitchell's Plain are geared to stimulate the filtering process in coloured housing by providing dwellings in a number of economic ranges. A consideration of Mitchell's Plain would thus constitute a study in itself and is therefore beyond the scope of this research.

The concern of this study is to identify social areas among the coloured population of Greater Cape Town and to see if these are recognisable in physical space; i.e. if once social areas are recognised, whether the patterns of differing social status can be related to that of adjacent areas. If socio-economic status is assumed to be reflected in housing type and quality, then physical environment and interaction between adjacent areas is likely to play an important role in creating an environment that is perceived to have a particular character and status. As a result, contiguity will have great significance in the emergence of distinct socio-economic sub-areas. It is therefore essential to concentrate a study of this type on adjacent ESD's and to regard the four major concentrations of residential areas as separate entities. By dealing primarily with contiguous ESD's each forming separate residential concentrations it carries no implication that there is no interaction between the different residential concentrations. Together they form a total system and areas of similar socio-economic status may develop in all or some of the residential concentrations.

In order to identify social areas the concept of social status must be considered. In chapter I therefore Social Area Analysis is closely examined and problems associated with the use of this technique are outlined. As it is on the basis of a social indicator that populations are divided into groups in social and / or physical space, the social indicators movement as a whole is discussed. Problems encountered in defining social indicators and a body of theory within which to view these indicators are discussed and some general applications of social indicators are reviewed. The choice of an indicator in this particular study is then justified and the applicability of the social area concept in the South African situation is evaluated.

Since the implementation of the Group Areas Act the coloured population has been regarded as a totally separate society here (Marais, 1957; Whisson, 1972; Davenport, 1977) and in this study they are considered as a separate group. Nevertheless the coloureds, according to Theron Commission Report (1976), share the aspirations and norms of white society and previously formed part of an at least partially single society.

In order to assess whether this assumption is valid and therefore whether as well it is appropriate to try to define social areas for coloureds, a description and analysis of the historical development of the coloured population is given in chapter II. Such an investigation provides a perspective on social organization among the coloured population as well as some insight into processes operative in the development of social differentiation, both past and present. Examples of social and residential differentiation are sought in what emerges as four distinctly differing periods in the history of the coloured population prior to the implementation of the Group Areas Act.

Chapter III deals essentially with the mapping of socio-economic status information and the patterns of socio-economic differentiation which result. Initially attention is given to the general problem of pattern recognition. Numerical classification is discussed and the clustering technique used to identify groups of similar socio-economic status in the study area is described and evaluated. Computation problems presented by the size of the data base are identified. The alteration of the existing computer program for cluster analysis in order to cope with such a large data set is then discussed. Initially plots obtained from mapping of descriptive statistics are described and general trends in the patterns are identified and explanations for their existence put forward. Patterns resulting from the mapping of cluster or calculated values are then described. These patterns are identified and discussed firstly on the basis of data for each ESD in the four residential concentrations considered separately. Finally patterns emerging when the residential areas are regarded as a single system are evaluated.

In chapter IV the spatial autocorrelation technique of Cliff & Ord (1969) is employed to determine spatial dependence among groups of ESD's of the same cluster. This technique establishes whether given a particular socio-economic status in one group of ESD's this necessarily determines the status of adjacent groups of ESD's; i.e. whether they will be of the same or a different status. The weighting factor is examined closely and different weighting systems are tested using both hypothetical and real data. Finally conclusions are drawn concerning the nature and distribution of socio-economic differentiation and its lack of representation in residential differentiation in coloured residential areas. The applicability of the social area concept and the link between social and physical space in the coloured residential concentrations of Cape Town is then assessed.

This study hopes to make a number of contributions in terms of identifying patterns of socio-economic differentiation. By establishing the extent to which there is spatial contiguity in terms of SES scores for coloured Group Areas, relevance and applicability of Social Area Analysis to the coloured residential areas can be measured. Correlation between the social status in a particular area and the adjacent areas (i.e. spatial dependence) may also be assessed. As a result, the effect on patterns of social differentiation of the Group Areas legislation and the allocative housing procedures as practised will be illustrated. Rather than precluding further study of social differentiation among coloureds it is hoped that this project will point to the need for more concerted study of:

a) social differentiation and social processes among coloureds;

b) the link between social and spatial differentiation;

c) the formulation of social indicators which will more accurately reflect socio-economic status among the coloured population;

d) the part which their past has played in contributing to the invidious position of the coloured in the present day social milieu of Greater Cape Town.

The aims and objective of this research are therefore two-fold. Firstly, to establish by means of mapping both descriptive data and calculated values (i.e. clusters) whether there are spatial patterns in coloured residential areas which reflect differences in social status among the coloureds. Secondly, the research aims to verify statistically the relationships between the visual spatial patterns resulting from mapping, by applying the spatial autocorrelation technique of Cliff & Ord (1969) to the data; i.e. to establish the degree of spatial dependence between the patterns of differentiation. A further aim is to implement and write into the computer program developed by Ward (1978) the weighting mechanism of Cliff & Ord (1969) which will standardise size differences for the irregularly shaped areas for which data is available. The nature of various types of weights and their effect on the correlation coefficient will also be tested.

Overall the study therefore represents an attempt to investigate the socio-economic status of the coloureds and its manifestation in physical space under the apartheid system in South Africa.

CHAPTER 1

SOCIAL DIFFERENTIATION

Social strata arise in a society as a result of increasing societal scale which is associated with differences in function; i.e. developments in technology result in varied occupations due to the need for specialization. Rewards therefore differ in conjunction with degree of specialization. Social stratification occurs in any society, whether capitalist, socialist or Marxist. According to Tumin (1967), forms of stratification will differ, depending on the society, as different emphases in rank, evaluation and reward arise in differing societies. The end result is, however, the same - inequality in the distribution of rank, power and privilege, leading to the formation of social strata within the community, which show greater similarities between individuals within the group than between individuals of different groups.

Increases in scale are presented as the main dynamic in the development of industrial society and provide the basis for structural concepts in the construction of the social area typology. It is unclear, however, whether the concept of increasing scale is intended to refer to an independent set of phenomena, concerned with social interaction, or whether it is merely intended as a general term reflecting the change from traditional to modern forms of society (Timms, 1971).

Shevky and Bell (1955), claimed that the social area typology was a logically demonstrable reflection of those major changes over time which have produced modern urban society. It is therefore implied that social phenomena are reflections of the wider and more general historical changes to modern forms of social organization, and are thus important indicators of social change. Social area analysis is therefore useful as a means of assessing these changes over time; and their effect on the stratification of society and the areal patterns that may emerge as a result.

I Social area analysis

The term "social area analysis", strictly speaking, applies only to that mode of analysis originally outlined by Shevky, Williams and Bell, in their studies of Los Angeles and San Francisco (Berry & Horton, 1970). These early social area analysts conceived social area analysis in a limited context of social change, in which change in form over time of the structural organization of the city was associated with the question of increasing scale of urban society. Increasing scale and complexity were assumed to give rise to increased social differentiation which was reflected in three major constructs and three indices describing these constructs. These three constructs viz. social rank, urbanization (family status) and segregation (ethnicity), were regarded as objective criteria reflecting the process of urbanization and stratification in modern society (Timms, 1971). The three indices, one per construct, were made up of from one to three census variables - designed to measure the position of census tract populations on scales of economic, family and ethnic status. A social space diagram could then be constructed, with each dimension of the space formed by one of the social area constructs (Figure 1.1). On the basis of statistics derived from census data, populations close to each other

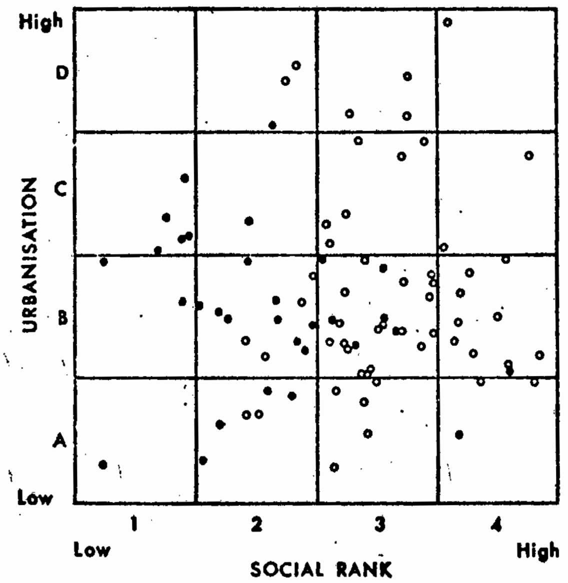


Figure 1.1   The social area classifications
Source: Herbert, 1972

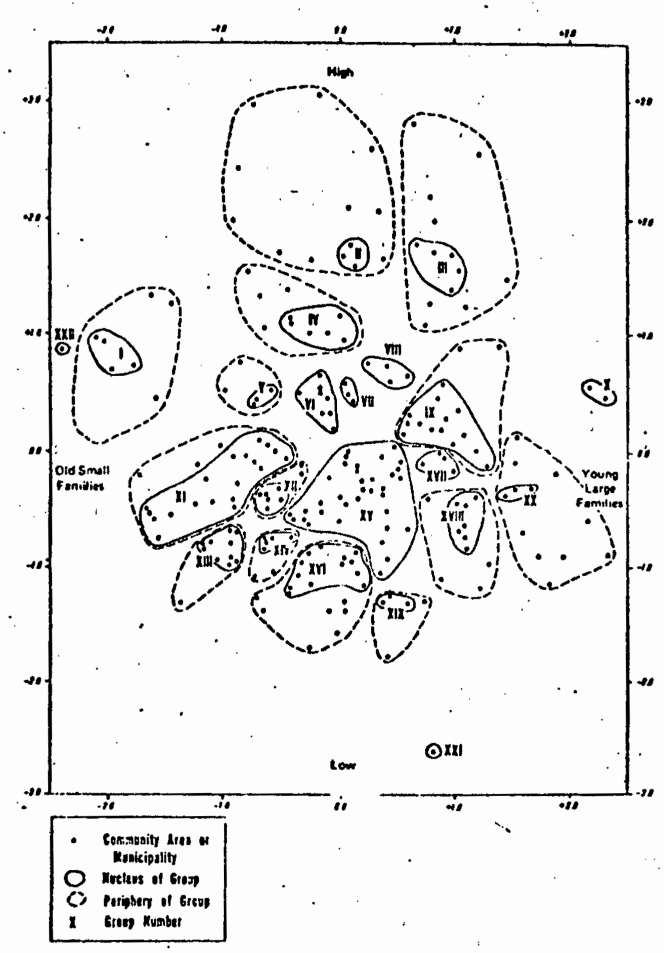


Figure 1.2   Groups in social space (Social Areas)
Source: Berry and Horton, 1970

were grouped together to form "social areas" (Figure 1.2). It should be noted, however, that there is no necessary relationship between social areas and the physical dimension of space (Timms, 1971).

According to Berry & Horton (1970), Social Area Analysis strictly defined, has been criticised both on the grounds of the theory underlying the constructs and for empirical reasons in particular the method of dimensioning the constructs. However, in an effort to meet the empirical objections that the measures used assumed the constructs to be correct and failed to provide a test of their validity, Bell in 1955 used factor analysis to show that in both Los Angeles and San Francisco the census measures used provided a structure consistent with Shevky's formulations (Berry & Horton, 1970; Rees, 1972). Van Ardsol, Camilleri & Schmid (1958) also used factor analysis in their work on ten American cities. Results from their study partly confirmed Shevky's original constructs, but also showed that in some cases there was need for modification. According to Rees (1972), these results suggested the inclusion of more variables than those used by Shevky and Bell, to fully explain the way in which census tracts vary according to socio-economic characteristics. From this rationale a large number of studies concerned with social dimensions of variation among residential area populations in cities have been carried out, using techniques of numerical classification, including factor- or some other form of multivariate analysis.

In this way the term "Social area analysis" has come to take on a broader meaning than was originally intended by the name (Figure 1.3). Social area analysis, broadly defined includes the works of its originators as well as later studies which make use of objective measures of similarity and grouping (Berry & Horton, 1970), (cluster analysis was Tryon's alternative to factor analysis and other multivariate analysis techniques) in order to produce indexes which may be used to delineate social areas (Figure 1.4). These social areas, by virtue of their base in census data have been shown to have an areal

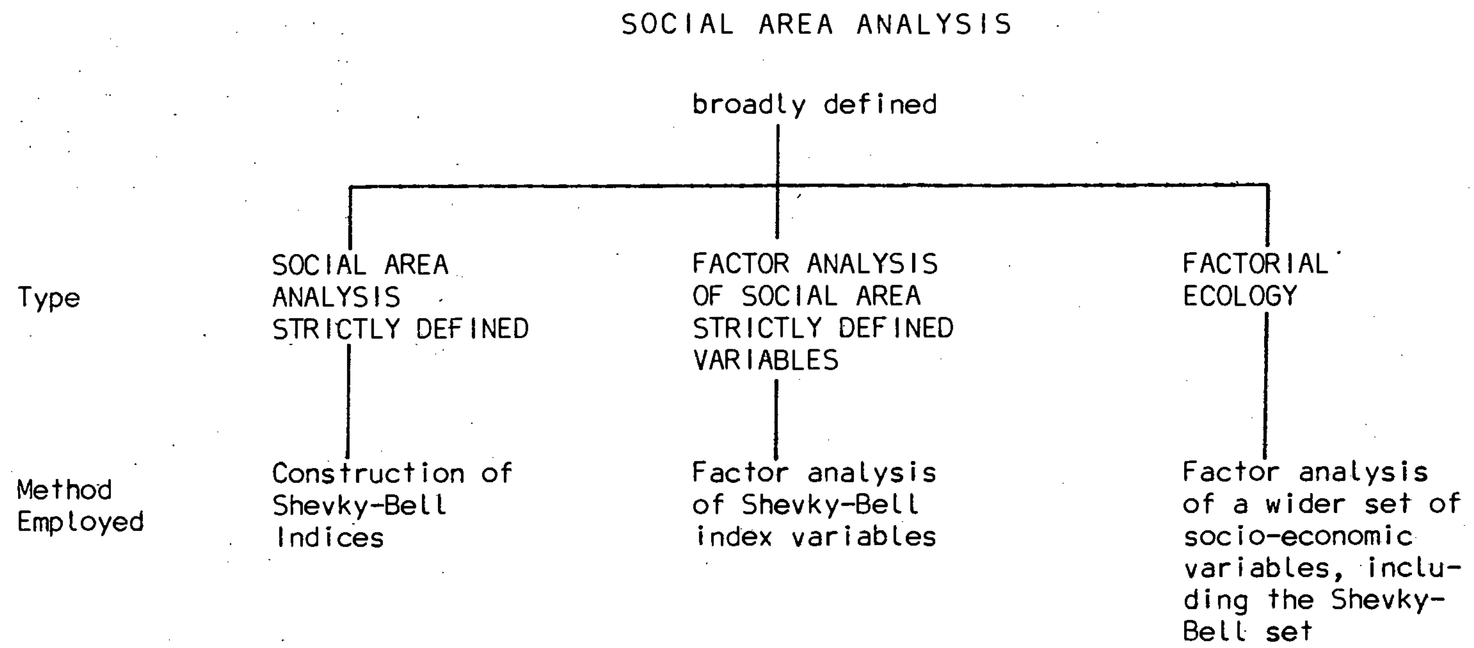


Figure 1.3 Typology of social area analysis

Source: Berry & Horton, 1970, p.315

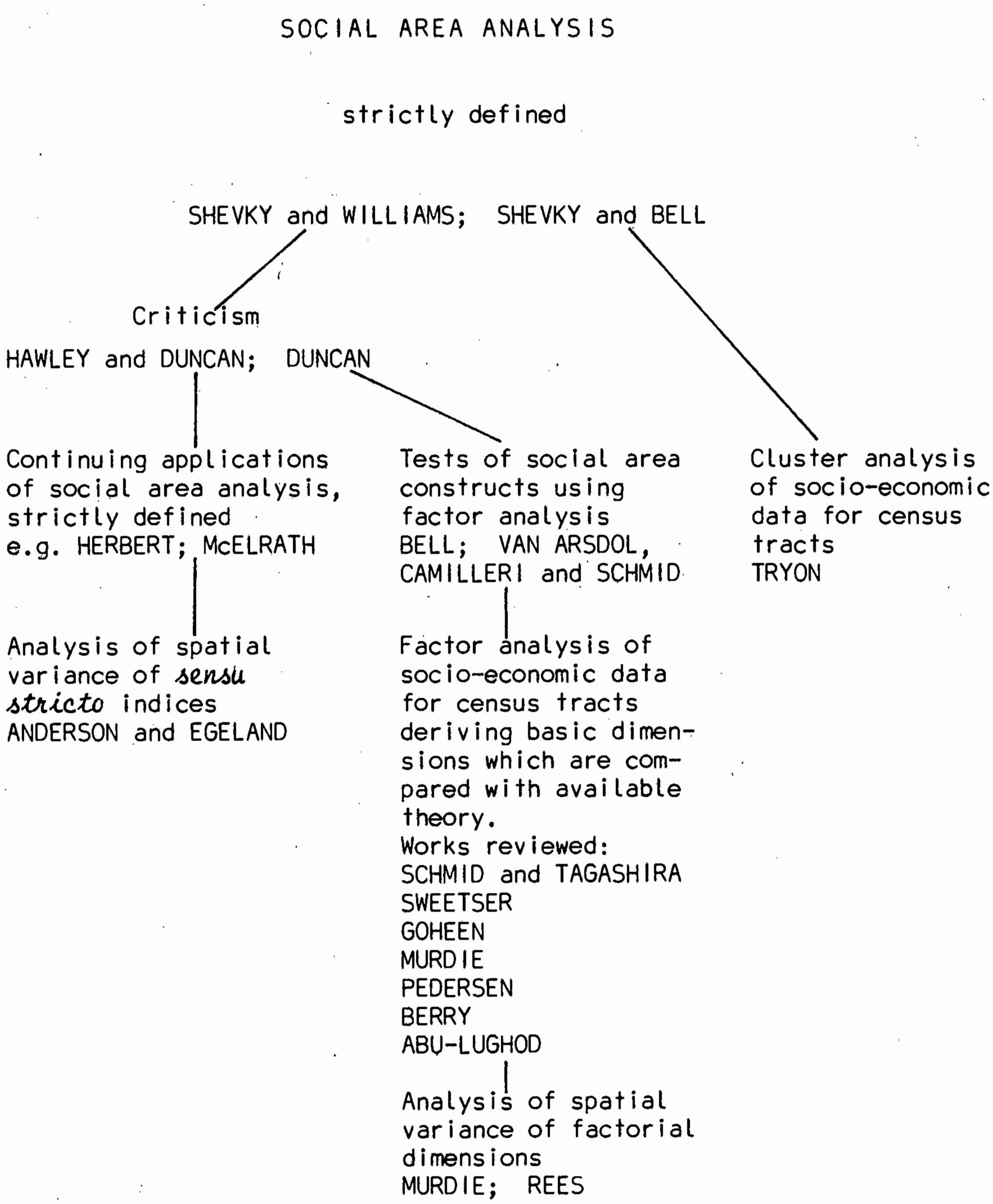


Figure 1.4 Development of social area analysis. Cluster analysis is Tryon's proposed alternative to factor analysis

Source: Berry & Horton, 1970, p.314

(latitudinal and longitudinal) component (Timms, 1971), however, although phenomena measured by the social area analysts may differ the space in which they are 'located' remains the same. Space exists independently of any phenomena or any material measuring device. Since there is no logically acceptable method for ascertaining the existence of any intrinsic spatial metric, there is no logic to support a belief in different types of space, viz. absolute, physical, geographical or social. Any particular concept of space is bound by whatever objects or phenomena the individual cares to take as points of reference. In other words; if a group of objects can be imagined as having a place component and each object is in some relation to the others, then that which comprehends all those places, can be called space. There is thus only one space, with many layers, filters or dimensions through which it may be seen. It is the relationship of these 'filters' or 'surfaces' to one another and to space, which forms the basis of the interpretation of social area analysis.

Social Area Analysis has been shown by the urban ecologists to be a useful tool for analysis embodying both the social and physical aspects of space, and is applicable in the study of the social geography of cities. Today Social Area Analysis attempts to relate the social structure of a community to the characteristics of the society in which it has developed. It is chiefly concerned with the description and measurement of social differentiation in urban communities.

The social area analysts may be contrasted with urban ecologists in that they begin with an hypothesis of social differentiation, identifying variations on a social surface which may then, if necessary be translated as physical differences. By contrast, urban ecologists have always sought initially to identify natural areas as geographical areas, and to study them in terms of their social characteristics (Herbert, 1972).

TABLE 1.1. Lughod's 'necessary' Conditions

Source: Berry and Kasarda (1977) p.132-3

Factor conditions	*Types of variables used*	*Necessary conditions based on Abu-Lughod (1968b)*
Socioeconomic status factor	Education, occupation, income	(1) That the effective ranking system in a city be related to the operational definition of social status; (2) that the ranking system in a city be manifested in residential segregation of persons of different rank at a scale capable of being identified by the areal units of observation used in the analysis
Family status factor	Family size, portions of the age pyramid, fertility	(1) That family types vary, either due to "natural" causes such as those associated with sequential stages in the family cycle, or to "social" causes such as those associated with other divisions in society, whether ethnic, socioeconomic, or other; (2) that subareas within the city are differentiated in their attractiveness to families of differing types at a scale capable of being identified by the areal units of observation used in the analysis
Disassociation between socio-economic status and family status dimensions		Either (1) that there exists little or no association between social class and family type; or (2) if there is some association between social class and family type, that (a) there is a clear distinction between stages in the family cycle, each stage being associated with a change of residence; and that (b) "subareas within the city offer, at all economic levels, highly specialized housing accomodations especially suitable to families at particular points in their natural cycle of growth and decline" at a scale capable of being identified by the areal units of observation used in the analysis, and (c) "cultural values permitting and favoring mobility to maximize housing efficiency, unencumbered by the 'unnatural' frictions of sentiment, local attachments, or restrictive regulations"
Separate minority group	Proportion of the subarea population in the minority group, measures of the relative concentration of minority groups	(1) That the characteristic(s) differentiating the minority group from the rest of the population be of perceived significance in the social system, i.e., that the urban population be truly heterogeneous; (2) that the minority group be residentially segregated from the rest of the population, at the scale of observation used in the analysis
Disassociation between minority group and socio-economic status dimensions	Measures used above for minority group, measures of education, occupation, and income (and the distribution of resources)	(1) That the minority groups be residentially segregated from the rest of the population, at the scale of observation used in the analysis; and either (2) that there exists little or no association between minority group status and socioeconomic status; or (3) that there is some association between social class and minority group status, that (a) the minority group still spans most of the social status range, though it may be concentrated at the lower end of the range; (b) a fairly full range of housing accommodation quality be available for families within the residentially segregated area
Disassociation between the minority group and stage-in-life-cycle dimensions	Measures used above for the minority group, measures of age distribution, family size, and fertility	(1) That the minority group be residentially segregated from the rest of the population at the scale of observation used in the analysis; and either (2) that there exists little or no association between minority group status and family status; or (3) that if there is some association between family status and minority group status that (a) minority groups still span most of the family status range, though it may be concentrated at one end of the range; (b) a fairly full range of housing accomodation types be available for families within the residentially segregated area

Social area analysis cannot be expected to apply to cities in which either technology or westernization are less evident than in western industrial cities (for which it was originally derived), as social strata are differentiated on the basis of different criteria for differing societies (Timms, 1971). That constructs, when applied in situations other than modern American should therefore show change or difference from the western 'norm' has been well illustrated by studies such as those of Indian cities (Berry & Rees, 1969; Berry & Spodek, 1969; McElrath, 1963) and particularly Abu Lughod's (1968) factorial ecology of Cairo. Her important contribution to the social area concept (or social area analysis broadly defined) was to recognise a set of "necessary conditions", without which patterns of social areas cannot be expected to emerge as they do in the American context. These "necessary conditions" are summarised in Table 1.1. Wherever social area analysis is to be employed, some sort of measures or criteria of similarity or difference must be used to group the population into their specific social areas, and distinguish them from other groups. These criteria are generally termed "social indicators".

II. Social Indicators

Social indicators are the signs whereby the condition of a society may be evaluated. On the basis of the access of individuals in the society to greater or smaller proportions of the facilities or commodities represented by the indicator, such an indicator may be used to measure social differences within the society. Groups of like individuals may thus be defined and social areas delimited. The choice of social indicators is complex, as the definition of a social indicator is open to much debate. The theoretical framework for definition and operation of social indicators is also much disputed and the use of indicators therefore questioned. A discussion of all these aspects of social indicators is necessary therefore, prior to an evaluation of the representativeness of social indicators and the choice of one for the identification of social areas.

a) Historical background

Interest in social indicators appears to have originated in the United States in the 1960's with the North American Space Administration (NASA)'s desire for quantitative measures of the social 'spin-off' of its activities. According to Gross (1966) there was a widespread realization that the traditional national accounting concepts were able to reflect only some of the elements of prosperity, welfare and opportunity relevant to social well-being. Therefore, after the publication of preliminary findings in 1966, social indicators established a firm footing in Federal Administrative thinking (Knox, 1975). In January 1969 the US Department of Health, Education and Welfare (HEW) published *Toward a social report.* This represented the first attempt to produce a social equivalent of the annual Economic Reports and provided a broad if relatively brief view of the nation on a wide range of social conditions (Smith, 1973). Six months later President Nixon created the National Goals Research Staff, responsible for developing social indicators capable of monitoring and predicting the American way of life (Cazes, 1972; Shonfield & Shaw, 1972). Congressional interest is reflected in a bill entitled "The Full Opportunity & National Goals and Priorities Act" introduced by Senator Walter Mondale in 1971 (Smith, 1973). The result of this urban revolution and interest in various social issues such as poverty and civil rights (Bixhorn & Mindlin, 1972) has been the development of the "social indicators movement" and a large body of accompanying literature.

The social indicators movement has also been felt in the United Kingdom (Moser, 1970; Smith, 1973; Knox, 1975) where it has been promoted by government statisticians and has been expressed in the publication by the Central Statistics Office in 1970 of *Social Trends,* an annual statistical report. In France too, the movement towards social indicators has been taken up by the office of the Commissioner-General for the Plan; while the United Nations Economic Commission for Europe decided in November 1969 to launch a programme of joint research to identify "societal" variables which can be used in social forecasting (Cazes in Shonfield & Shaw, 1972).

Cazes (1972) identified two themes underlying the early interest in social indicators viz. the scientific knowledge of society (or a description of the social state of society) and "social planning" or the deliberate interference with the structure and performance of society and the whole complex of social phenomena. In "social" planning, the state of society is ascertained in order to identify those parts which require attention and aid, so that provision may be made for these. He suggested that whether the aim is to assess the impact of a given phenomenon on any specific part of the social system, or to observe the state of the system and endeavour to detect the factors which explain the changes in that system over time, there is a need for social data regularly accounted for. The character and diversity of this data should make it possible to evaluate a high number of programmes. The regularity of its collection should also make possible comparisons over time and projections for the future. A point that must be emphasized, vide. Land (1970) Kamrany & Christakis, (1970) and Carlisle, (1972) is that social indicators should be seen as applying to a social system and as such the systems approach of linkages and interrelationships between components within the system, is an important feature to be remembered in both derivation and application of social indicators. The interest in social indicators therefore (Cazes, 1972), reflects a period of optimism in the social sciences about the possibility of quantifying social facts. Sheldon & Freeman (1970) argue that the social indicator movement can contribute only to its more modest goals of:

1) improved descriptive reporting;
2) analysis of social change; and

3) prediction of future social events and social life.

Along with Kamrany & Christakis (1970), Plessas & Fein (1972) and Knox (1975), they assert that the claim that social indicators can be used to develop a system of social accounts is not reasonable because there is no social theory capable of defining the variables of a social system and the interrelations between them and such a theory is an essential prerequisite to the development of a system of social accounts.

The aim of developing a system of social accounts would surely be for identifying inequalities in the social well-being of a society or a nation which may be used in order to direct planning for greater equality in the future. In order to assess the state of the society or various parts of it, statistics describing particular attributes of the population may be chosen which will reflect differences within this population. An evaluation of these differences would allow for an overall assessment of the social well-being of the society and plans for the removal of inequalities could then be made. For this, no theoretical framework is necessary. The claim of Plessas & Fein (1972) that social theory is an essential prerequisite to the development of a system of social accounts may therefore be disputed.

Directly connected with the problem of application of social indicators is what Plessas & Fein (1972) call the basic dispute about social indicators which is related to the relative importance of goals versus means. The early consensus (Gross, 1966; Biderman, 1966; Duncan, 1967) was that there was a need to develop measures of systematic operation in the hope that such measures would ultimately progress towards systematic goals. This reflected the necessity for immediate action (i.e. the need to formulate and derive social indicators in order to assess social states for future planning) versus the equal importance of scientific validity (Bauer, 1966; Gross & Springer, 1969; Perle, 1970; Sheldon & Freeman, 1970). Writers such as Bauer (1966) preferred to seek improvement by the use of available data while emphasizing the defects and limitations of such material. Authors such as Sheldon and Freeman (1970), however, maintain that it can hardly be claimed that even potentially a set of social measures exists that parallel the economic ones without the guidance of theoretical formulations concerning significant variables and their linkages. They feel that the role of social indicators in the overall process of national policy planning must be clarified and the connectiveness of social indicators to national or societal objectives and goals must be established first. What is needed is the development of a "national social theory" to which "social indicators would provide the necessary social information" and from which "national social policy measures" are derived (Sheldon & Freeman, 1970). No theory, tools, system of information or indicators such as those in the economic sphere, exist in the social sphere.

It is important at this point then, to investigate the conceptual framework within which social indicators have been developed and to consider the problems of both such a framework and the resultant difficulties associated with the use of social indicators.

b) Relation to theory

The conceptual framework described by the majority of writers on the subject appears to be a set of somewhat loosely defined, if common, ideas on the requirements for the recognition of criteria which social indicators must meet, rather than a structural model for their development or a logical background or framework within which they can be viewed and evaluated.

The general feeling on the subject of social indicators is that a unified body of theory is conspicuously lacking (Kamrany & Christakis, 1970; Garn & Flax, 1972; Plessas & Fein, 1972; Knox, 1975; Smith, 1979). It is felt that this is partly due to what is considered as the lack of definition of the field and objectives of social indicators. A major difficulty in creating and using social indicators (Plessas & Fein, 1972) is separating them effectively from economic indicators and the national economic accounting system. The other social sciences, they maintain, lack a unified body of theory such as that governing economics. They also lack a single basis for comparability, such as money in economics. Thus the problem of deciding on what social measurements should be made, how they should be made and how they should be evaluated, are obstacles which will have to be overcome before a theoretical framework can be achieved.

According to Kamrany & Christakis (1970) social indicators have to meet some necessary and sufficient criteria before they can be said to have a theoretical basis. These criteria are: <u>Completeness</u>: which refers to the embodiment of all variables and elements that impinge upon the quality of a good life. <u>Specified Levels of aggregation</u>: Aggregation of social accounts masks specific problems or specific problem areas. Basic statistical difficulties exist in establishing indicators and indexes, and a great deal of research is needed to develop appropriate levels of aggregation and disaggregation for various socio-economic information. <u>Geographical delimitation</u>: Indicators have varying degrees of sensitivity and significance with respect to the level of geographic disaggregation, which varies from the earth, down to a lot in a specific location. Further problems relate to the boundaries for which information or data is available but they should be easily disaggregated by geographic areas because such sub-areas and their growth patterns have important implications for public decision making. In addition Knox (1975) feels that where possible, they should refer to the outputs of the system - such as achievement in any sphere, rather than inputs such as expenditure in that sphere and relate to public policy goals, as well as be available as a time-series.

In the search for a body of theory within which to develop or assess social indicators therefore, the uses to which indicators will be put must not be obscured. Social indicators should be seen as tools which may be employed for describing the social state or conditions of a society. The exact purpose in employing them depends on the objectives of the investigator concerned. Indicators will be defined and used in different ways depending on the requirements of the researcher. From this point of view, the contributions of Carlisle (1972) and Springer (1970) are extremely useful as both emphasize the importance of the use to which the indicators will be put. Kamrany & Christakis (1970) and Carlisle (1972) also suggest some criteria (depending on their uses) for defining and developing social indicators.

If there are specific criteria which should be met when formulating social indicators and if the choice of indicator and the manner in which it is applied depends upon the requirements of the particular research being undertaken, there appears to be little real need for a body of theory within which indicators must be developed and evaluated. As long as the requirements for the formulation of indicators are fulfilled and indicators are adequately defined along with specifications for their use, and the manner in which they are to be used is described, they may well be considered as valuable tools for assessing the social well-being of a society (see Knox, 1976).

c) Definition

According to Plessas & Fein (1972), many writers on social indicators fail to provide their readers with an adequate or even partial definition of their subject. This sentiment is echoed by such writers as Carlisle, 1972; Cazes, 1972; Smith, 1973; Knox, 1975, and many others. The lack of definition of the subject and scope of social indicators these writers claim, may be one of the causes of the difficulties in establishing a conceptual and theoretical framework or guidelines along which social indicators may be developed and the social indicators movement may proceed. Lack of definition and direction results in fragmentation of interest in the subject and confusion of objectives for the development of theory and application of social indicators.

Social indicators have been variously defined (Smith, 1979; Knox, 1975; Kamrany & Christakis, 1972; Bixhorn & Mindlin, 1972; Garn & Flax, 1972; Cazes, 1972; Carlisle, 1972; Perle, 1970), and as Sheldon & Freeman (1970) have pointed out, there is little agreement on the defining attributes of social indicators beyond the notions that:

1) Social indicators are a time series that allow comparison over an extended period; and
2) Social indicators are statistics that can be disaggregated (or cross-classified) by other relevant characteristics.

Land (1971) claims that these two criteria do not allow for a distinction between social statistics and social indicators and suggests that it would be preferable to have a definition that would facilitate a resolution of the proper functions of social indicators with respect to social accounting and social reporting. Social indicators tell sociologists something about the functioning of society. He proposes therefore that the definition should include the fact that social indicators are components in a social system model or some particular segment or process thereof. The important point, Land (1971) maintains, is that the criterion for classifying a social statistic as a social indicator is its informative value which derives from its empirically verified nexus in a conceptualization of a social process. Carlisle (1972) provides just such a definition of a social indicator - it is "the operational definition or part of the operational definition of any one of the concepts central to the generation of an information system descriptive of the social system". For Carlisle (1972) the use to which the information is to be put determines its definition as a social indicator. Her emphasis is placed on the operational or functional aspects.

The definition of a social indicator therefore appears to be dependent on the social system under consideration, the information available for describing that social system, and the use to which the information is to be put; i.e. what is to be investigated or reported about which part of the social system. The social system concerned must therefore be modelled; relevant and available social statistics reviewed and the use to which these are to be put, evaluated. It is the usefulness of social statistics for describing social systems which determines whether they become social indicators.

d) Uses of social indicators

Notwithstanding the many criticisms levelled against social indicators and the problems associated with the formulation and application of these indicators, they have been widely used (see Kahl & Davis, 1955; Agocs, 1970 and *Social Indicators research* 1974 - 1979). Spence (1968) used a number of factors in order to regionalize British counties on the basis of employment. Gordon & Whittaker (1972) describe prosperity in the south western region of England. An assessment of the effect of north-sea oil on two small towns on the Scottish coast is made by Knox (1976) on the basis of evidence derived from subjective indicators. These three applications are all examples of the use of social indicators for the specific purpose of describing and assessing the state of a society. Social indicators, particularly socio-economic status indexes and the social areas delimited by these are important in urban modelling. In cases where direct information for an area is not easily accessible or available, areas delimited by the use of social indicators, as being of a certain status may, on the basis of this status be assumed to have certain characteristics. These assumed characteristics may then be used for predictive and planning purposes (Wilson & Rees, 1979). The use of social areas as a basis for modelling behaviour; i.e. shopping patters (Nader, 1969) and travel patterns (Forrest, 1974) are examples of this. Attitudes and reactions of particular ethnic or socio-economic groups can be estimated and predicted and planning undertaken on the basis of social areas (possessing certain characteristics) delimited by the use of social indicators; e.g. attitudes to migration have been shown to vary with social status group (Hart, 1973). The consideration of welfare among blacks in metropolitan and rural regions has been a factor in the planning for optimum city size in the USA (Mills, 1972).

While Perle (1970) expresses certain reservations, supported by Sheldon & Freeman (1970) he suggests that the usefulness of indicators may be divided into five major themes:

1) Improved descriptive reporting on the state of society;
2) The analysis of social trends and social change;
3) Assessing the performance of society;
4) Anticipating alternative social futures;
5) Social knowledge for societal control.

The wide range of uses have led Kamrany & Christakis (1970) to suggest that where aggregation ought to be made in terms of regional indicators the overall behavioural and environmental conditions should be translated into three kinds of indicators, namely:

1) <u>Absolute indicators</u> - which refer to those "scientific" indexes for which a substantial agreement among experts has been reached; e.g. minimum requirements for clean air.

2) <u>Relative indicators</u> - those for which time-series data and cross comparative data are available and for which no optimum value is available. (The boundaries of an optimum, reasonable or expected rate of change for relative indicators could be established; e.g. number of parks per population.) These are similar to the predictive indicators described by Carlisle (1972).

3) Autonomous indicators - indexes which reflect social, economic, institutional and cultural values of specific regions, e.g. socio-cultural attributes of a particular ethnic group. (Carlisle, 1972, called these informative indicators.) This idea was expounded by Carlisle (1972) by the addition of two extra types of indicators associated with policy; viz. Problem oriented and programme evaluation indicators.

As both relative and autonomous indicators are subject to various, continuous socio-economic changes which the nation experiences, a system of continuous monitoring and interpretation of these changes on both types of indicator are necessary.

Regardless of whether they are finally to be used for social reports, predictions or planning, autonomous social indicators are initially used to assess the condition of a population with regard to particular criteria. Unfortunately variables which might be considered to be measures of a social state are not always readily quantifiable (Etzioné & Lehman, 1967). Furthermore social indicators are generally based on more than one social measurement in order to make meaningful statements on the social conditions within any group or area. Individuals in the population then, may be grouped together or separated from one another as a result of differences in their ability to attain a proportion of the facility or commodity represented by the indicator, and areas of different status thus emerge. Social areas are therefore identified as the direct result of the division of the population into groups on the basis of social indicators. Possibly the most commonly encountered social areas are those which consist of individuals which have been grouped together on the basis of their socio-economic status. This sort of grouping is often (at least among advanced western societies) reflected in the grouping of these people also in specific geographical areas and identifiable by such characteristics as type of housing.

III Social Indicators and Residential Differentiation

Social indicators have been shown to be useful tools in social analysis, especially as a means for delimiting and differentiating between social areas. Although social area analysis filters out social differences from the urban mosaic the resulting patterns may be transposed onto the physical surface. This is possible as the statistics used to delineate social areas on the social surface are drawn from census tracts. Once the social area has been located on a social space diagram, the census tract may then be located on the physical surface as an area which has a particular set of social characteristics.

Implicit in the social area concept then, is the idea that it embraces two dimensions of space - the physical and social surfaces. Depending on the city under consideration, one or other of these surfaces may be seen as the more relevant filter through which the particular space may be viewed. It is possibly a lack of understanding of the dual components of the social area concept which has led to the formulation of the major criticisms against it, namely: that choice of indicators is regarded as arbitrary and that they lack independence from one another as well as the criticism that the typology is not universally applicable.

The identification of the physical and social components and their relative importance and relationship to one another in any particular application of the social area technique must be ascertained before any meaningful assessment of the situation under consideration can be made. Although the physical and social surfaces are both dimensions of space, they are not interchangeable. Rather, they are filters through which space and each other may be viewed. The degree to which housing or the physical surface may be used as a filter through which differentiation on the social surface may be seen, will vary between societies and reflects the emphasis placed on residential differentiation as an indicator of social differentiation (Harvey, 1973a).

In urban-industrial society according to Shevky & Williams (1949) the unfailing indicators of the social position of others readily accessible to everyone are houses and areas of residence. People of like socio-economic status tend to locate in the same or similar residential areas, close to areas of higher socio-economic status, and as far as possible from areas of lower socio-economic status. In this way the social distance is minimised between individual and the population he wishes to emulate, and maximised from the groups he wants to leave behind (Timms, 1971).

The degree to which housing type and quality can be used as an indicator of social stratification depends on the degree of differentiation in housing. Residential differentiation is thus not a stable indicator. For example, in advanced industrial capitalist dominated cities there is a great variety of housing types. Here social differentiation (the social surface) can be seen through the housing filter (the physical surface). In other cities where there is less variety in housing this cannot be done.

Harvey (1975) puts forward several hypotheses relating residential differentiation to social structure. Firstly he claims that residential differentiation should be interpreted in terms of the reproduction of the social relations within capitalist society. Secondly, residential areas ("neighbourhoods" or "communities") provide distinctive milieus for social interaction from which individuals to a considerable degree derive their values, expectations, consumption habits, market capacities and states of consciousness. Residential neighbourhoods are, in other words areas in which the socialization processes of particular groups take place. Harvey (1975) further claims that residential differentiation is produced by the organization of forces external to the individual or even the collective will of a particular social group. Financial and government institutions control and co-ordinate housing-market behaviour and in so wielding their influence produce certain broad patterns in residential differentiation.

In terms of western industrial cities, American cities in particular, where residential differentiation is an accepted criterion for the evaluation and reflection of social status (Harvey, 1975), residential differentiation may be used as an indicator for identifying strata in the society and assessing their relative distances from one another on the social surface (Harvey, 1973a). In other societies, however, emphasis may be placed on other phenomena as indicators of social status and social differentiation. Abu Lughod (1969) for example, in establishing the factorial ecology of Cairo, found that social differentiation although real could not be seen through the filter of a physical or housing surface.

In South African society where the whole population is legally separated on the basis of race, and not all groups have the same access to the modes of production, interesting variations may be presented for the application of social area analysis, and in the link between social space and residential differentiation.

IV Applicability of the Social Area Concept in the South African Context

Social areas can only be recognised once some criteria have been chosen as a basis from which to assess similarity or difference between individuals in the population under consideration. Prior to any attempt to apply the social area concept in the case of South African society then, relevant indicators must be identified and utilized to asses the state of this society.

Very little, if any influence of the social indicators movement has been felt in South Africa. No official attempt has been made to set up a national social accounting system and no plans for identifying social problems and for social forecasting are operative on a country-wide scale. Statistics which could supply social information such as figures on housing, income, population growth, home language, religion, medical facilities, agricultural yields, health, education and occupation are collected on a fairly regular basis, both by census every decade and at substantially shorter intervals by the Bureau of Statistics. Other sources of such statistical information, though not as general nor as regularly collected for comparable areas, include the publications of financial and mining houses and the various agricultural boards. On the more local level, reports of the Medical Officer of Health, as well as other municipal organizations and councils provide a wealth of social statistics. These could be used for assessing the state of society, especially in terms of differences in distribution of these attributes and services between various areas and population groups. Distinct and differing social areas could therefore be identified, particularly as population groups are so well defined and separated here. Such statistics might therefore be used as indicators of the conditions of South African society.

In seeking to apply social area analysis (broadly defined) as a tool for assessing and comparing social and physical structure of cities, it should be noted that it is the concept rather than the rigid formulation which should be utilized. Further, the important distinction between the social and physical components of space and their relationship to one another in differing societies must be considered when attempting an analysis or comparison of social and physical structure of cities.

South African society has been described by Fair (1969) as a western industrial society with a modern money economy superimposed on an indigenous tribal society with an economy based on subsistence. This dichotomous system has resulted in an overall social and city structure which is therefore not directly comparable with western capitalist orientated industrial cities. The turbulent social history and accompanying rapid changes in social structure has led to the development of a complex and unique social form in South Africa. As is the case in the United States, the population is extremely heterogeneous, a combination of many cultures. The uniqueness of modern South African society however, lies in the social development of the country as a result of political forces which have played a major role in this development particularly in the last three decades (Davenport, 1977). As a result of the ideology of the ruling political party, legal and economic constraints have been imposed on South African society, resulting in broad stratification between white and non-white. Further stratification within these groups may have occurred in response to other factors. Political ideals have also found expression in residential separation of ethnic groups - white, black, coloured and asian by specifically delineated areas for the use of each population group in every city.

Davies (1963) has suggested a tentative model for South African cities which resembles the Hoyt sectoral model with the addition of semi-independent non-white townships. South African society, Davies (1976) claims in broad terms, strikingly resembles the form of colonial society. Secondary stratification is however, based on purely socio-economic criteria. Concomitant with opportunity, range of socio-economic stratification for whites is wide but is limited for the non-white groups. The South African city is primarily arranged in relation to class race divisions and secondarily in relation to socio-economic status.

Among the white population group, residential differentiation appears to be a reflection of social status and it has been used as an indicator for identifying strata in the society (Hart, 1968, 1975, 1976). Social differentiation among the white population group may thus be seen through both the physical and social filters.

In population groups other than White, however, emphasis may be placed on phenomena other than residential differentiation as indicators of social stratification and social status. Legally imposed economic constraints, the allocative housing system, housing uniformity and shortages (for the majority) and limited space for expansion of certain ethnic groups within their group areas, inhibits the extent to which residential differentiation can occur.

Among the Asian (Indian) population, social stratification occurs primarily on the basis of language and religion (Hey, 1961; Watts, 1971) but is increasingly becoming a reflection of economic and occupational as well as educational status (Brand, 1966; Watts, 1971). Meer (1968) finds evidence of residential differentiation ranging from wood and iron shacks, council built sub-economic and economic rented accommodation and dilapidated housing to affluent owner-built houses which accommodate extended families. It would appear that Group Areas legislation has restricted the development of Indian housing (Brand, 1966; Meer, 1969), but where less restricted development is possible social stratification may be reflected in residential differentiation.

Brandel-Syrier (1971) and Wilson & Mafeje (1963) find that among the Black population group stratification occurs as a result of economic, occupational and educational differences. Although this differentiation is generally not visible in the form of residential differentiation (Kuper, 1965) as a result of the lack of available housing and Group Areas restrictions, where circumstances allow its development, residential differentiation reflects social status (Brandel-Syrier, 1971).

The Theron Commission Report (1976), Van der Merwe (1962), Thomas (1976) and Stone (1972) all find evidence of social stratification in terms of education, occupation and income among the coloureds. Notwithstanding the restrictions of Group Areas legislation and a lack of available space, the full spectrum of housing, from sub-economic rented dwellings, to high income owner occupied ones exists in each of the concentrations of coloured residential areas considered in this study (S.A.I.R.R. 1950-1978; Phillips, 1971; Cape City Council, 197 ; Ferrario, 1979). It is therefore suggested that socio-economic stratification may thus find expression in terms of occupation of differential housing by different social groups within the coloured residential areas.

While for the majority for the non-white groups therefore social stratification is not reflected in residential differentiation, there is evidence that among certain sectors of the non-white population, social differentiation may be seen through the physical filter as residential differentiation.

It thus appears that the social area concept when applied in the non-white South African context, should first be separated into its social and physical components. It is necessary to establish whether in the light of the constraints posed by Group Areas and allocative housing procedures, any spatial differentiation has occurred. Thereafter relationships between social differentiation and physical (not necessarily residential) differentiation can be evaluated to gain an idea of the true nature of any social areas that may exist.

The specific aim of this study is to identify groups and residential areas of similar social status within the coloured population. Statistics describing the social state of this group are available from census data. Those which provide the means for classing people as socially similar or dissimilar may therefore be defined as social indicators.

Using socio-economic status indicators, an attempt is made to ascertain whether areas of differing socio-economic status reflecting the characteristics put forward by Timms (1971) occur in coloured residential areas in Greater Cape Town, and if these are geographically recognisable in terms of differences in housing type.

As well as socio-economic status (Timms, 1971) ethnic identity affects desirability of an area for residential location. The degree to which both of these affect the locational choice depends on the extent to which the ethnic population is stigmatized and whether socio-economic status differentiations in the group are similar to those of the core (or dominant) society. Where the ethnic group is small and the majority confined to a single social status, ethnicity and socio-economic status become synonymous. To a large extent this is the case with the coloured population in Greater Cape Town. Although coloureds form the majority of the population here, they are a minority group in South Africa, and are anyway regarded as a separate group by law. Legislation to a large extent also determines and confines the range of occupations open to coloureds and the incomes for these occupations. Seen within the context of the wider society of Cape Town, these people form a distinct ethnic group generally of a lower socio-economic status. This study therefore, represents an attempt to see whether the coloured population of Cape Town does form a uniform social status group, and if not, to establish if patterns of differentiation exist and to investigate aspects of socio-economic differentiation within the coloured group.

In a society where prejudice against an ethnic group is great (Timms, 1971) (as in South Africa (my parentheses)), residential areas adjacent to those of the ethnic group become highly undesirable to the core society, regardless of the socio-economic status of the individuals of the ethnic group located there. These areas become highly desirable for the ethnic group though and individuals of higher status within the ethnic group generally locate there. Social distances between the ranks of the ethnic group remain marked however, and patterns of social desirability tend to mirror those of the core society. The coloured population of Greater Cape Town is confined to specific geographical areas because of the Group Areas Legislation. Analysis of the coloured residential areas using socio-economic status indicators should show not only whether social differentiation occurs within the group, but whether it parallels social differentiation among whites here, and further, whether the coloured residential areas adjacent to white residential areas show higher status than others. Social indicators are thus to play an extremely important part in the identification and analysis of patterns of social differentiation among the coloured population of Greater Cape Town.

In order to ascertain the existence of social areas, some specific measure or indicator of social similarity or differences within the coloured community must be adopted.

V Social Areas, Social Indicators, and Coloured Residence in Greater Cape Town

A number of variables are available which characterise the coloured residential areas and could be used as indicators for defining social areas among this population group. These indicators include the following:

a) Residential indicators

i) Types of dwellings: In Cape Town types of dwellings in residential areas are classified as House, Flat or Other. Numbers of each are given for each ESD. This is however too broad a classification upon which to assess residential differentiation particularly as it is not specified how many houses are rented as opposed to owner occupied; how many are sub-economic as opposed to economic rented dwellings; how many dwellings are council owned or privately owned and neither is the size, quality or type of house defined. Dwelling type by ESD is thus of little use in determining patterns of spatial distribution of housing types from which areas of a particular socio-economic status might be inferred.

ii) Population Density: Total population by ESD is meaningless unless related to number of dwellings or rooms per ESD to give a measure of residential density. Density defined as number of housing units per area; of persons per room or of families per dwelling are all commonly used indicators of social status - either in the western sense of preferred low density (of persons per room, families per dwelling and dwellings per ESD), or in the non-western situation where the extended family concept leads to higher average densities.

In the case of the coloured population of Greater Cape Town though, the usefulness of the measure is limited by the acute housing shortage (Personal communication, Cape City Council Housing Branch, Divisional Council Housing Branch) which leads to multiple family occupation of single family dwellings by necessity rather than by choice. Nevertheless Louw (1979) found some limited evidence of distance decay patterns which reflect the theoretical models as outlined for western and non-western cities (Berry & Horton, 1970). In the case of the coloured residential areas of Greater Cape Town, Louw (1979) found that no consistent relationship between high income and distance from city centre existed. Table 1.2 may thus be evidence of a non-western situation or of pressure on available housing.

b) Family status indicators

These include education, occupation and income. One or all of these indicators may show some degree of social differentiation among the coloured population, as it would be expected that, as in the case of

TABLE 1.2 Density and Income Characteristics of Selected ESD's in Coloured Areas

ESD	Av. family income (R)	Density m^2 per person	House density m^2 per house	Average SES score
C2003	1617	91,4	-	58
C0182	1336	80,5	-	51
C0045	1138	155,3	-	46
W2135	2039	97,8	680	52
W2145	1336	97,9	643	54
W2211	1227	46,3	-	51
W2209	1416	104,5	634	52
W2204	1295	90,8	520	51
W2013	1035	279,3	-	48
W2294	1555	151,3	1238	43
W2296	996	75,2	455	38

Source: Louw, 1979

the white group, that better education secures better jobs and therefore higher wages. However, there are problems associated with the use of these indicators to identify social differentiation among the coloured population group in Greater Cape Town. There is a ceiling on incomes for coloureds, due to the wage policy of the government; the availability of education is limited, and the range of occupations open to coloureds is limited by job reservation. These indicators must therefore be adapted in some way in order to make them more meaningful as indicators of social status and differentiation in the context of the coloured population group.

For the purpose of this study, which seeks to identify the spatial patterns of those characteristics which may be interpreted as indicators of social differentiation or stratification, the above characteristics should be seen in combination with some, and related to all of the variables discussed above.

c) Socio-economic status scores

A variable which relates income, education and occupation and combines them into a composite indicator is the socio-economic status (SES) score. The combination of such variables is commonly regarded as an acceptable indicator of social differentiation (Kahl & Davis, 1955; Knox, 1975; Smith, 1979 Differences in the scores imply differences in social status of the residents in any area. These scores may therefore be used as informative indicators to report on the spatial component of social status differentiation as suggested by Sheldon & Freeman (1970) in particular. Scores of this type were developed by the Technical Management Services section of the City Council for each enumerator's sub-district of Greater Cape Town. According to the TMS Working Paper 6-76 (1976), the compilation of a socio-economic rating index represents their attempt to use census information to construct a valid measure of socio-economic status for all residents in Greater Cape Town. Separate scores were derived for each population group (white, coloured and asian). This means that it should be possible to identify within group differences in socio-economic status. Combined scores for all races were also given, but this implies that classification into socio-economic groups would be much broader probably grouping non-whites together.

The SES scores were derived as follows:

Using data for all coloureds in the 01 Region, cumulative percentage groups were constructed for each of the of the parameters education, occupation and income. The categories in each component were adjusted so that about ten percent of the persons in the universe fall in each tenth of the distribution of the item. In order that each individual should have three separate ratings to form the average score calculations were based on each income earner. The overall SES scores calculated in such a manner show a normal distribution for the entire population of Greater Cape Town.

For each ESD (Enumerator's sub-district) the ratings of individual income earners for education, Income and occupation were then calculated from the standardised percentiles. A numerical example illustrates the method clearly

Education:	A	96	B	31	
Occupation:		64		43	
Income:		90		74	
		250		148	
SES scores		83		50	etc.

These values were then used to calculate the average score for all income earners in the ESD. An average SES score for the ESD as a whole and frequency of occurrence of individual scores by percentile provide a full profile of characteristics of that ESD, viz.

	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-100	Av. Score
ESD W2116	2	18	28	38	76	105	88	101	32	10	57

A single meaningful indicator is therefore available by ESD for the coloured population. This can be used to identify whether spatial patterns of social stratification exist within the coloured residential areas.

Although there are problems associated with the use of socio-economic status scores as status indicators, there are a number of factors which justify the use of these scores in this particular case. Firstly, in Greater Cape Town in particular other available indicators such as housing type; housing and population density; persons per household; rent paid or income are not adequate indicators of socio-economic status among the coloured population because of the situation of the majority of coloureds in the wake of the group areas implementation and due to state imposed wage and job restrictions. This has resulted in the majority of coloured people living in council owned rented accommodation. Further, the housing shortage for these people has resulted in artificially high residential densities, as young single and married people are frequently forced to continue to live in the parental home, when they might have preferred to set up independent households. Secondly, SES scores are based on information about all coloureds in Greater Cape Town. The scores therefore represent variations among this particular group and are thus a suitable tool for a status study of a single racial group. Also they are made up of three variables each of which reflects the actual distribution patterns among the Cape Town coloured population and the TMS developed SES scores may therefore be considered a balanced and representative indicator of social status, especially as the components are related to each other and tend to be complementary in affecting status.

Problems concerning evaluation of socio-economic status which might arise out of the forced extended family situation among the coloureds is avoided by the use of SES scores. Since the scores are based on all income recipients there is little loss of information. In the first place, the proportion of non-income earners who markedly affect the social status of a family, or who have high social status is likely to be small. In the second, every individual, male or female, who may contribute to the household economic status is included. Among the coloured group in Cape Town this means that there is no loss of information such as would occur if status were evaluated by family or head of household.

Low average incomes mean that there are often more than one income earner in a family, and this is reflected by the nearly bi-modal nature of the profiles of 11% of the ESD's. It is possible that two groups of differing status within a single ESD may reflect differences between male and female income recipients. If this is the case in any one household it may in fact lead to an underestimation rather than an overestimation of the amount of money available in a family for the acquisition of outward symbols of social status, e.g. housing.

Although SES scores were calculated for a particular time (data were drawn from the May 1970 census), this is not likely to affect status rankings at the present time. The aim of the study is not to monitor individual scores or to consider the values themselves but to use SES scores as a tool for differentiating among the coloureds according to status. The method used for calculating SES scores, in which the averages of three cumulative percentile distributions are combined means that even if actual scores alter, relative status position is unlikely to change markedly over time.

Finally, not only are scores available as averages for each ESD, of which the population is also known, but the frequencies by groups of ten are also given. This means that a profile of the socio-economic character of every ESD may be constructed. As data are available by ESD for the entire coloured group area of Greater Cape Town it is possible to locate in physical space where coloureds with a particular SES score are living and to use the SES scores to establish whether spatial patterns of socio-economic differentiation emerge in the demarcated coloured residential areas.

Social stratification has been shown to occur in any society (Tumin, 1967). One means of defining or recognising the social strata which emerge is social area analysis (broadly defined). On the basis of one or more specified criteria, this approach identifies differing groups or strata in a population. These groupings or divisions take place in social space and it is debated whether they can be translated into, and easily recognised in, physical space. The criteria by means of which individuals in the population are classed as similar or not are termed social indicators. The definition and derivation of these indicators has given rise to much debate primarily in terms of the theoretical framework within which their derivation and definition are couched and therefore in terms of their reliability and representativeness as measures of the state of the society.

The question of the applicability of the social area concept to the South African situation (and particularly the coloured population group of Greater Cape Town) has been raised and the identification of an appropriate social indicator for grouping this population into social areas has been discussed.

The aim of the present study is to establish whether, since the implementation of the Group Areas Act, with residential confinement to specific geographical locations, social areas can be identified within the coloured community and whether these differences are reflected in physical space as well. An investigation of the history and development of the coloured people is therefore needed to establish whether social stratification among these people has been recognised in the past and therefore whether it is likely to occur today.

CHAPTER II

PERSPECTIVE ON THE COLOURED POPULATION

IN THE GREATER CAPE TOWN AREA

Social differentiation within any society results in the grouping of similar individuals in social space, into social areas. These may find their spatial expression reflected in the type and quality of housing occupied by a particular group at a particular time. In cases where this occurs, housing type can be used as an indicator of social differentiation. As the social and economic situation improves for this group, it moves into better housing, vacating accommodation in lower status areas for occupation by another less well-off group. It is by means of this filtering process in housing that distinct social areas develop within the residential structure of a city (Berry & Horton, 1970). The purpose of this research is to establish whether or not this filtering process has been operative in Cape Town, especially in terms of the coloured group and whether housing type may be used as an indicator for identifying social areas in both social and physical space. In order to understand the complexities of the coloured population and the nature of social stratification in the broader milieu at the Cape as well as within the group itself, it is necessary to examine the evolution of the coloureds within the framework of Cape Town society as a whole.

In the Republic, with the exception of some Malays, coloureds are the product of miscegenation among the ethnic groups (Davenport, 1977). There is general agreement about their origins. Marais (1957); Wilson & Thompson (1969); Whisson (1972) and Davenport (1977) indicate that the indigenous Khoikhoi had largely disappeared from the Western Cape as autonomous entities by the late 18th century. Defeated in battle and decimated by smallpox, they intermarried with slaves and others, including whites (Davenport, 1977). Marriages between whites and half-castes - descendants of the Free Blacks (Boeseken, 1977) and liaisons between black and white (Davenport, 1977), also contributed to the formation of the Cape Coloured people.

The evolution of the coloured population may be divided into four periods, each of which is associated with distinct changes in society as a whole.

I). The period of Dutch occupation at the Cape (1652 - 1806) when society consisted of a servile group (slaves) and a non-servile group.

II) The period of British occupation (1806 - 1910) when social stratification became more complex as a result of the introduction of British customs, and the abolition of slavery.

III) The period of South African government prior to the implementation of the Group Areas Act (1910-1950's). In this period social differentiation became increasingly complex as a result of greater industrialization and increased urbanization.

IV) The period after Group Areas implementation to the present time - in which ethnic groups have been separated into geographically defined residential areas, and normal forces operating to cause stratification no longer apply to society as a whole, but may do so within the distinct ethnic groups.

The importance of social change which has occurred within these periods, lies in the effect which it has had on social stratification in Cape Town society as a whole, and within the coloured population group in particular. The physical manifestation of this stratification in western society and cities and is assumed to be recognizable in type and place of residence which shows distinct patterns of differentiation in each of these eras.

I The Period of Dutch Occupation at the Cape 1652 - 1806

The pattern which emerges in this period is one of stratification or separation in the whole society on the basis of servility or non-servility which was (with the exception of the Free Blacks) coincident with colour and originally with Christianity. While Mentzel (1921) distinguishes four classes within White society at the Cape (Company officials, free Burghers, wine and corn farmers and pastoralists, i.e. trek boers), class distinctions in the European sense did not apply within the White community (Wilson & Thompson, 1969). The White colonists were essentially a single community with a common heritage (Northern European Protestant) strongly unified by the Dutch Reformed Church.

The other group of non-servile people at the Cape were the Free Blacks. These were slaves who had been manumitted and were free to live as and where they chose. Table 1.1 shows numbers of Free Blacks in the Cape district during this period. Boeseken (1977) mentions several cases of Free Blacks living among the Free Burghers in the 17th century and describes the homes of at least two families of Free Blacks living in Zee Straat. These people were entitled to all the rights of free citizens at the time. "The Free Blacks in the 17th century were not treated as a separate group" (Boeseken, 1977, p.97).

TABLE 2.1 Free Blacks in the Cape district

Year	Men	Women	Children	Total	Prop. Free Burgher Population
1685	4	2	2	14	7,6%
1711	18	18	6	42	6,0%
1730	55	61	103	219	16,3%
1750	94	98	157	349	15,8%
1770	97	139	109	345	12,8%

Source: (Elphic & Giliomee, 1979)

These people, both White and Black were all free citizens, property owners, and slave owners or the masters of Hottentot servants. They were thus distinctly different from the servile class of slaves and Hottentot labourers. This distinction was accentuated by a legal and social divide separating the free men who had civil rights, i.e. the right to marry, own property and bring or defend an action in a court of law from slaves who had no civil rights (Davenport, 1977; Wilson & Thompson, 1969).

During Van Riebeeck's sojourn at the Cape schools were established to improve the working and selling potential of slaves but other children could attend. As early as 1662, however, slave and white school-going children were separated, as a school was established for whites (Laidler, 1952). On the other hand while there was a difference in type of housing for the two different social classes, there was no spatial separation of the groups in terms of separate quarters of the town for the different classes. Free men lived in houses, in town or on their farms, behind which were their slaves' quarters. The slaves belonging to the Company lived in the slave lodge just below and on the town side of the gardens (Figure 2.1). Although domestic slaves were generally better treated than predial (Davenport, 1977) or Company slaves (Laidler, 1939), there is no evidence of stratification among them. Slaves who came from the Rio de la Goa were considered inferior and kept in the basement of the slave lodge (Mentzel, 1857) but this was probably a distinction drawn by the Company rather than a self-developed pattern among the slaves.

II The Period of British Occupation at the Cape 1806 - 1910

This period is marked by distinct social change. Firstly, the previously unified and dominant but essentially white sector of the society became more diversified as a result of the introduction of English norms and customs (Wilson & Thompson, 1969). Secondly, it was during this period that the emancipation of slaves took place. This meant that all men were free with equal legal rights and therefore, in terms of Ordinance 50 of 1828, could own land (Davenport, 1977). Their socio-economic status would thus no longer be predetermined.

At the beginning of this period, the total population of Cape Town was 16 451 of whom 39% were Europeans. Figure 2.2 shows population composition, the growth and the increasing proportion of whites throughout the period until by 1904 the total population had reached 188 781 of whom 55% were white.

According to Davenport (1977) the Cape during British occupation was dominated by an English-speaking element, largely urban, official, mercantile and professional in character, which rested on a numerically larger but recessive Dutch (Afrikaans)-speaking community. Both these groups in turn depended on the services of coloured people - the Malay artisans and fisherfolk and coloured labourers. Although

Boeseken (1977) maintains that all traces of the Free Black population was lost in the 18th century, Davenport (1977) claims that the Khoisan, freed slaves and half-castes were all identified as "free persons of colour" after emancipation in the 1830's.

Stratification in the community as a whole still existed. Within the white group there seem to have been different economic and social levels. This is evident in the letters of Lady Anne Barnard (Wilkins (ed.), 1910) especially between British Government officials and the Afrikaner Burghers. Shorten (1963) also indicates a commercial class, a class of tradesmen and a poor white group. It would appear that the division became one based essentially on occupation which in turn reflected a colour division. Table 2.2 shows that in 1891 population "other than white" dominated in domestic (60%) and agricultural (76%) types of occupation, while whites dominated (90%) in the professional occupations. The implication is that people of colour were generally labourers. Coloured workers were also paid lower wages than whites for the same jobs (Houghton & Dagut, 1972, p.194-5).

The picture was essentially unchanged in 1904 but the Malays were differentiated as a separate group and proportionally more were involved in commerce (10% vs. 6,5%) and industry (33% vs. 21.5%) than were the coloureds (Table 2.3).

Stratification is thus evident in terms of occupation and although non-white people were clearly of a lower economic and social status than the whites some degree of stratification did exist.

The economic and associated social stratification within the Cape society as a whole also emerged in the residential layout of Cape Town and the outlying suburbs or villages. Although the population in all areas was mixed, the trend during the British Occupation appears to have been one of a general reinforcement of White dominance in the

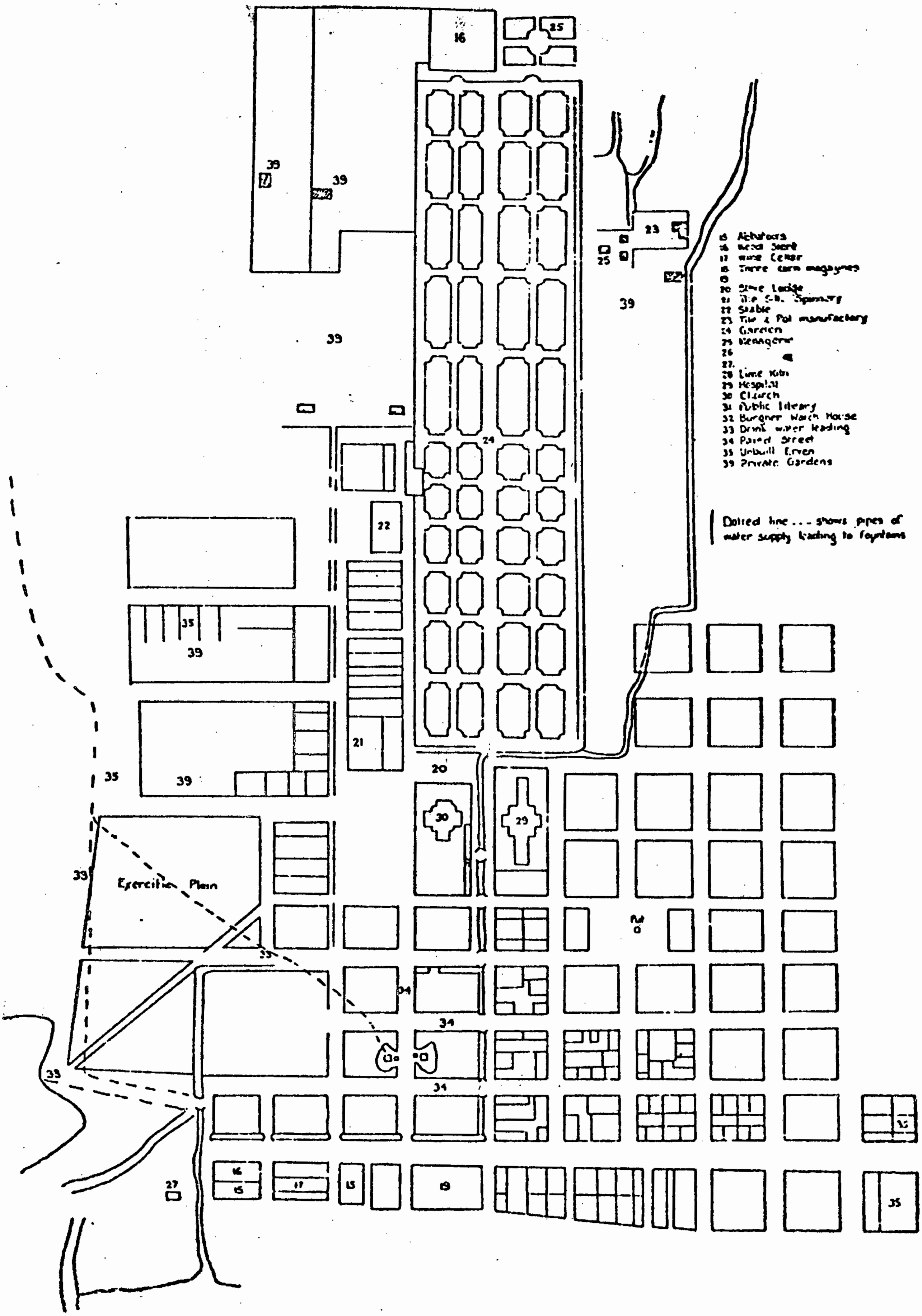
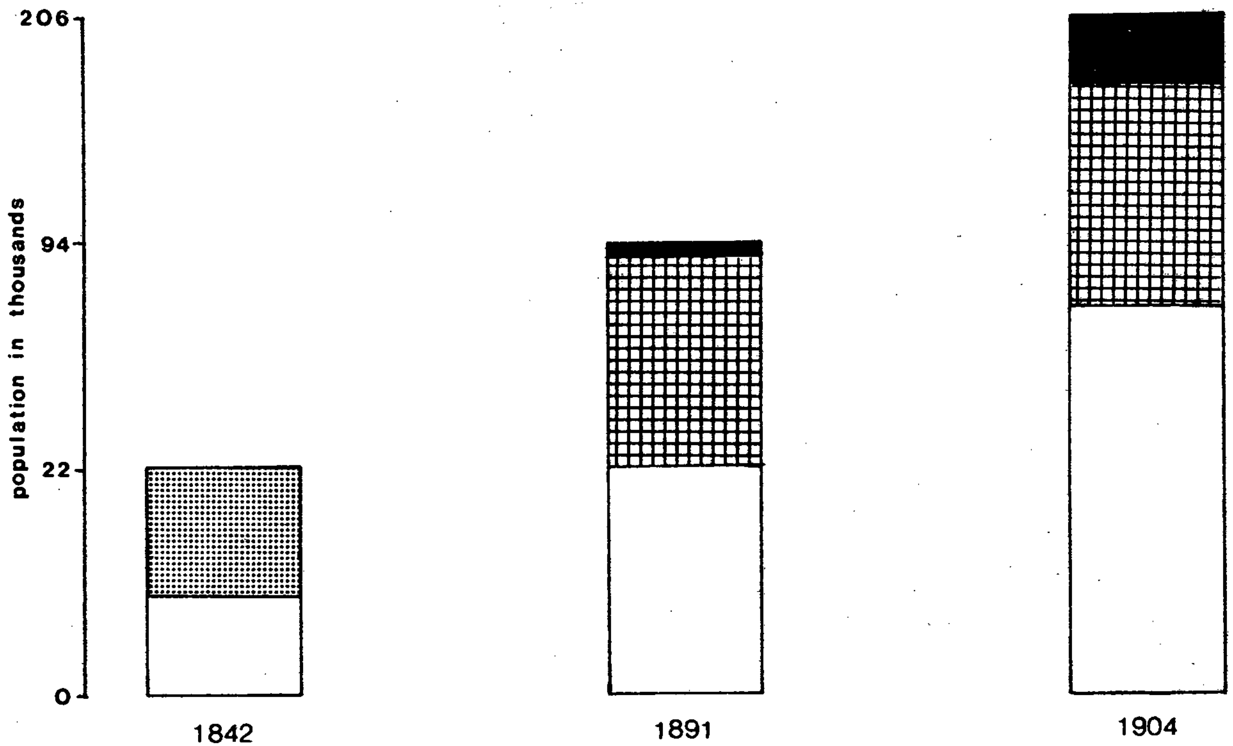


Figure 2.1 Plan of Cape Town 1767
Source: Laidler 1952



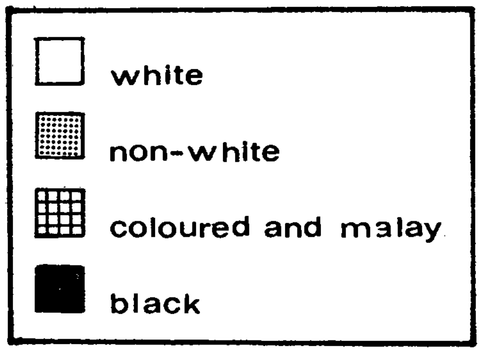


Figure 2.2 Population composition of Cape Town during the British occupation 1806 - 1910

Sources: Cape of Good Hope Blue Book 1891
Laidler, 1952
RSA Bureau of Statistics 1960

TABLE 2.2 Occupations of "adult" population in Greater Cape Town 1891

Class	Total Population No.	%	White No.	%	Other than white No.	%
Professional	10 713	6	9 638	90	1 075	10
Domestic	78 075	45	31 567(F)	40	46 508(F)	60
Commercial	16 557	10	12 378(M)	75	4 179(M)	25
Agricultural	20 263	12	4 834(M)	24	15 429(M)	76
Industrial	39 123	23	20 480(M)	52	18 643(M)	48
Unspecified	6 053	4	3 372	56	2 682	44
Total economically active	170 784		82 269		88 516	
Total Population	271 983		135 310		136 673	

M or F indicates dominant sex constituting this class

% indicates percentage of total population represented by each class, and percentage of white and non-white in each class

Source: Cape of Good Hope Census, 1891

TABLE 2.3 Occupations of "adult" population in Greater Cape Town 1904

Occupation	Total Population	White	Malay	Coloured	Other
Professional	33 203	29 942M	66M	1 195	2 000M
Domestic	264 628	96 052F	4 761F	80 823F	40 325F
Commercial	65 442	44 750M	996M	11 633M	8 063M
Agriculural	405 670	104 957M	532M	44 325M	225 856M
Industrial	168 708	65 618M	3 311M	38 313M	61 466M
Unspecified	13 718	7 740	146	1 998	3 843
Total economically active	951 369	349 059	9 812	178 287	341 553
Total Population	1 489 691	553 452	15 615	285 382	635 242

M or F indicates dominant sex constituting this class

Source: Cape of Good Hope Census, 1904

prime locations, e.g. Wynberg and Claremont and a drift of coloured people to areas such as Parow, Maitland, Klipfontein and Serapta on the Cape Flats (Table 2.4). A more affluent community was to be found on better land nearer the mountain and along the coastline (Putterill & Bloch, 1978), in areas such as Green and Sea Point, Camps Bay, Wynberg, Rondebosch and Mowbray. The "handsome mansions of the rich" also stretched from Campground to Claremont in the 1840's (Laidler, 1939). The 1891 census shows that these areas were predominantly white (Figure 2.3) and that the proportion of whites had increased by 1904.

Kalk Bay is described as a fishing port which Whisson (1972) suggests had a large number of coloured residents, while Simonstown was a naval base where in 1857 "the Malay population (approximately 20% of the total) exercised some influence on the habits of the Christian settlers" (Laidler, 1939, p.415), and distinctly separate residential areas for whites, Malays and coloureds occurred within the town itself (Whisson, 1972).

Noordhoek and the surrounding rural areas were residentially mixed, and groups of coloured fishermen lived at Kalk Bay, Kommetjie, Hout Bay and Witsands. Salt River, still countrified in 1881, had fishermen's cottages dating from the days of free fishermen in 1665. That some of these fishermen's hovels were removed from the Rogge Bay area in 1875 (Laidler, 1939) suggests that this sector of the population were extremely poor, and that their accommodation presented both a sanitary problem and an eyesore to the wealthier public who thus discriminated against them.

Davenport (1977) comments that the tag of "free persons of colour" after emancipation provided the Cape coloured people with an identity which many of them continued to regard as artificial. Whether stratification occurred within the coloured group itself cannot be established as there is little evidence or documentation of stratification as such. However, from 1875 the Malays were separately

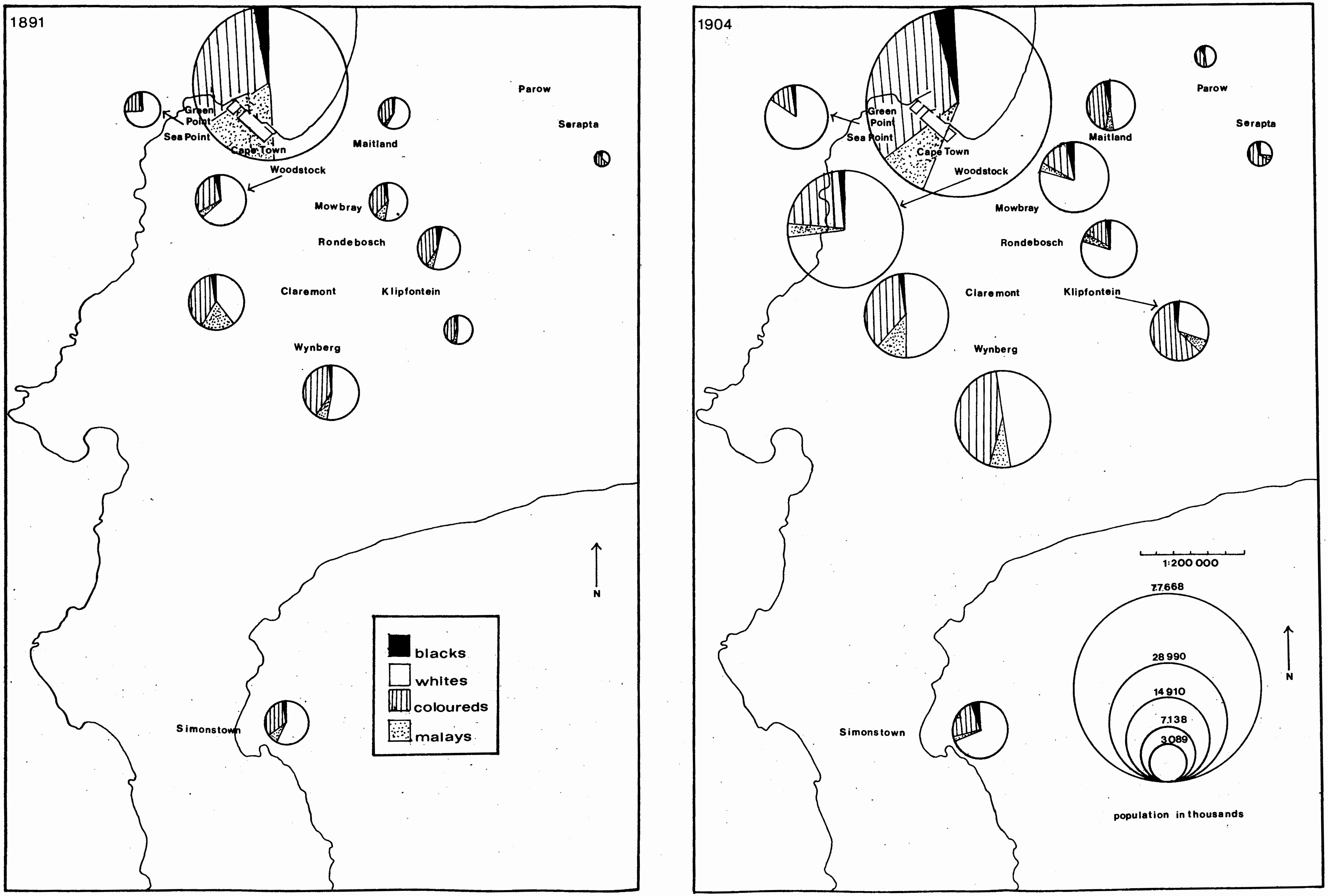


Figure 2.3 Distribution of population 1891 and 1904

TABLE 2.4 Distribution of coloured population 1891 and 1904

Suburb	% of population classified coloured					
	1891			1904		
	45%	45-55%	55%	45%	45-50%	55%
Cape Town		X		X		
Green & Sea Point	X			X		
Camps Bay	X			X		
Wynberg		X		X		
Rondebosch	X			X		
Mowbray		X		X		
Claremont			X	X		
Parow	-	-	-		X	
Maitland	X				X	
Klipfontein		X				X
Serapta			X			X

Source: Percentages derived from census data 1891 and 1904 enumerated. In addition specific areas where Malays lived are mentioned - the Malay Quarter in central Cape Town (Pollock & Agnew, 1963); parts of Claremont and Wynberg (Laidler, 1952). These people clearly formed an accepted stratum of the coloured population; their cultural background and economic status, coupled with their religious cohesion and residential differentiation made them distinctly recognisable.

The fact that some coloureds resided in predominantly white areas and that a limited number were engaged in higher status occupations suggests that they might have become integrated into the community as a whole, while the existence of fishermen's cottages and areas with greater than 60% coloured communities such as Parow, Klipfontein and Serapta could be taken as evidence of a lower social stratum whose residential area was spatially separated from the rest of the community. Ellis *et al*, (1977, p.114-5) mention squatter settlements on the Cape Flats, emphasizing the squalid conditions of Rondebosch extension, Maitland and Blaauvlei in Retreat, in the early 1900's. This implies a fourth and lowest stratum within the coloured community living in distinct shanty areas within the city.

The pattern which emerges by the end of the period of British Occupation at the Cape is one of social stratification primarily based on occupation rather than entirely on the basis of colour (see Figure 2.3). During the period a degree of residential stratification apparently emerged and ethnic separation became evident to a very limited extent as reflected in recognition of the Malay groups.

III The Period After Union and Prior to the Implementation of the Group Areas Act (1910 - mid-1950's)

Between 1910 and 1950, during the period of Union, although there was a numerical increase of 203 716 coloureds the relative proportion remained more or less constant at 46% (Table 2.5). It was only at

TABLE 2.5 Cape Town Coloured Population Growth 1911 - 1951

	1911	1921	1936	1946	1951
Coloureds	39 302	108 489	169 392	228 944	297 018
Total Population	197 332	248 484	372 744	514 040	632 013
Coloureds as % of total	47%	44%	45%	45%	48%

Source: Bureau of Statistics, 1960

the time of the First World War (1914 - 1918) however, that problems of overcrowding were recognised in Cape Town and the first scheme for Council Housing (Bloemhof Flats in District Six) was put foward by the Health and Building Regulations Committee in 1916 (Shorten, 1963). The 'Flu epidemic of 1918 apparently brought home to ratepayers the conditions under which a large section of the population was living and prompted the Provincial Council to empower the City Council to grant housing loans to people of limited means. This could not have affected many coloured people, but for those able (economically) to take advantage of such a scheme, it meant the ability to own a home and thus the stratum within the coloured community was extended to those able to own their own homes.

Although Council Housing for employees was undertaken on a small scale at Maitland, Claremont and Mowbray and for coloureds in general and blacks at Milner Estate (now Gleemore), it was not until 1929 that the City Council considered schemes for sub-economic housing. In 1931 the flat building schemes aimed at clearing slums and alleviating accommodation pressure in District Six and in 1933 the cottages that represented the City's first letting schemes in the Athlone area (Bokmakierie) were completed. In 1934, the Slums Bill became Act No. 53 of that year (Shorten, 1963). This meant that housing classified as being of slum standard had to be renovated or demolished and better housing provided for the inhabitants. Between 1939 and 1949 housing schemes were initiated for lower income families to the east of the main axis of the southern suburbs and concurrently housing for low-income coloured families was established on the Cape Flats. The provision of mass housing at Kewtown set the pattern of fringe development and housing estates continued to be built on the Cape Flats from Athlone to Grassy Park and along the southern side of the Goodwood-Bellville railway line, from Elsie's River to Bellville South (Putterill & Bloch, 1978). In District Six and the Malay Quarter, repairs and reconstruction were undertaken (Shorten, 1963).

According to Marais (1957), the State, since Union, has influenced the employment of skilled coloured labour in two ways: firstly, by discriminating against them in Civil Service appointments; and secondly, by laying down the conditions of employment in private undertakings. This has been achieved through a series of Acts such as the Apprenticeship Act of 1922, the Civilized Labour Policy and the Industrial Conciliation Act of 1924 and the Wage Act of 1925 (Houghton & Dagut, 1973). All of these Acts constitute barriers to high level economic achievement for Coloured people. The result is that in the main, coloured people are poorer and cannot afford housing in the more favoured areas of the city. They are thus forced to locate in the low rent and sub-economic Council Housing schemes (described above) in the peripheral areas of the city, or in older houses immediately adjacent to the commercial centre (the Malay Quarter and District Six).

Ethnic grouping of residences was not really well defined until significant urban growth got underway (Putterill & Bloch, 1978). However by 1946 the coloured population in Cape Town numbered 228 944 (45% of the total population) of which 57 236 or 25% were economically active, and by the 1950's a hierarchy of residential areas occupied by coloureds had emerged.

Analysis of the maps produced by Scott, (1955) (Figures 2.4, 2.5 and 2.6) shows that there is a distinct relationship between these coloured residential areas and quality of residence. High and medium grade residential districts are also areas of detached houses and appear to have been occupied almost exclusively by whites. Less than 20% of the coloured population group were located in areas of high grade detached housing in Rondebosch, Claremont and Wynberg. These form well defined pockets in predominantly white areas and reflect a difference in income within the coloured population group evident in the ability of this small percentage to live in high grade housing in the more favourable areas. The upper stratum would include home-owners such as those in the Milnerton area. This supposition is related to the relationship between economic status and residential differentiation based on the fact that approximately 2% of the economically active population are employed in professional occupations (Table 2.6)

In addition clear evidence of spatial separation on ethno/religious grounds within the coloured group occurs in that distinct areas of Malay population (Figure 2.4) are defined. The Malays are described by Scott (1955) as the aristocracy of the Cape coloured people due to the fact that traditionally they have been skilled craftsmen and thus in a better economic position than most other coloured people. Their religious homogeneity may have given them a distinct social and residential identity as they have tended to cluster in the areas around their mosques. Approximately half of the areas of coloured residence however, comprised semi-detached houses, terraces and tenements and are located near commercial sub-centres and along the suburban railway line where they are classified as "old" and "low

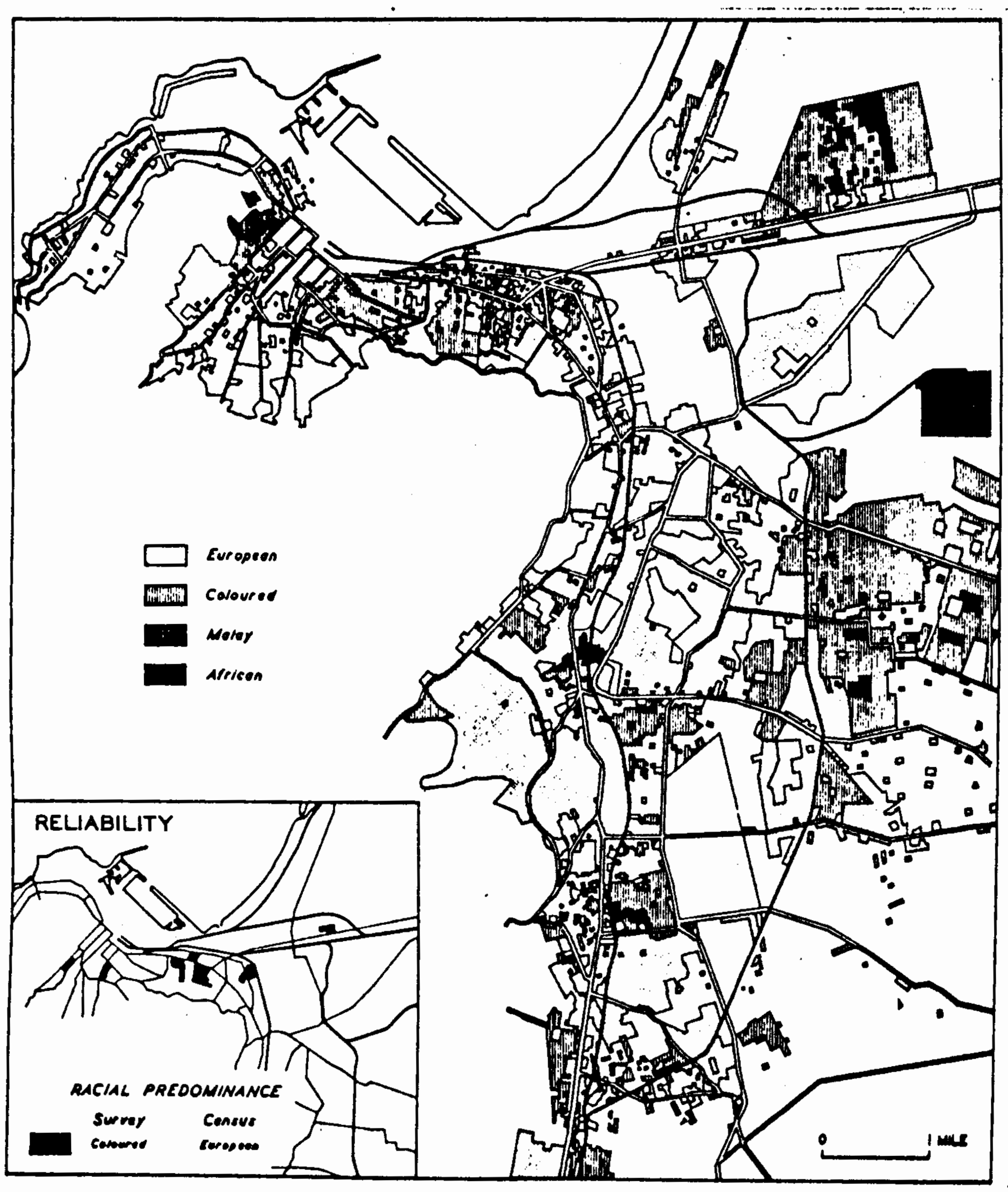


Figure 2.4 Ethnic areas in Cape Town
Source: Scott (1955) Figure 4

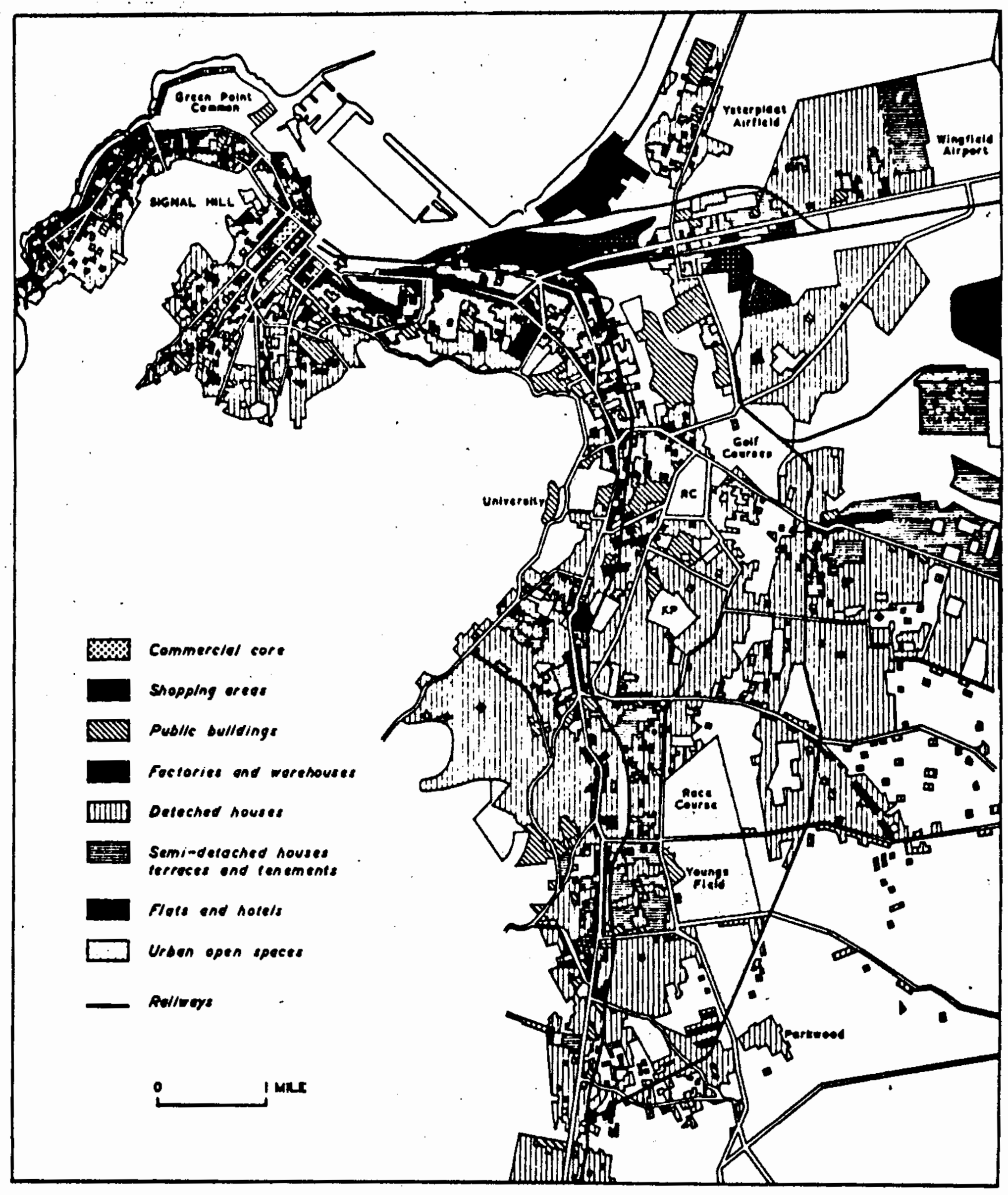


Figure 2.5 Functional areas in Cape Town
Source: Scott (1955) Figure 5

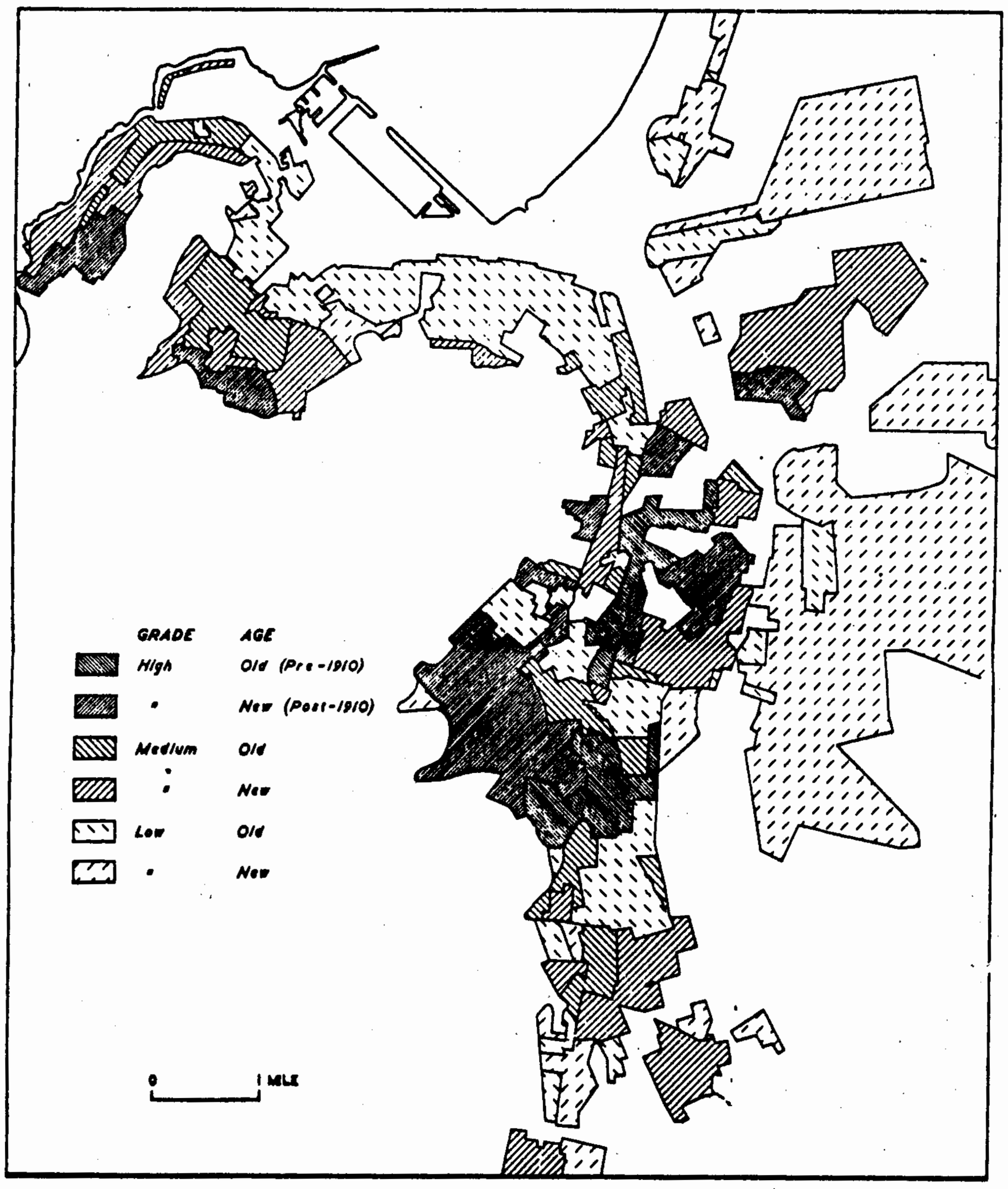


Figure 2.6 Residential districts in Cape Town
Source: Scott (1955) Figure 3.

stratum were the squatter elements (Ellis *et al*, 1977) which even by 1942 were located in areas such as Philippi, Grassy Park, Elsie's River and Windermere.

Social and residential differentiation were therefore well developed in the coloured community and according to Marais (1957) and Whisson (1972) the spatial expression of this stratification - residential patterns - was determined largely by economic criteria prior to the implementation of the Group Areas Act of 1950. Economic restrictions imposed on the coloured population group during the period between Union and the middle 1950's resulted in reduced potential to obtain housing or accommodation in the favourable areas for the majority of the coloured population. It did not, however, prevent those who could afford it from living where they chose nor did it residentially separate white and coloured on the same economic stratum. It allowed for a pattern of stratification within the coloured community not as a separate entity but as a part of the wider society of Greater Cape Town.

IV The period from the Implementation of the Group Areas Act (mid-1950's) to the Present

The advent of the Nationalist Government in 1948, with its ideology of separate development for the various population groups (Davenport, 1977) marks another turning point in the social history of South Africa and the residential structure of its cities. The Group Areas Act (No. 41 of 1950) provided for the legal separation of residential areas, in which occupation or ownership or both are restricted to persons belonging to a specified racial group (S.A.I.R.R., 1951).

In Greater Cape Town the coloured population group was thus forced to move into the areas set aside for them forming artificially created and geographically defined ethnic areas. The first Group Areas proclamations in the city were not gazetted until 1957 and these did not deal with all the Greater Cape Town area nor with the entire area mentioned in the Group Areas proposals for Cape Town in 1954-55. These proposals were that the entire area between the Atlantic Coast and the suburban railway as well as a considerable proportion of land east of the railway line, between Observatory and Wynberg, be set aside for whites. With the exception of Bellville South and a strip along the south-western border of Bellville municipal area, Bellville too, was to be reserved for whites. Battswood estate in Wynberg was to be reserved for Malays, while the coloured population was to have Bellville South and the south-western border strip, certain portions east of the railway line between Diep River and Retreat and near Zeekoe, Princess and Ronde Vleis. Areas on the Flats were also set aside for coloured group areas. Table 2.7 lists the Group Areas proclaimed since 1957 while the present distribution of coloured group areas is shown in Figure 1.

As a result of the Group Areas proclamations, large numbers of coloured people have had to move out of established homes in designated white areas and find accommodation in the housing schemes and in other coloured areas on the Cape Flats. Exact numbers of people moved by Group Areas proclamations are not available, but figures quoted by the South African Institute of Race Relations give some idea of these. In 1961 2,6% of the white population as against 25,7% of the coloured population were affected by Group Areas proclamations (Table 2.8a). By 1963 proclamations of Group Areas in 1957, 1958 and 1961 involved the removal of approximately 7 500 whites, 94 200 coloureds and 4 700 asians, and by 1970 just over 15 000 coloured families had been resettled according to the Minister of Community Development.

The scale of the operation is clear when 7 133 coloured families had been moved from District Six alone by 1977 and 2 534 were still to be

TABLE 2.7 Group Areas Proclamations for Greater Cape Town since 1957.

1957-1961	Proclaimed White	Entire Table mtn. area, including Sea Point and other Atlantic Coast Suburbs. Partly industrial and mixed residential areas east of city centre - Brooklyn, Pinelands, etc. Entire area NE of railway line to interior and branch to Somerset West.
	Proclaimed Coloured	Housing schemes at Athlone and other areas on Flats (inland from railway line to Muizenberg) and at Kensington and inland from Brooklyn. Bellville South and undeveloped land on Flats. NB Malay Quarter separately classified for Malays.
1961	Proclaimed White	Entire region west of suburban railway (except small area adjoining Athlone and Battswood Estate) - from Salt River to Retreat, and Hout Bay to Zeekoevlei, Ottery and land round Zeekoe and Rondevlei.
	Proclaimed Coloured	Battswood Estate, land to S. of Athlone, Retreat. Lotusrivier Estate area, including Princess and Little Princess Vlei.
1963	Proclaimed White	Epping Garden Village
1965	Proclaimed White	Part of Lotus River Area
	Proclaimed Coloured	Grassy Park. Part of Lotus River area. Land in Wynberg adjoining Athlone.
1966	Proclaimed White	Major Portion of District Six. Frazerdale (between Rosebank & Rondebosch and Athlone), Claremont above railway line.

continued -

TABLE 2.7 (continued)

1967	Proclaimed White	Kommetjie, Simonstown, Fish Hoek, Kalk Bay, St James, Clovelly, Lakeside, Muizenberg
	Proclaimed Coloured	Slangkop (Ocean View)
1968	Proclaimed White	Large part of Woodstock, entire coastal strip from Three Anchor Bay to Duiker Point (including Sea Point, Bantry Bay, Bakoven, Oudekraal, Sandy Bay, Hout Bay Beach)
	Proclaimed Coloured	Hout Bay harbour area. Further areas on Flats

Source: Survey of Race Relations in South Africa, SAIRR, 1961, 1963, 1965, 1966, 1967, 1968.

TABLE 2.8

a) Numbers affected by Group Areas in Cape Town 1961

Group	1957	1958	1961	Total
White	100	1 829	5 442	7 371
Coloured	14 900	34 604	44 644	94 148
Asian	1 000	1 511	2 147	4 658
	16 000	37 944	52 233	106 177

b) Numbers to be resettled by Group Areas in South Africa 1972

Group	NUMBER OF FAMILIES		
	Disqualified	Resettled	To be resettled
White	1 598	1 433	165
Coloured*	76 544	41 199	35 345
Indian	38 561	26 294	12 267
Chinese	1 233	68	1 165

* 30% coloured population of South Africa lives in Greater Cape Town

continued -

TABLE 2.8 (continued)

c) People in South Africa Removed in terms of Group Areas Act, 1975

Group	Families	Persons
White	15	47
Coloured *	5 631	29 281
Indian	677	3 385

d) People in Cape to be resettled in Group Areas 1976

Group	Families	Persons
White	365	1 351
Coloured+	47 061	244 718
Indian	2 134	10 670
Chinese	142	526

* Note 30% coloured population of South Africa lives in Greater Cape Town

+ 35% coloured population of Cape Province lives in Greater Cape Town moved. Today the total estimated coloured population of Greater Cape Town is 821 430 (Ferrario, 1979) and despite the moves 9,3% of these people still do not live in proclaimed coloured group areas. The majority however, do live in the thirty designated suburbs shown in Figure 1.

The change in residential location is reflected in population numbers of what are now the predominantly coloured residential areas of Greater Cape Town when compared with these areas in the era prior to Group Areas implementation. From Table 2.9 it may be seen that while the Cape Town and Simonstown areas have maintained a fairly constant proportion of coloured population largely reflecting natural population increase a marked increase in population occurred in Bellville, Wynberg and in Grassy Park between 1951 and 1970. These suburbs contain a large proportion of designated coloured group areas, but unfortunately detailed figures are not readily available over the time period for individual suburbs, neither are they directly comparable.

TABLE 2.9 Percentage population coloured in selected areas of Greater Cape Town 1911 - 1970

Area	1970	1960	1951	1946	1936	1921	1911
Mag. Distr. Bellville	59	57	49	50	47	53	55
Mag. Distr. Cape Town	45	48	44	42	43	39	44
Mag. Distr. Simonstown	33	31	42	35	47	37	35
Mag. Distr. Wynberg	58	53	50	53	53	55	53
Greater C.T. Total	54	51	44	42	43	39	44

Source: Bureau of Statistics 1960, 1970

TABLE 2.10 Occupational grouping of coloureds in the 01 economic region

Year	Total Population	Economically active	Profess-ional	White collar workers	Skilled labour & production workers	Service workers	Unskilled labour
1970	606 075	232 867	6 177	133 172	130 976	41 491	6 489
		100%	3%	14%	56%	18%	3%
1960	417 414	165 622	4 559	5 131	75 677	45 173	7 514
		100%	3%	3%	46%	27%	7%

Source: Bureau of Statistics, 1970, 1960

TABLE 2.11 Types of housing in coloured Group Areas

Suburb	Rented			Owner occupied	
	Sub-economic	Economic	Private	Lower income	High income
Malay Quarter	X	X			
Walmer Estate			X	X	
Wynberg				X	X
Heathfield					X
Retreat	X		X	X	
Lavendar Hill	X				
Parkwood	X	X			
Grassy Park	X	X			
Mitchells Plain				X	X
Lansdowne			X	X	
Hanover Park	X				
Athlone	X	X		X	
Kewtown,) Bridgetown) Silvertown)		X			
Vanguard & Welcome				X	
Heideveld	X	X			
Manenberg	X				
Bonteheuwel	X	X			
Bishop Lavis			X		
Matroosfontein				X	
Elsies River	X				
Bellville South				X	
Kensington & Factreton	X	X	X	X	
Belhar					X
Charlesville					X

Sources: Ferrario, 1979; SAIRR, 1950-1978; Divisional of the Cape, 1978

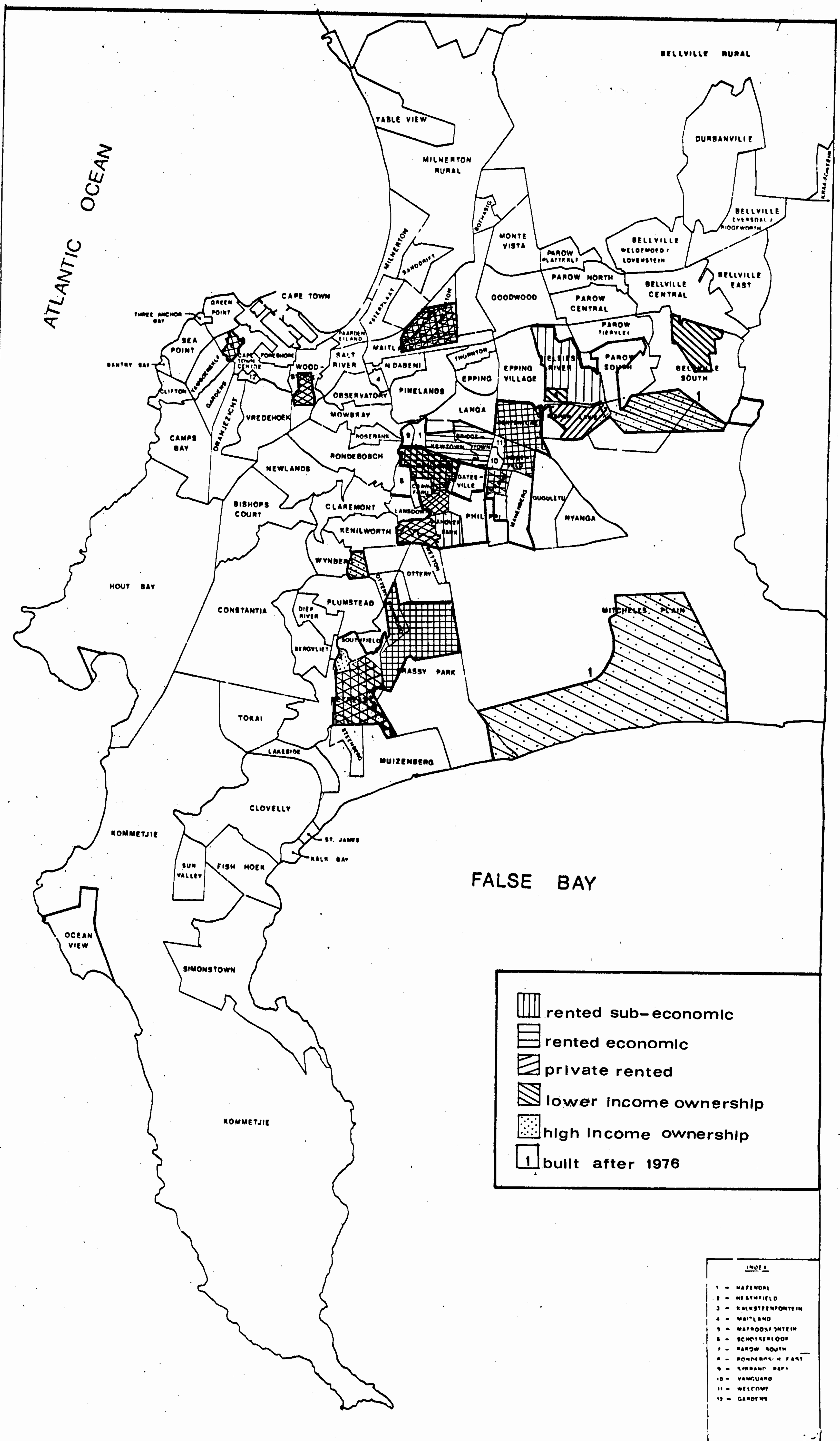


Figure 2.7 Distribution of housing types in coloured Group Areas

Over the 20 year period from 1951 to 1970 a change in occupation also occurred for the number employed as skilled labourers or production workers rose by 10% to 56% in 1970, with a concomitant drop in the proportion of unskilled and domestic work (Table 2.10). Also, while the percentage employed in professional capacities remained constant at 3%, that of white collar workers increased 5 times to 14% in 1970. It appears therefore that there is a relatively wide range in economic status among the coloureds and therefore that the ability to pay for accommodation varies within the group. However, for many coloured people, the forced moves have meant a decline in their economic potential as a result of losses on previous properties (S.A.I.R.R., 1970) and thus a reduced ability to find comparable housing elsewhere ((Western, 1978). Much of the area set aside for coloureds was to a large extent undeveloped (S.A.I.R.R., 1971). Although many people have been obliged to seek accommodation in Council Housing Schemes, which operate on an allocative basis, a variety of housing types and accommodation may be found (S.A.I.R.R., 1950 - 1978; Ferrario, 1979; Cape City Council, 1977; Divisional Council, 1979). These vary from sub-economic rented accommodation to affluent, privately-owned dwellings (see Figure 2.7 and Table 2.11).

The Theron Commission Report finds evidence of social and residential stratification within the coloured community, (Theron Commission Report, 1976) and identifies four broad status categories:

1) The top group, about 20% of the coloured population, consisting of employers, professionals and skilled white collar workers who would be able to live in private (affluent) dwellings in areas such as Belhar, Charlesville, Squarehill, Elfindale or Puntz Estate (Heathfield), Wynberg or in Mitchells Plain.

2) The middle group, about 40% of the coloured population, which may be divided into upper and lower halves:

(a) The upper half consists mainly of semi-skilled workers who would probably have the finances to live in lower income home ownership schemes such as Mitchells Plain, Matroosfontein, Walmer Estate, Kensington, Bellville South or Lansdowne; or in economic rented dwellings in areas such as Kensington and Factreton, Vanguard and Welcome, Bridgetown, Kewtown and Silvertown and Athlone.

(b) The lower half who are mostly unskilled workers, whose household incomes are not much more than the supplementary living level and would only be able to afford rented sub-economic dwellings in areas such as Bonteheuwel, Parkwood, Elsie's River, Manenberg, Hanover Park, Heideveld, Bishop Lavis and Lavendar Hill.

3) The bottom group, also about 40% of the coloured population, who are unskilled labourers and often unemployed or underemployed. They live in a situation of chronic poverty. Although some might live in sub-economic rental areas as mentioned above, the majority including the under- and unemployed are located in the squatter camps in areas such as Vrygrond, Strontyard, Lotus River area, Elsie's River and Ravensmead, where the total number of coloured squatters is between 120 to 180 000 (Ellis *et al*, 1977).

A hierarchy, corresponding to the economic (reflecting census occupational groupings) and social strata of the coloured community in the Greater Cape Town area may be recognised. The implication of the existence of distinct economic strata within the coloured population group is that there is differential ability to pay for housing or accommodation within the coloured group areas. This suggests that because a gradation of housing types is available there is the possibility of a filtering process in the coloured community. It therefore implies that distinct social areas, corresponding to the different social strata could develop within the coloured group areas.

The history and development of the coloured population group of Greater

Cape Town falls naturally into four distinct eras, each of which has shown changes in social and economic status for the coloured population group. Throughout that time although the largest proportion of coloureds were in the lower economically defined strata, it was primarily between 1910 and the mid-1950's, that legal restrictions began to influence the economic situation of the coloured group. The majority of the coloured group became entrenched as the lower echelons of society and with reduced economic potential, class and race became largely synonymous. Better defined ethnic areas began to develop as the coloured lower income groups were forced to reside in the cheaper, less favourable residential areas. Since the implementation of the Group Areas Act in the mid-1950's ethnicity of certain residential areas has been further strengthened and the coloured community has been forced to live in an artificially created and geographically well defined area. Stratification still exists within the coloured community much as it did in the past and is still primarily a reflection of economic status. As a hierarchy of housing types has been identified this socio-economic stratification may still find its spatial expression in the emergence of distinct residential areas within the coloured group area.

The nature of the patterns of social differentiation among the coloureds in the wake of the Group Areas removals have yet to be determined. Using socio-economic status scores derived from the 1970 census data by the TMS of the Cape Town City Council as indicators of social differentiation, this study attempts to determine and examine the patterns of socio-economic stratification which emerge both in social and physical space for the coloured population group of Greater Cape Town.

CHAPTER III

PATTERNS OF SOCIO-ECONOMIC STATUS

Geographical phenomena distributed in space may take the geometric form of points, lines, areas and volumes and their distribution can be classified according to intensities, associations between them or flows between phenomena (Harvey, 1973b). Great emphasis has been placed on the measurement of locational patterns or areal distributions as the focal point of contemporary human geography (Smith, 1975). These patterns can be summarised in various ways, compared with one another and finally used as a tool to synthesize the whole. Synthesis in geography can be interpreted as assessing the extent to which pattern might suggest process, or the links of cause and effect which lead things to be as they are (Harvey, 1973b; Smith, 1975).

I Pattern Recognition

Although patterns of points, lines and areas have all been analysed in terms of nearest neighbour techniques (Walford, 1973) these methods, along with quadrat sampling have been used primarily in the analysis of point patterns. Nearest neighbour provides a simple, objective technique for measuring spatial pattern, while quadrat sampling is concerned with the probability of finding 0,1,2...n, points in an area of a given size (i.e. a quadrat). Both provide a convenient way of objectively describing some of the general characteristics of point patterns.

The patterning of areas in particular however, presents serious problems. Analysis tends to be descriptive rather than analytic because of the mathematics involved in numerical analysis. Patterns are usually therefore analysed by descriptions of their relationship to one another or other patterns in terms of their location and variation and the values of the variable. Techniques for numerical analysis of areal patterns are however, available (Walford, 1973). By means of the contiguity ratio it may be determined whether statistics for each unit area are distributed at random or whether they form a pattern (Geary, in Berry & Marble, 1968). If the actual ratio found is significantly less than unity (probability under a random distribution) by reference to the standard error, then the units themselves can be regarded as contiguous. The ratio, therefore, establishes not only the fact, but the relative strength of contiguity. The general method Dacey (1968, in Berry & Marble) uses is to apply a variant of the four colour problem to any set of non-overlapping regions. Areas are then tested for contiguity in terms of the presence or absence of the attribute under consideration. Related analyses of associations between areal data by Moran (1950) and Geary (1968) in particular, were the basis of Cliff & Ord's (1969) statistic of autocorrelation. Autocorrelation is a measure of the spatial dependence of a variable distributed over an area. It is determined by means of simple correlation applications within a single data set.

Autocorrelation is of special interest to geographers because it expresses a characteristic of distributions which is inherently spatial (Olson, 1975). She suggests several ways in which autocorrelation in data can be tested. Kendall's Tau coefficient serves reasonably well as a measure of pattern or orderliness for each individual class interval. Several other methods of measuring autocorrelation were however, applied to deal with the pronounced influence of identical neighbours. One method was a simple proportion of identical neighbours. Average difference between neighbouring values is another way of assessing neighbour relationships. In order to indicate the interaction between neighbouring units, the number of clusters of like values may be counted: the fewer the regions or clusters, the greater is the spatial organization. Finally the proportion of identical neighbours method may be adapted in an effort to find a means of giving weight to clusters, but also some weight to smoothness of transition between, and irregularity in shape of, clusters of like values. Every pair of values contributes to the difference between their class values.

Olson (1975) does not aim to provide a quantitative measure that accurately predicts subjective judgment (of map patterns) in an absolutely sound and scientific manner. Rather, her emphasis is on the utility of measures such as those given above, as indicators of the degree of organization rather than orientation or specific pattern.

Analysis of patterning of areas is particularly relevant in terms of this research. The data input for this study (i.e. socio-economic status information) is first to be mapped. Resulting patterns may then be described in terms of their shape and nature. The data will then be subjected to the Cliff & Ord spatial autocorrelation procedure, whereby any spatial dependence in the distribution of social status may be ascertained. Although this procedure might imply some of the underlying processes leading to social differentiation i.e. that these processes result in regular, random or clustered dispersal of social status, to fully explain process, these patterns need to be set in a wider context (Smith, 1975) of which this study is merely the first stage.

Data input for pattern recognition may be a single generalized descriptive statistic (representative of the whole set) or a class of values (within the set) determined in an objective manner. The descriptive statistics are single statistics which are used to summarise the varying values within a data set. One value is chosen to give as reasonable an approximation as possible to what is "normal" i.e. to summarise the data (Gregory, 1973). Three measures of central tendency in common use are mean, median and mode. Two of these measures (i.e. mean and modal class) are available from socio-economic status data. The objectively determined class of value should be based on the full profiles of each ESD. The ESD's may be classified as similar or not similar in profile and thus divided into clusters. Once these have been identified, they may be plotted on maps. The extent to which ESD's in the same cluster are contiguous may be assessed and the existence of spatial patterns thus established. The problem which arises is that of assessing similarity or dissimilarity of ESD's on the basis of the percentage of population falling in each percentile group between zero and one hundred and is part of the more general problem of numerical classification.

a) Classification

Briefly, the aim of numerical classification is to group similar individuals together and to separate them from dissimilar individuals on the basis of objective measurements of certain specified criteria (Berry, 1958). A technique must be selected from the many available for measuring similarity and dissimilarity. Classificatory strategies and a process for determining choice of a suitable technique are summarised by Williams (1971) in the flow diagram (Figure 3.1).

The selection of a classificatory procedure depends partly on the aims of the researcher, but more specifically upon the nature of the data to be classified (Johnson, 1968; Spence & Taylor, 1970; Clifford & Stephenson, 1975). Numerical classification is in this respect, as Johnson (1968) states, subjective.

The data set for this study (full profiles of population in each ten percentile group for each ESD) renders the first phase of the classificatory process - that of classification into broadly similar groups, unnecessary. The first two choices in the flow diagram (Figure 3.1) are thus already determined as grouping of the data is both exclusive and intrinsic.

The second phase of classification according to Spence & Taylor (1970) is that of actually measuring the similarity of individuals "objectively" on the basis of certain specified criteria. The form of the profiles for each ESD (population in ten percentile groups between zero and one hundred) is essentially hierarchical. Further, as the data is a measure of socio-economic status in the coloured areas of Greater Cape Town, any form of agglomeration is likely to result in basic heirarchical groupings of low, middle and high social status. Differences in status imply a hierarchical arrangement. It therefore appears that for this study, a hierarchical grouping method, (choice 3) typically illustrated by a dendrogram (Spence & Taylor, 1970) would be most appropriate. Since there is only one variable, (SES scores) and the data reflect the percentages of the ESD population occurring in each ten percentile group, multivariate and discriminant analysis are inappropriate as classificatory procedures. Similarity of the variate therefore suggests that an agglomerative approach (choice 5) be adopted for this purpose.

Three programmes for calculating similarity coefficients and grouping data by hierarchical agglomerative methods (i.e. performing a cluster analysis resulting in a dendrogram as output) that are readily available include:

1) BMDP (cluster analysis of cases P 2 M)

2) Bray-curtis

3) Canberra metric.

The nature of the data set determines the choice of programme for calculating the similarity coefficient. BMDP (Brown, 1977) is a package programme which is both easily accessible and capable of application to many types of data. The limitations of this programme lie in its nature as a package. Poor documentation gives no specifications of requirements for data input. However it appears that BMDP may be a non-hierarchical cluster programme. Clusters are formed on the basis of one of four distance measures, viz. Euclidean distance, absolute differences, chi-square statistic and phi-square statistic. None of these measures is appropriate for application to the data set for this study. Euclidean distance is a measure of physical, straightline distance. In this study ESD's must be grouped on the basis of their distance from one another in social space (i.e. socio-economic terms). The use of absolute differences maximises the distance between points in the data set and in this case the aim is to minimise the distance. The latter two are specifically non-parametric tests while the data set for the study is normally distributed. The programme was regarded as inappropriate in terms of the particular data set.

Both the Bray-Curtis and Canberra Metric programmes are hierarchical and agglomerative in nature. Both also offer optimal strategies for sorting and clustering data. These include nearest neighbour; furthest neighbour; centroid; median and group average (Lance & Williams, 1967). Centroid, median and group average are inappropriate strategies for this study as they are generalized measures and use of any of these would result in the loss of profile detail.

The sorting strategy chosen for the study was furthest neighbour. Using this method the least similar ESD's are isolated leaving those that are most similar grouped together on the basis of the profile characteristics. More intensive clustering is obtained by this method than would be the case if nearest neighbour sorting were used (Clifford & Stephenson, 1975).

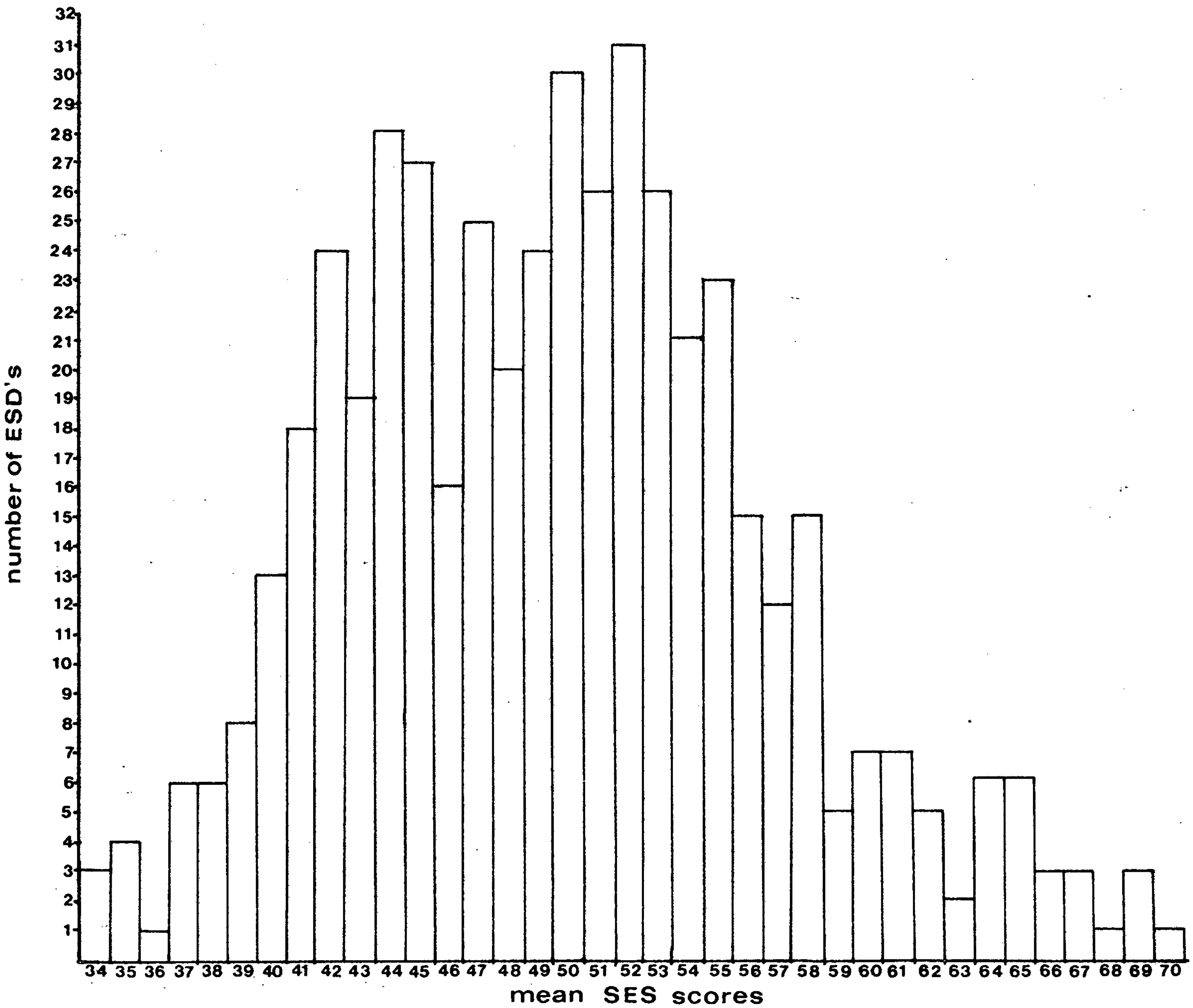
In order to render profiles for each ESD directly comparable without loss of information, numbers of individuals in each class of the profile should be converted to percentages. The matrix thus created has relatively few extreme values, but a large number of zeros in the higher and lower classes. Since the Canberra metric coefficient is more effective for data sets in which there are few zeros but occasional extreme values (Williams, 1971), it too was considered inappropriate for manipulation of the data set in this study. This left the Bray Curtis as the only alternative.

Although it is more sensitive to extreme values, the Bray-Curtis coefficient is not affected by large numbers of zeros in the data matrix. The Bray-Curtis programme was thus considered the most appropriate in terms of the data set to be used. (For copy of programme see Appendix 1.)

As is usual from this type of programme, the output is a listing of clusters of similar ESD's in which the degree of clustering varies from 0 - 100%; and a dendrogram. In the latter, observations which are highly connected group at an early stage and as the threshold for joining a group is lowered, more observations or groups merge to form classes at lower levels of generalization until all group together in a single class. Given this output, it is necessary to choose a specific level (from the variety which can be used) at which clusters can be identified.

The level of similarity chosen as a cutoff point determining the degree of clustering and the number of cluster groups, is dependent on the requirements of the researcher. According to Moll, Campbell & Probyn (1976), there is no recognised objective procedure for selecting

Figure 3.2 Distribution of average SES scores in the study area



Different wage and salary scales would similarly decrease the index for income and consequently educational opportunities would be commensurately lower, again reducing the score. The influence of the 40 - 45 bulge is very limited for the data set is statistically normally distributed and all the values lie within three standard deviations of the mean. The coefficient of skewness is 0,34 and that for kurtosis is 2,86[1], as calculated using the Statjob unistat 2 computer programme (M.A.C.C., 1976).

Table 3.1 Statistical characteristics of the study area

Suburb	Range	Mean		Modal values	Median	
North	26	51	6,4	49; 50; 52;	52	- skew
Central	36	51	7,3	51; 53	52	- skew
East	26	53	7,8	44	46	+ skew
South	35	49	8,1	47; 50; 52	48,5	+ skew

1. For skewness a value of zero (0) and for kurtosis a value of 3 indicate perfect normality (Gregory, 1973, p.59).

Although mean values may hide differences in detailed characteristics between individual ESD's they do give some idea of the likelihood of trends or patterns in the distribution of social status over each of the four major residential concentrations. The mean calculated for each residential concentration shows that there is little obvious difference between the areas, a conclusion supported by the fact that the standard deviation has a range of only 1,7. When, however, the relationship between mode, mean and median values is considered, slight differences between the areas emerge. The eastern and southern suburbs show a positive skewness with a larger proportion of scores below the mean, while the northern and central suburbs are negatively skew. This is most marked in the northern area where there are two modal values higher than the mean. This suggests that higher status than average is found in two of the concentrations, viz. north and central, while the eastern and southern areas have below average socio-economic status.

In order to get a clearer picture of the distribution of mean scores, the surface of average SES scores for each of the residential concentrations (Figures 3.3. - 3.6) has been drawn with the aid of the Saclant Graphics Computer Package (U.C.T. Computing Service, 1979). To ensure that the contour diagrams are directly comparable, the same scale is used for each area. The size of each residential concentration is thus proportionally correct. Actual means were plotted at the point in the centre of each ESD and contours interpolated between given values.

Northern suburbs (Figure 3.3)

There is very little variation in mean SES scores in this area. Although the range of values is only 26 (Table 3.1), most scores vary between 50,5 and 60,5. Mean and median almost coincide. Variation in the mean scores is greatest from west to east (Figure 3.3(b)). The general trend is one of a ridge in the south west (Figure 3.2b(i)) a sudden sharp decrease towards the low of 34 in the centre and a steep increase towards the north eastern plateau before falling again to the north east corner low (39). The flat character of the plateau of higher scores (60,5) is illustrated in Figure 3.2b (iii) and reflects the location of a single ESD which is proportionally the same as the other smaller ESD's.

Central suburbs (Figure 3.4)

This area also shows slight negative skewness and though the overall mean is the same as that for the northern suburbs, the range is larger (Table 5.1) and therefore greater contrasts can be expected. In fact the profiles (Figure 3.4b (i & ii)) give an impression of generally higher scores. Except for a small outlier to the east of the indian residential area, a gradient from high mean values in the west (near white residential areas) to low ones in the east (where there is more coloured housing) is clearly evident (Figure 3.4b i). The north south pattern is one of gentle undulations but shows a steep rise to the plateau of high values (70,5) in the south.

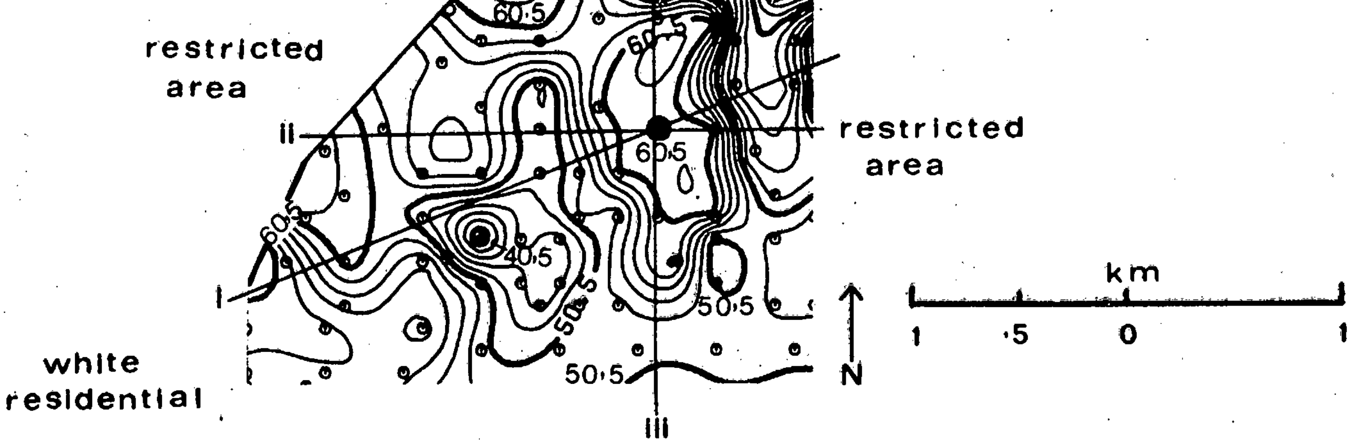
Eastern suburbs (Figure 3.5)

As in the case of the northern suburbs there is not a great deal of variation in mean values in this area (range 26, Table 3.1). The eastern suburbs have the highest mean, despite having the lowest modal value of any of the residential concentrations. This high mean is a reflection of a few high values within a majority of low scores corresponding to areas of scores of 40 and below, accounting for the slight positive skewness. The profiles (Figure 3.5 b) are generally flat with small undulations but the trend is one of decrease from west to east (Figure 3.5 b ii). A plateau of mean values in the 50's to 60's runs almost diagonally from north west to south east, broken in the central area by an intrusion of lower values (40's) from the east. The entire panhandle area has extremely low scores (below 30) and are contained within a single contour.

Southern suburbs (Figure 3.6)

The relatively large range (Table 3.1) and fairly low mean give rise to a positively skewed distribution in this area. This suggests that the scores will be unevenly distributed and generally below the average for the other residential concentrations. Profiles (Figure 3.6 b) show higher values along the northern and eastern

a) Contours



b) Profiles

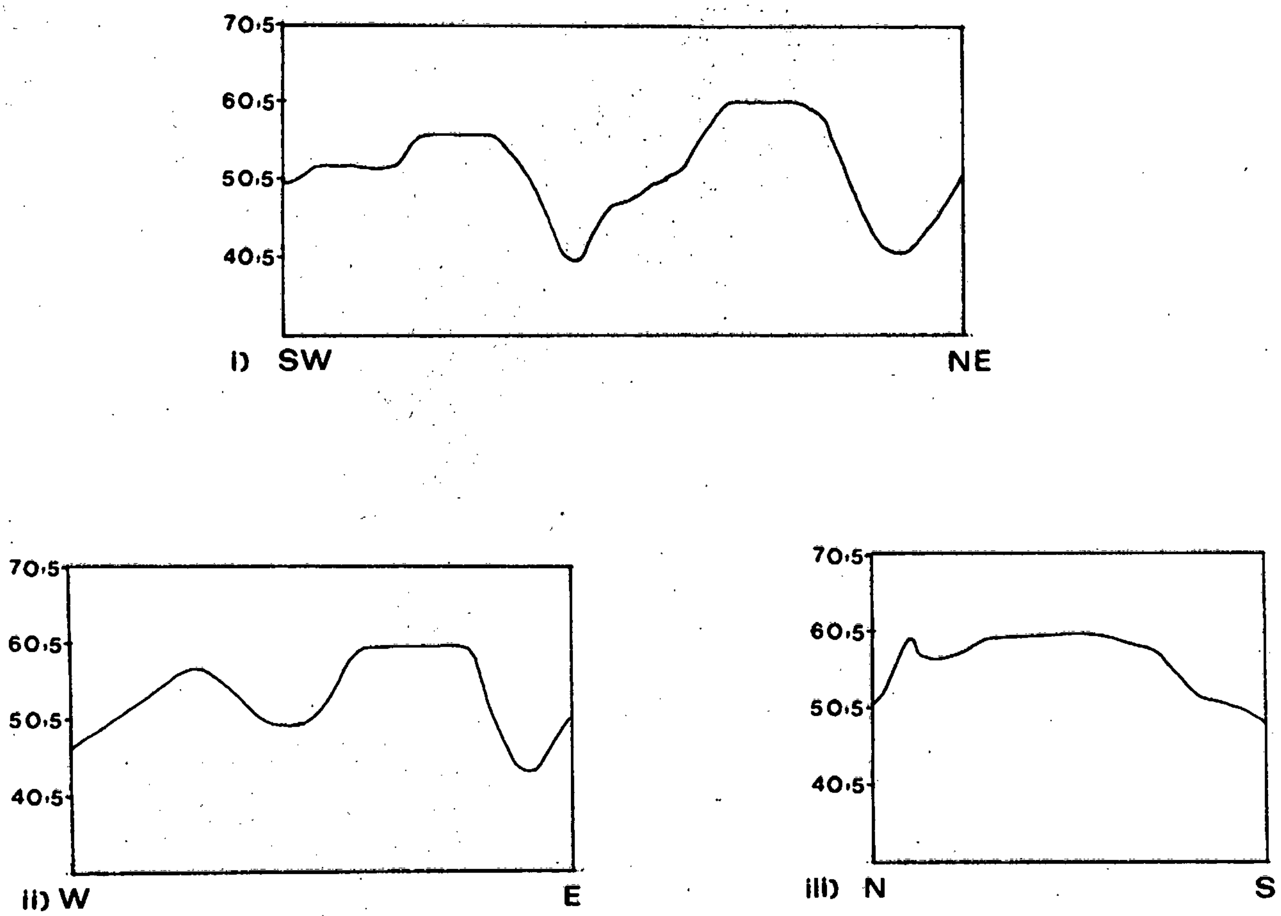
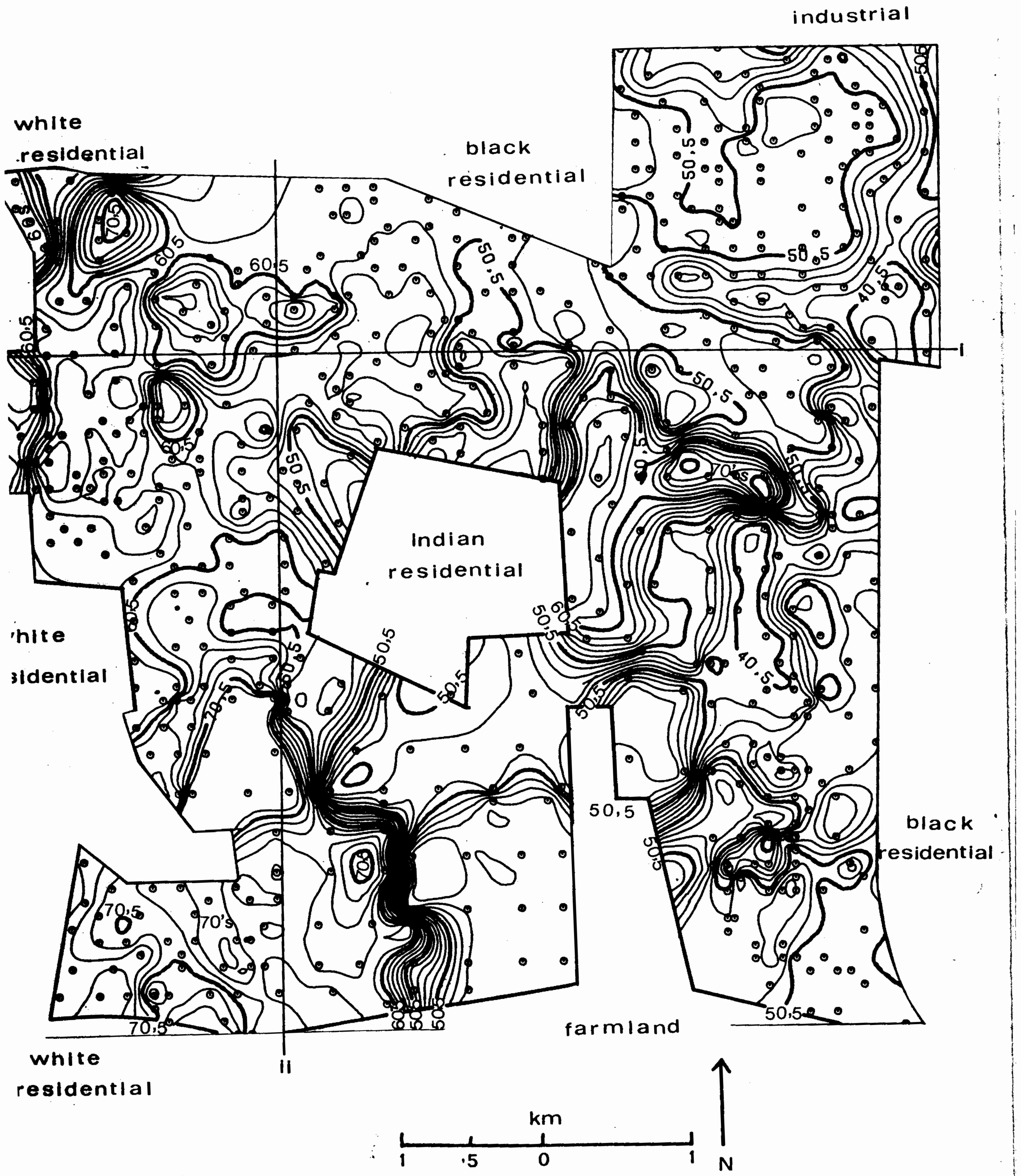
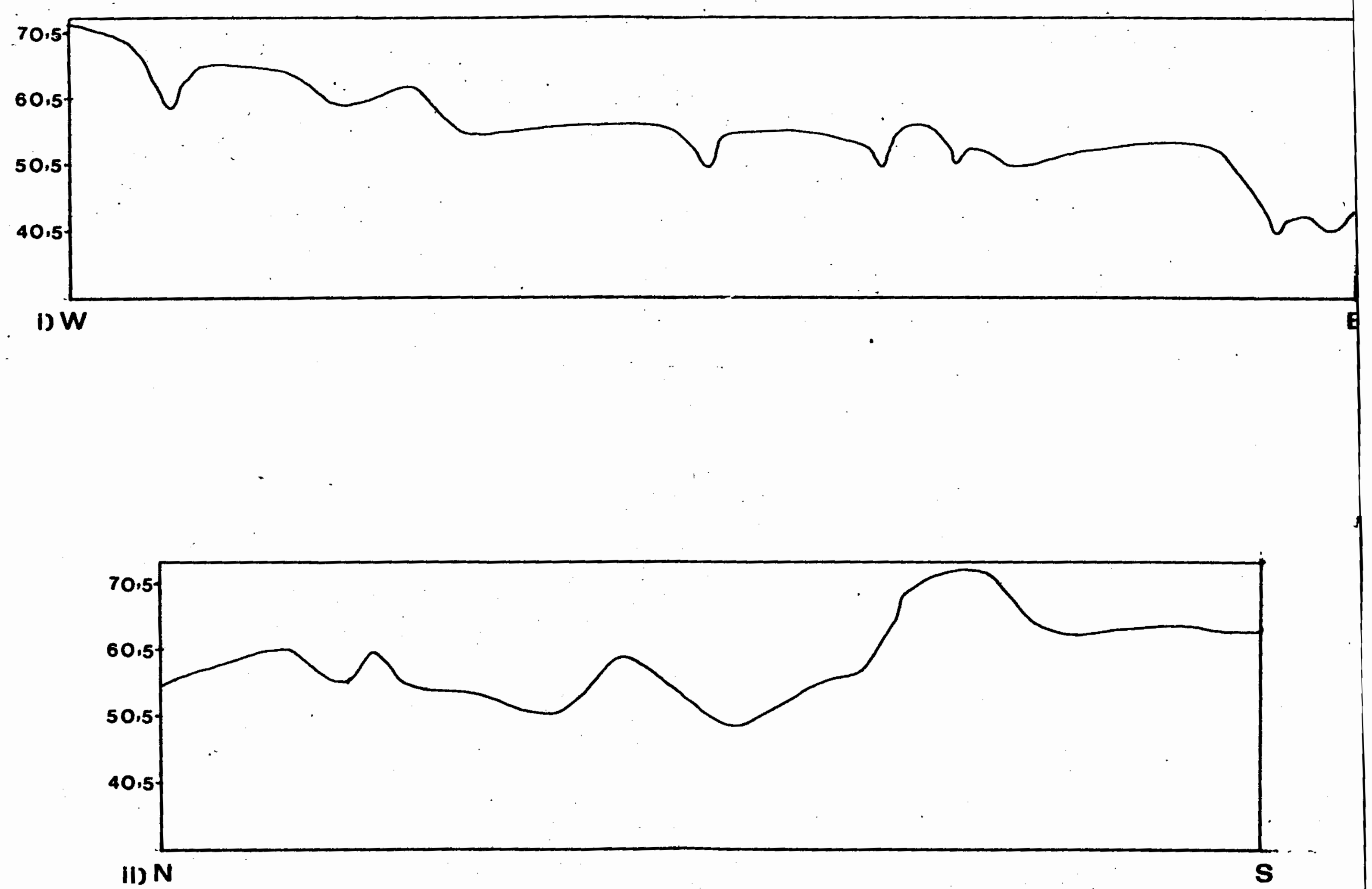


Figure 3.3 Mean SES scores: northern suburbs

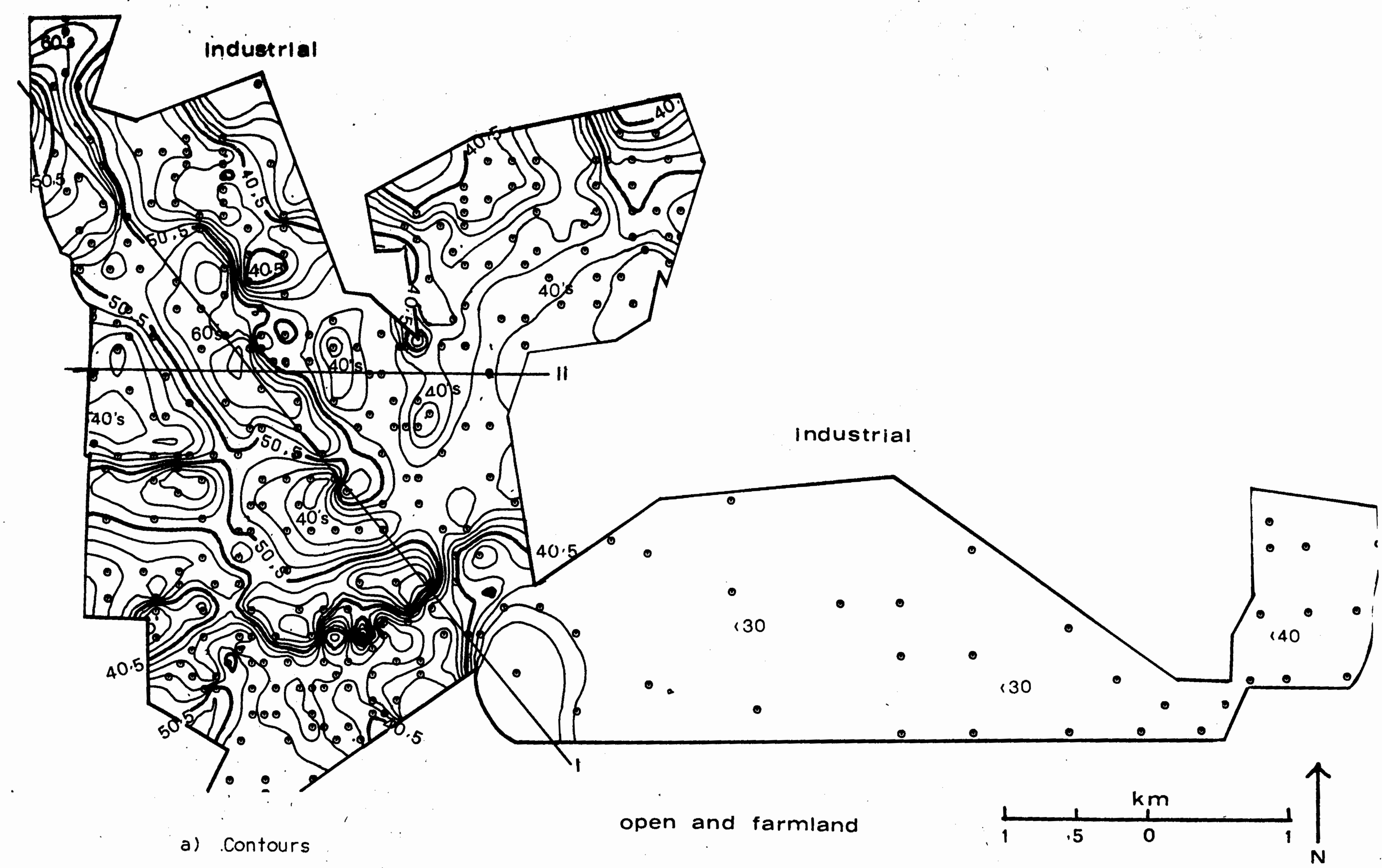


a) Contours



b) Profiles

Figure 3.4 Mean SES scores: central suburbs



a) Contours

b) Profiles

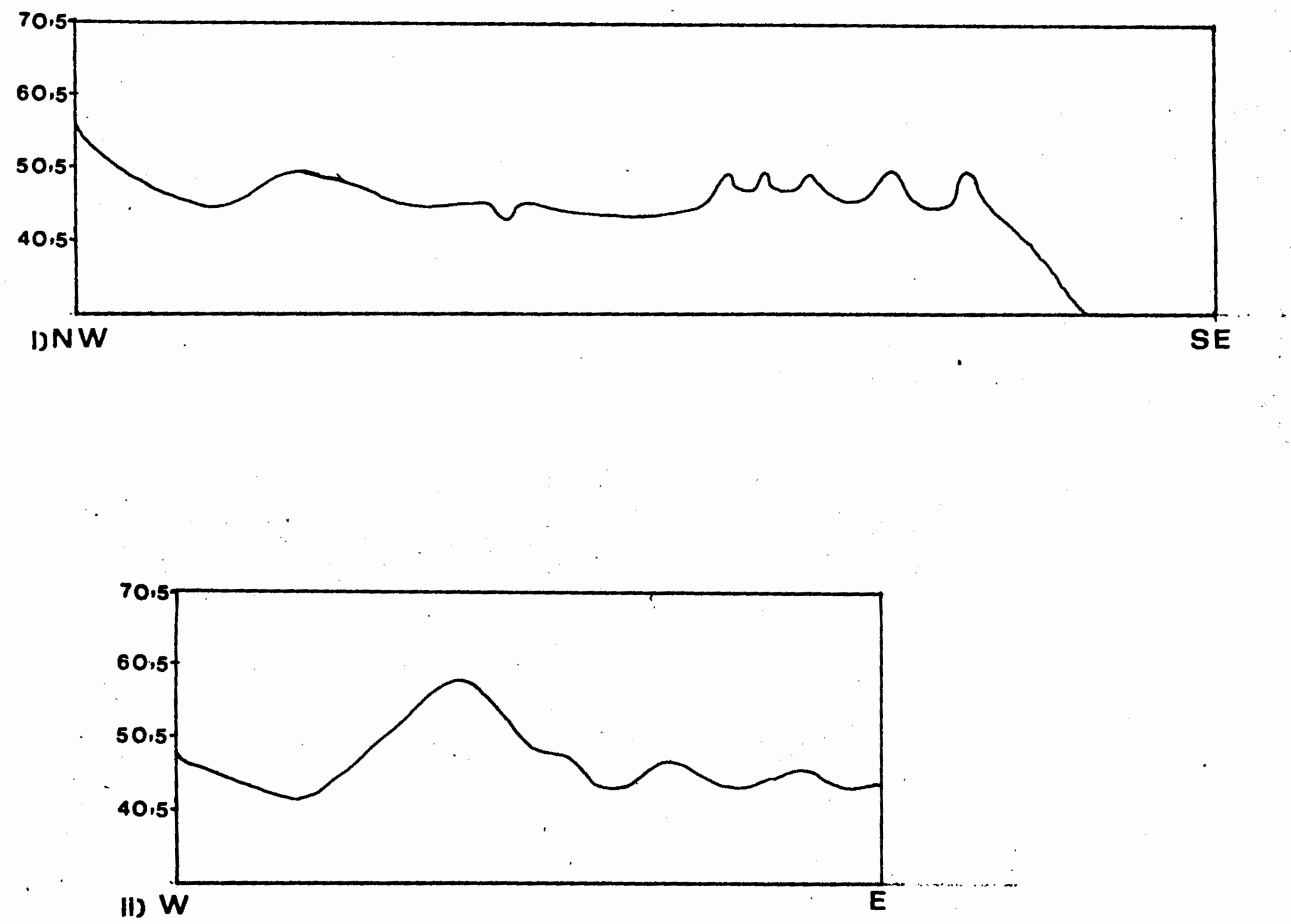


Figure 3.5 Mean SES scores: eastern suburbs

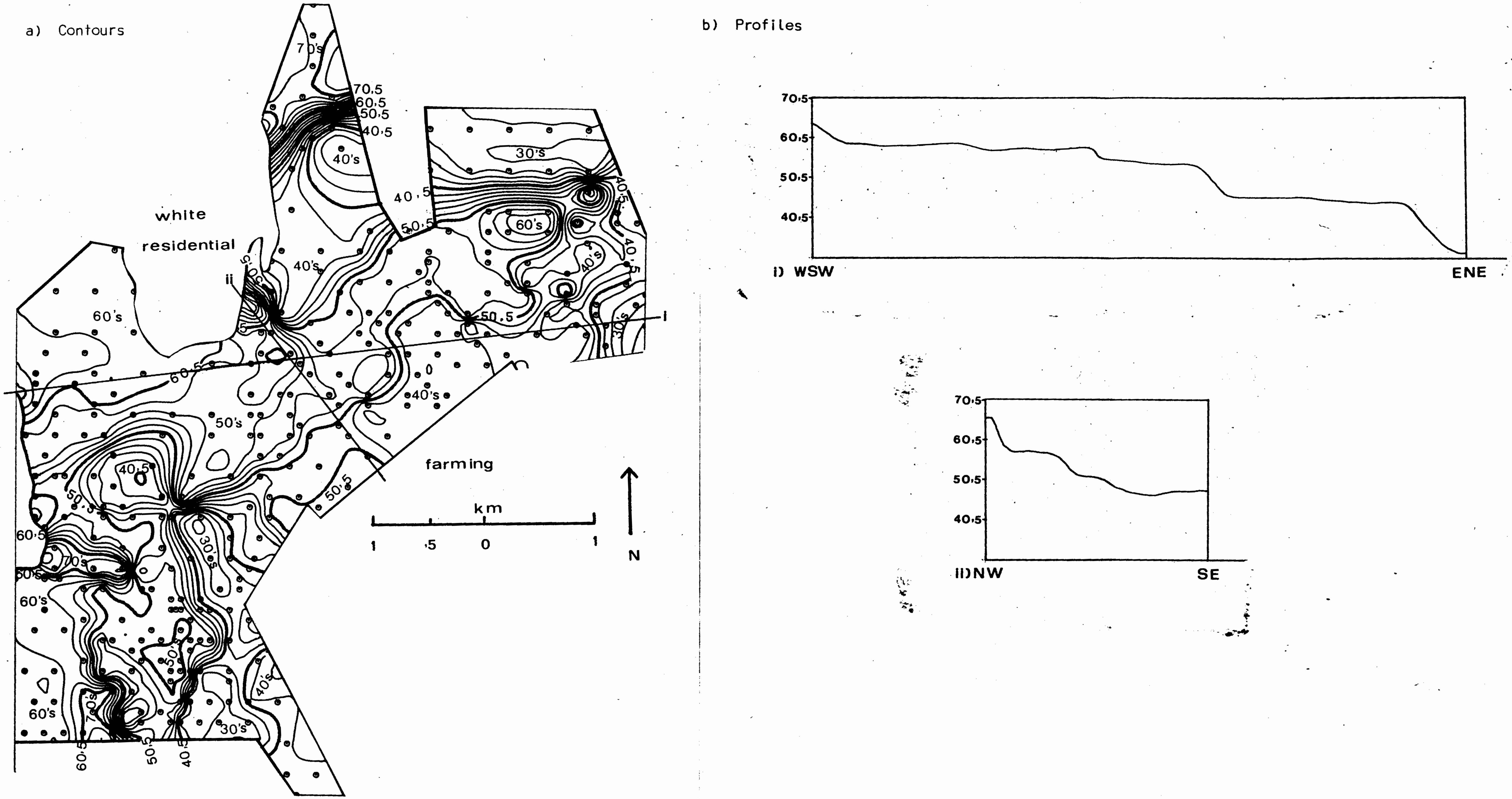


Figure 3.6 Mean SES scores: southern suburbs

edges, dropping sharply to lower values and then flattening out towards the southern and western boundaries. This gradient from the northern and western edges to the southern and eastern edges may again, as in the case of the central suburbs be related to the proximity of the convex edge to white residential areas and places of employment. The concave edge borders on an area of smallholdings and farmland.

In two of the four areas (central and Southern) gradients from higher to lower mean SES scores can be identified. The northern suburbs shows two ridges of high values separated by a trough of lower scores. Overall, mean score plots show complex patterns of local level undulations. This suggests that broad patterns may be identified but owing to the relatively small range found in mean scores, they need to be seen in conjunction with other measures of social status.

ii) Modal class patterns

Modal class values indicate the percentile into which the majority of income earners in an ESD fall. Assuming that the dominant social status group is most likely to influence the status of each ESD as a whole, use of modal classes should allow for a more realistic assessment of socio-economic status than average values. If anything, use of the modal class would tend to underestimate the actual or average social status of any ESD where more than one individual contributes to the status of the household. Modal class may also reflect a link with housing as allocation of housing takes place on the basis of income of head of household. Where this is the case a high proportion of people of a certain income are likely to fall into the same socio-economic class and in the profile would then emerge as the modal class.

Within the study area, modal classes show a fairly wide distribution range as they extend over seven of the ten possible categories. In 59% of the ESD's one modal class emerged as distinctly dominant over other classes (Figure 3.7 a) and the class value could therefore be regarded as representative of the socia status of the majority of the wage earners. In only 11% (53) of the ESD's the profile was bimodal (Figure 3.7 b) and in every case the two modes were adjacent to each other. If both values are taken into account these ESD's would be over represented. Arbitrary selection of one or other is unsatisfactory. The boundary value between these adjacent bimodal classes was therefore chosen as most representative and would also tend to under- rather than over-estimate actual social status within the ESD.

In 30% of cases, however, the modal class does not stand out as dominant over the other classes, which may have almost the same number of occurrences (Figure 3.7 c). It was felt that although the modal class was not obviously representative of the ESD it did reflect the highest frequency of occurrence and as it tended to be centrally located in the distribution was a satisfactory measure.

Midpoint values for each modal class were plotted for each ESD in the four residential concentrations. Again using the Saclant Graphics Package (U.C.T. Computing Service, 1979) contour diagrams (Figures 3.8 - 3.11) were produced. Modal class was also mapped by ESD using a number code for successive modal class. Besides giving a simpler visual image than the contour diagrams, it was felt that this would best illustrate the degree of contiguity of ESD's of the same modal class. Perspective diagrams (Figure 3.12) were also constructed by the Saclant Graphics Package (U.C.T. Computing Service, 1979). These diagrams show the nature of the "topography" of each area and obviate the need for profiles. The diagrams are viewed from the south east corner of each residential concentration and the angle of elevation is 80° from the horizontal, so that lower features in the background are not obscured by peaks or ridges in front of them.

The distribution of modal class midpoints gives distinct patterns of

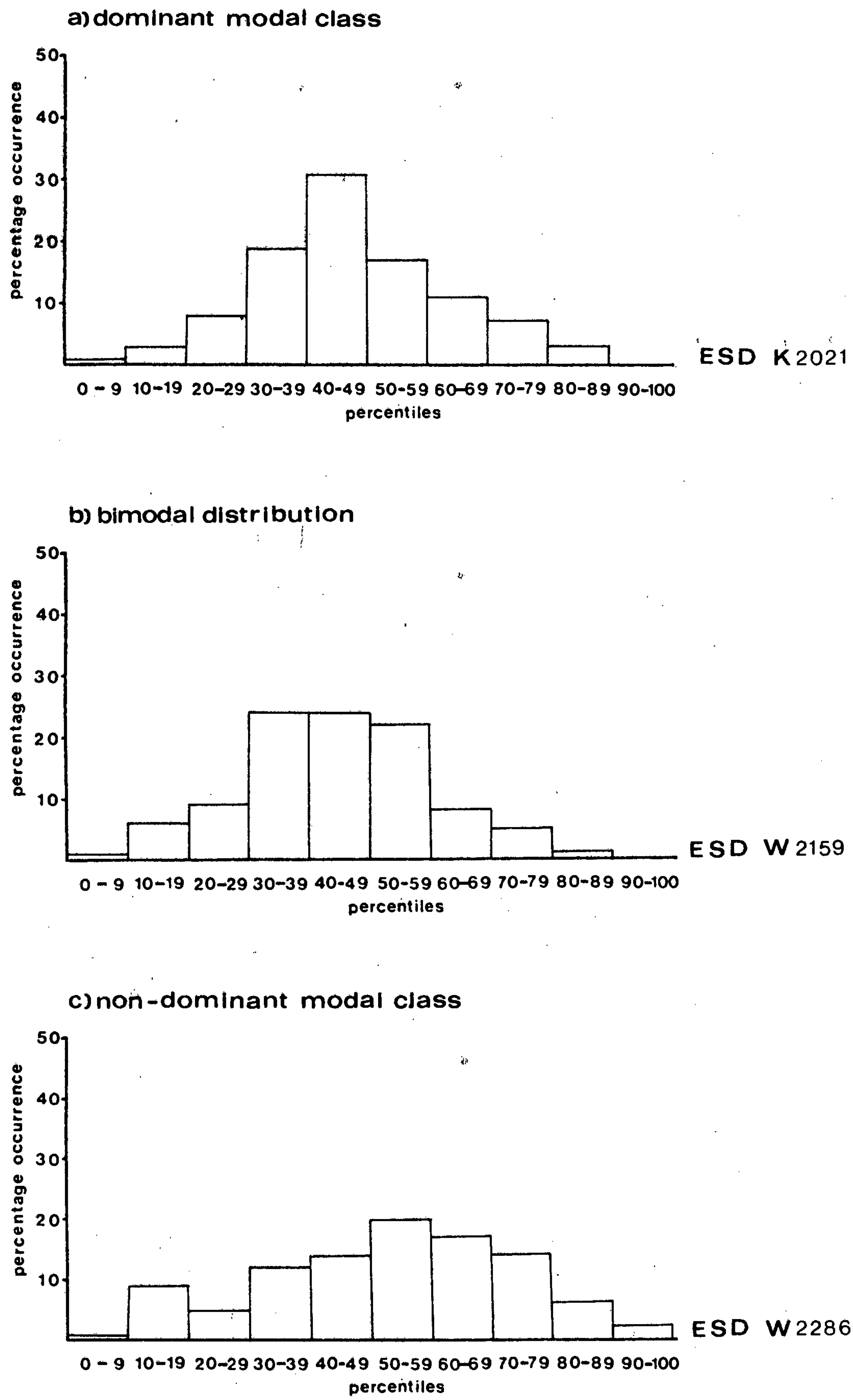


Figure 3.7 Representative ESD profiles

differing socio-economic status and reinforces the general trends shown by average SES scores.

Northern suburbs (Figure 3.8)

The modal classes 40 - 49 and 50 - 59 contain the majority of ESD's in the area (45% and 30% respectively) (see Table 3.2). The north south trending ridge of high modal class (Figure 3.8 b) is actually a single large ESD (Figure 3.8 a) which is flanked by a sharp drop of two classes to the east and to the west a more gentle gradient down to the single very low cell. The population in this ESD is similar to that of the other ESD's in the northern suburbs and density is thus lower which may be related to the higher social status as reflected by a modal class of 70 - 79. The south west corner of the area, surrounded on three sides by white residential areas also has high modal class values. These two ridges of high value separated by a trough of lower values are outstanding features on the perspective diagram (Figure 3.12). At the centre of this trough is an ESD of modal class 30 - 39. This cell has a population of 604 while the mean for the area is 435. Density in this ESD is thus extremely high and the area has been identified as a squatter settlement (Simkins, 1978). The remainder of the ESD's are of approximately the same size and form two groups of four and two groups of seven contiguous ESD's (Table 3.2).

Central suburbs (Figure 3.9)

The most obvious feature is the large plateau of high scoring ESD's in the south east and a smaller one northeast of the indian residential area. These are formed by 22 and six respectively of the 36 ESD's of modal class greater than 70 and are well illustrated in Figure 3.12. A further five are one ESD from the main concentrations. In only two cases are contiguous ESD's in the percentile immediately higher or lower. The break is far sharper for the rest - a drop of at least one and in some cases two percentiles. The eastern sector is generally low scoring, with small isolated pockets of lower and higher modal values. In class 40 - 49 the groups contain an average of 35 cells and 93% and 66% of the ESD's in this sector respectively form groups of more than three contiguous ESD's (see Table 3.2). Overall the topography is relatively smooth.

Eastern suburbs (Figure 3.10)

The modal class 40 - 49 dominates the eastern suburbs, accounting for 55% of the total number of ESD's in this area. Of these, 95% consist of more than three contiguous ESD's all located in the main section (Table 3.2). Classes 50 - 59 and 30 - 39, although forming smaller proportions of the total (24% and 17% respectively) are grouped together and 83% and 62% respectively of their members comprise more than three contiguous ESD's. It can also be seen from the map that on average, groups are formed of six contiguous ESD's in these two modal classes. The two percentiles of highest value (60 - 69 and 70 - 79) ESD's are single isolated ESD's (Table 3.2).

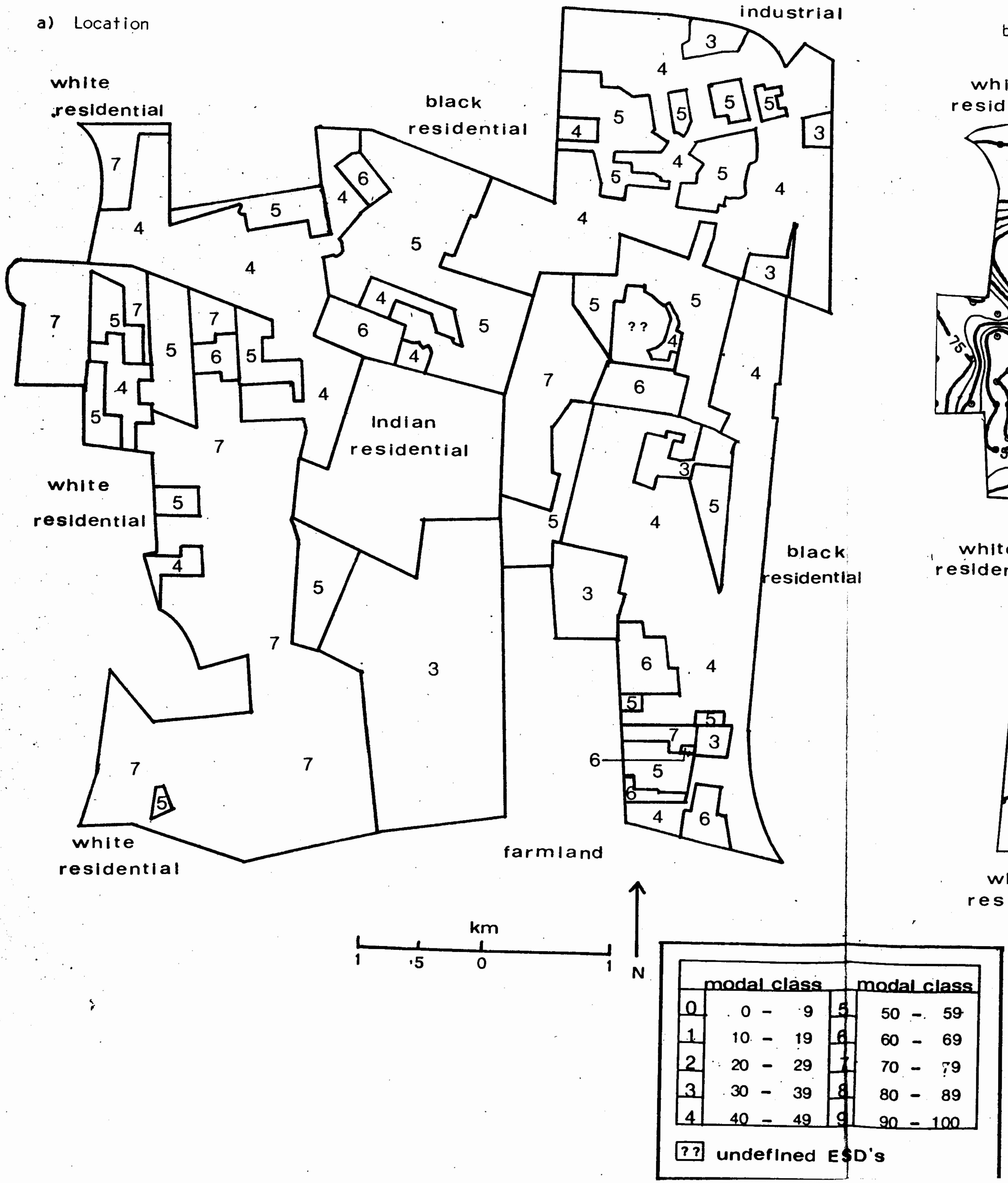
Other than in the case of a single ESD of high modal class (70 - 79) in the centre of the main area which is surrounded by ESD's of two or three modal classes lower, adjacent ESD's are generally only one percentile removed from their immediate neighbours. The generally north west to south east trending ridge which emerged in the plot of mean SES scores although still identifiable (Figure 3.10) is not very high. The panhandle is made up of three large, sparsely populated cells, flanked by industry and open land to the north and farmland to the south. It contains the lowest modal class values for both the eastern suburbs and the whole study area (one ESD with the modal class of 10 - 19). The overall impression, substantiated by the perspective diagram for this area (Figure 3.12) is that of gentle undulations in modal class values.

Southern suburbs (Figure 3.11)

Despite the majority (68%) of the ESD's in the area falling into three modal classes, the modal class contour diagram (Figure 3.11 b) gives a more broken pattern than the gradient of high scores on the north western edge to lower ones on the south eastern side shown by mean SES scores. This is clearly evident in the perspective diagram (Figure 3.12). Islands of high value amongst low scores and vice versa are now evident in the south western and far eastern sections of the diagram. This corresponds to groups (containing on average six contiguous ESD's) of the modal classes 30 - 39; 40 - 49 and 50 -59 in which 54%, 84% and 90% respectively comprise groups of three or more contiguous ESD's (Table 3.2). A fairly well defined plateau of values in the 70 - 79 percentile in the north west and north and one of modal class 40 - 49 in the south east, are the most obvious features. Although only 10% of the total ESD's fall into the class 70 - 79, three of these are separate outliers at least one percentile higher than all the surrounding ESD's. The remainder form a single group (Table 3.2) of five contiguous cells, adjacent to the white residential areas and make the plateau. There is a sharp drop of modal class values (from two to four percentiles) from this plateau to the adjacent ESD's.

In each of the residential concentrations which make up the study area then, the trends identified by the plotting of average SES scores are still recognisable but are intensified by the grouping inherent in plotting modal classes. Visually a distinct difference of more than one percentile occurred between adjacent groups of ESD's which suggested that the small groups of contiguous ESD's in the same modal class should be readily identified and socio-economic status was more similar between these cells than it was to other ESD's although they are adjacent.

It is of interest to establish whether groups of ESD's identified from modal class plots could be related to suburbs which are named sub-areas



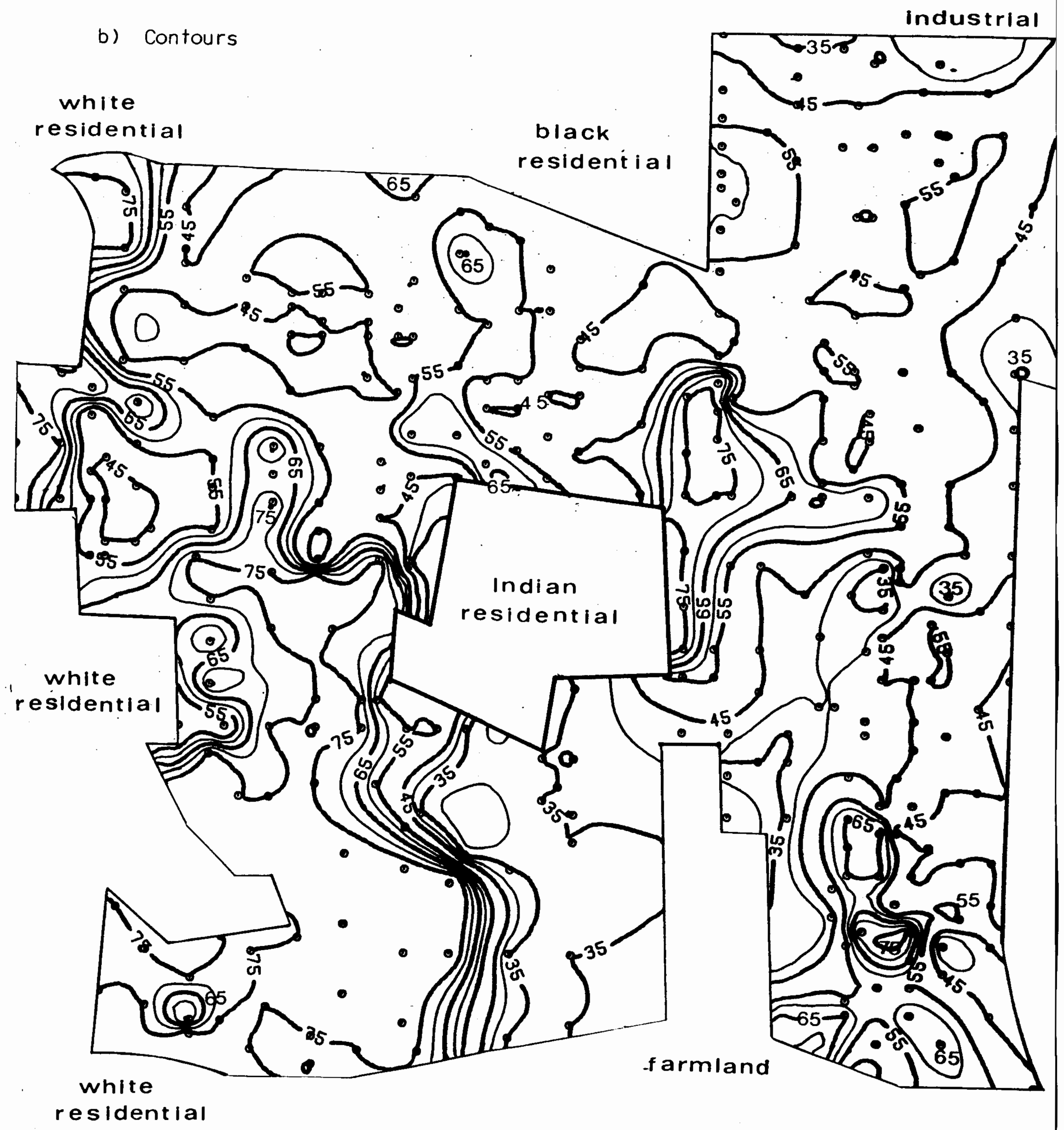
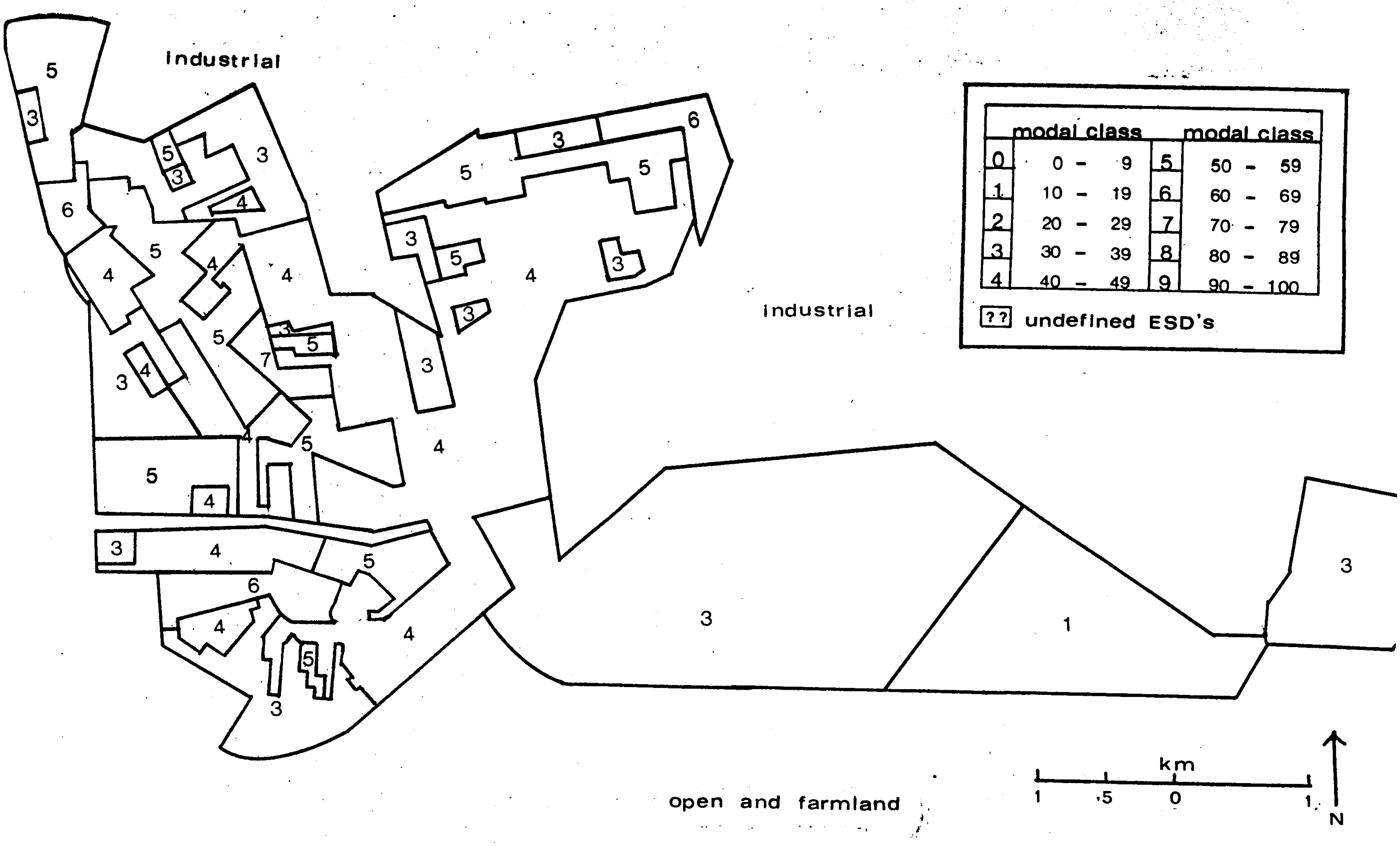


Figure 3.9 Modal class patterns: central suburbs



b) Contours

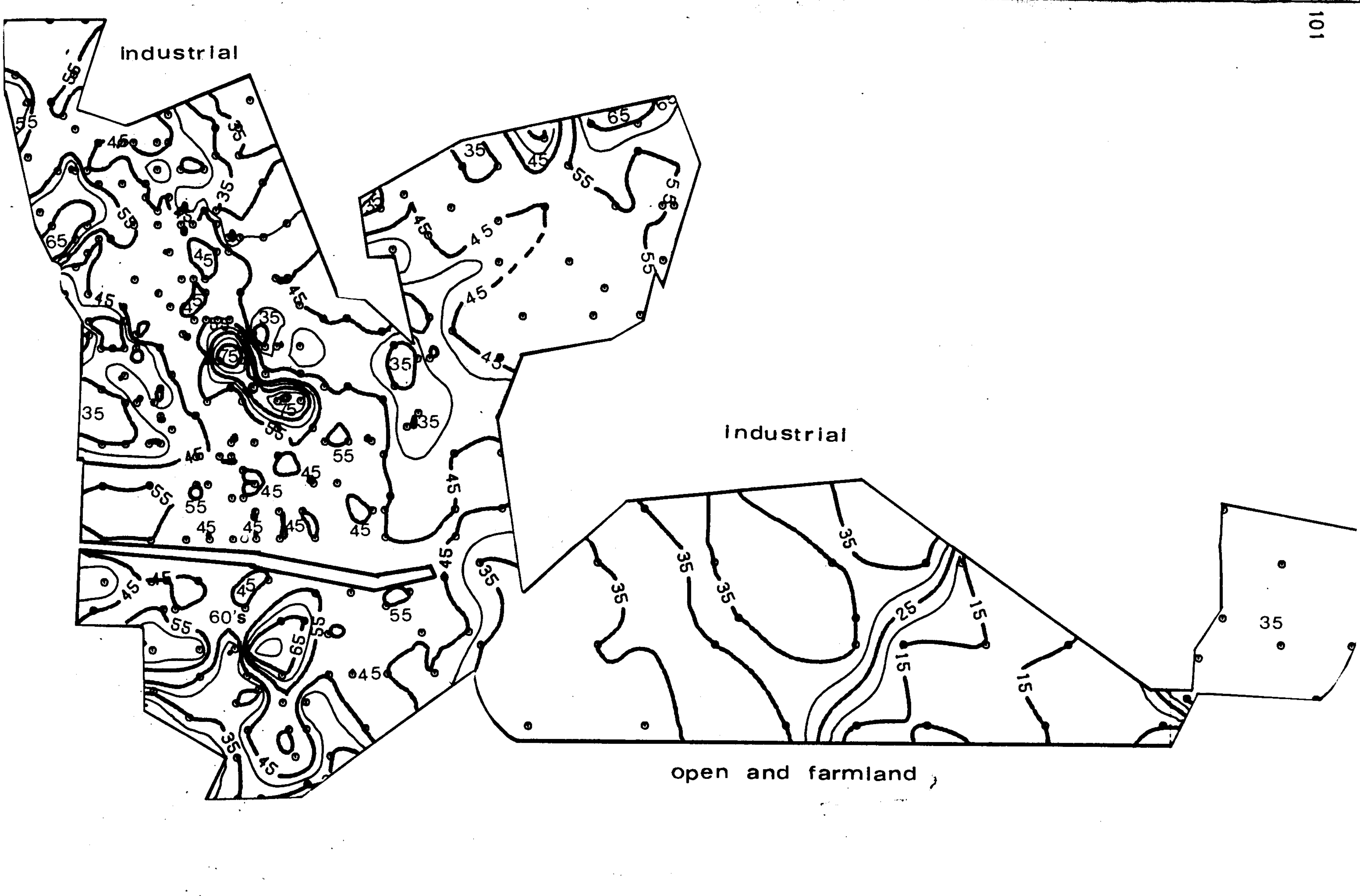
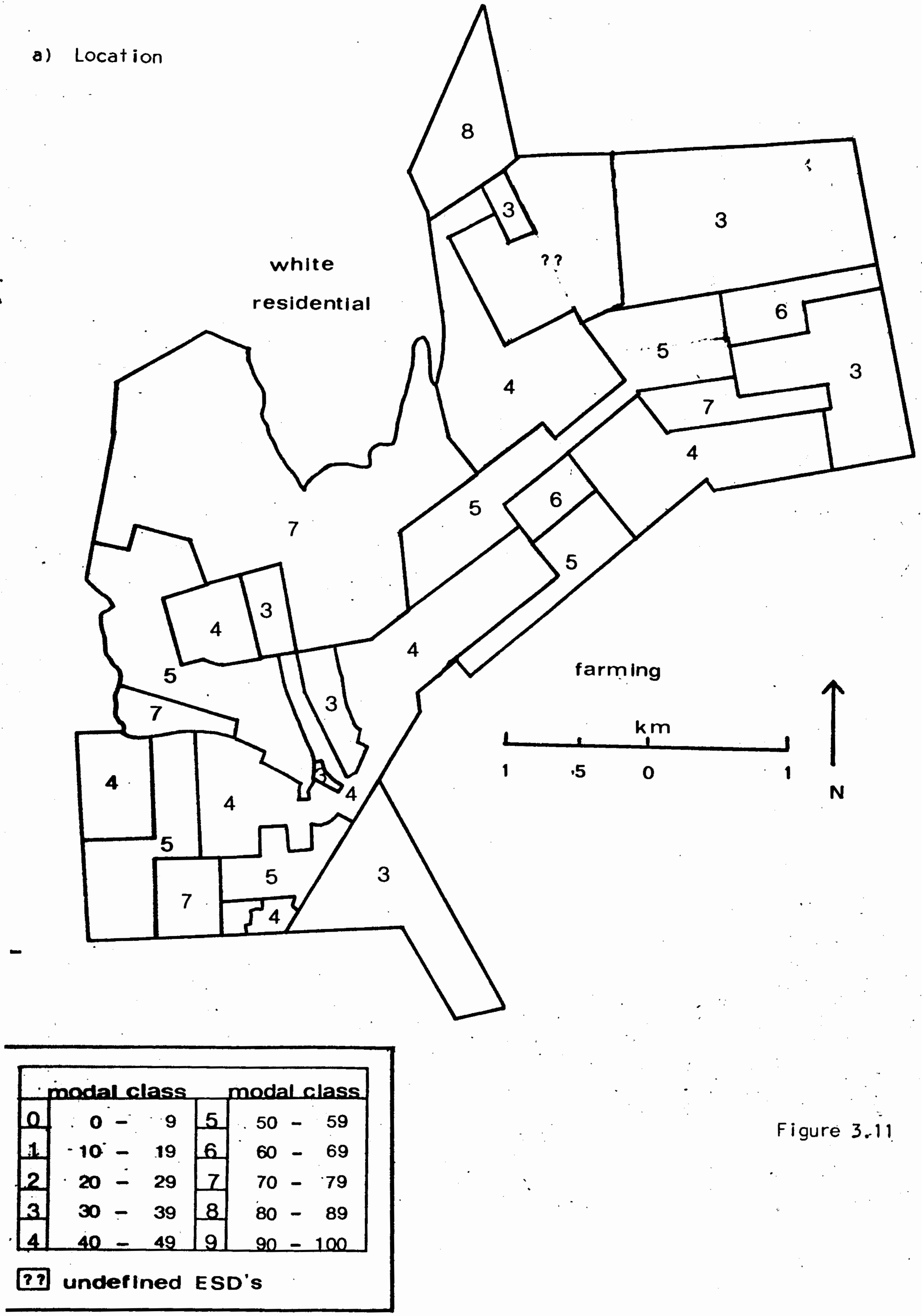


Figure 3.10 Modal class patterns: eastern suburbs



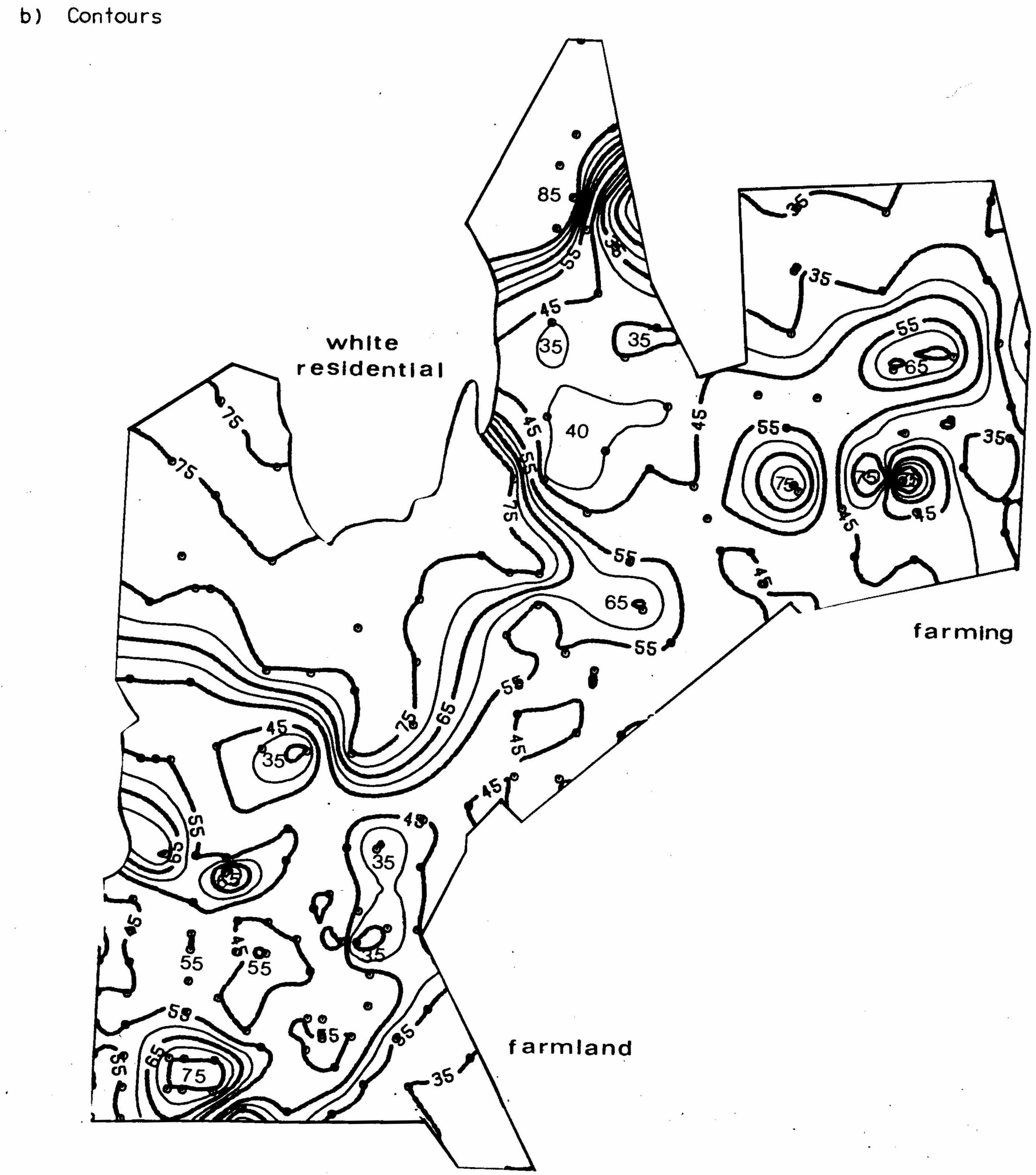


Figure 3.11 Modal class patterns: southern suburbs

Figure 3.12 Perspective diagram of modal class distribution in the study area

TABLE 3.2. Continuity of cells in modal classes (see Olson, 1975).

	Modal class	Groups of 1 ESD	Prop of class	Groups of 2 ESD's	Prop	Groups of 3+ ESD	Prop	No.	Prop. of total ESD's
Northern Suburbs	30 - 39	1	33%	1	67%	0	-	3	10%
	40 - 49	0	-	0	-	3	100%	15	45%
	50 - 59	0	-	0	-	1	100%	10	30%
	60 - 69	1	100%	0	-	0	-	1	3%
	70 - 79	2	50%	1	50%	0	-	4	12%
								33	100%
Central Suburbs[1]	30 - 39	4	50%	2	50%	0	-	8	4%
	40 - 49	6	6%	1	1%	3	93%	106	49%
	50 - 59	11	22%	3	12%	7	66%	50	23%
	60 - 69	7	88%	1	12%	0	-	9	4%
	70 - 79	2	6%	1	6%	3	88%	36	17%
								209	97%
Eastern Suburbs	10 - 19	1	100%	0	-	0	-	1	1%
	30 - 39	10	38%	0	-	4	62%	26	17%
	40 - 49	2	3%	1	2%	5	95%	80	55%
	50 - 59	4	11%	1	6%	5	83%	35	24%
	60 - 69	3	100%	0	-	0	-	3	2%
	70 - 79	1	100%	0	-	0	-	1	1%
								146	100%
Southern Suburbs[2]	30 - 39	5	33%	1	13%	2	54%	15	16%
	40 - 49	2	8%	1	8%	4	84%	26	28%
	50 - 59	0	-	1	10%	4	90%	22	24%
	60 - 69	2	100%	0	-	0	-	2	2%
	70 - 79	3	33%	0	-	1	67%	9	10%
	80 - 89	1	100%	0	-	0	-	1	1%
								75	81%

1. All except 9 ESD's are taken into account as numbers 2214 - 2222 are not defined

2. All except 18 ESD's are taken into account as numbers 2028 - 2032 and 2038 - 2050 are undefined

of the city frequently associated with certain mental images. When the boundaries of suburbs in each of the concentrations are overlaid on the modal class distribution, seven of a total of 28 (i.e. 25%) were found to correspond to a group in the same modal class. Generally these were the smaller suburbs. In the southern residential concentration Heathfield and Southfield (Figure 1) together comprise five ESD's of modal class 70 - 79. This presents a marked contrast to the surrounding area, while the three ESD's that make up Lavendar Hill fall into modal class 30 - 39. Lansdowne, Wetton, Hanover Park and Vanguard in the central concentration all consist of ESD's of modal class 70 - 79, while Kalksteenfontein (which consists of only two ESD's) has a modal class of 40 - 49. Finally, Matroosfontein in the eastern residential concentration forms a block of contiguous ESD's of the same modal class (50 - 59). Suburbs of distinct socio-economic status can thus be identified in the study area on the basis of modal class mapping.

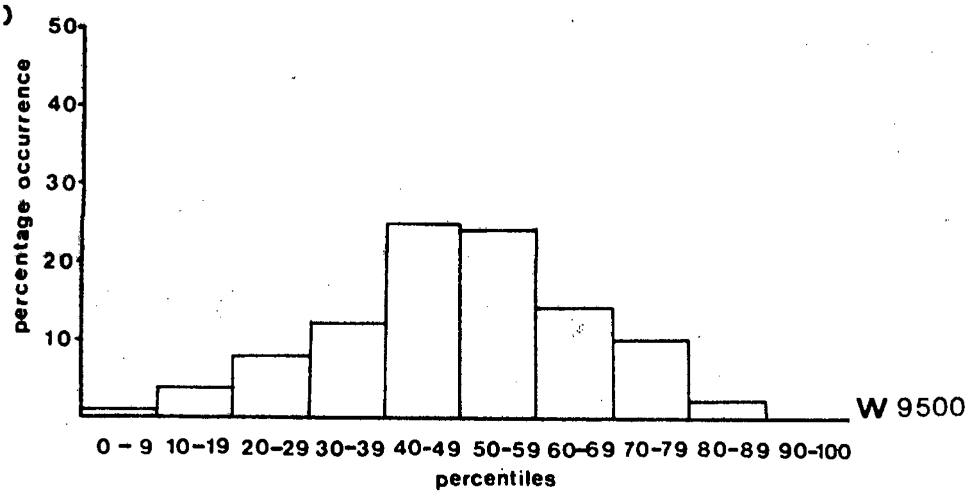
Plotting of descriptive statistics (i.e. mean SES scores and modal classes) has led to the recognition of patterns of differing socio-economic status. The more generalised the statistic, the less specific and distinct are the patterns. From mean SES scores (which are the broadest summary of the nature of the distribution of scores in each ESD) only overall trends of socio-economic differentiation can be identified in each of the residential concentrations. Modal classes, in representing the class of scores common to the majority of wage earners in each ESD, are slightly more specific measures and give rise to more distinct patterns of socio-economic differentiation within the coloured areas. Since the likelihood of finding patterns of socio-economic differentiation among coloureds has thus been established, it is important to use the total information on the socio-economic character of each ESD (i.e. full profile) to ascertain the nature of socio-economic grouping and differences among the coloureds of Greater Cape Town.

b) Patterns emerging from classification

The use of modal class to identify spatial patterns of social differentiation involves the selection and use of only one class (albeit the class containing the majority of the population for one ESD) of the ten which make up the profile of social status for each ESD. While the modal classes of two ESD's may be the same, their overall profiles may be very different; e.g. in Figure 3.13 the modal classes each contain 25% of the wage earners, yet Figure 3.13(a) has 50% of the income recipients which fall above the mean while the other (Figure 3.13(b)) has 50% of the earning population represented by scores below the mean. The actual status of the ESD's given by the full profile may thus be lower or higher than that suggested by the modal class. The complete range of values in the profile should therefore provide a more realistic picture of the socio-economic character of a particular ESD.

Cluster analysis was carried out first using the full profiles for each ESD separately. The data base for this study is large, viz. 490 ESD's. The available corespace in the Univac computer at U.C.T., the programme used to perform the cluster analysis and the size limitations determined by subroutines for the plotter prohibited the manipulation of data for the entire study area at one time. Even separating the study area into the four individual residential concentrations meant that the number of cells in the central and eastern suburbs was still too large for the machine. The programme was thus altered so that more than 65K of core space in the computer (limit for average length and size jobs) could be utilized and plotting subroutines were rewritten to allow the dendrograms to be drawn. From the dendrogram the nature of the similarity of groups may be established according to the criterion that members of a group are more similar to one another than to members of another group. Prior to grouping each cell is completely unique. The criteria for joining a cluster however, become less stringent as grouping proceeds and when all the cells have been connected, there is complete generality (Abler *et al.*, 1972).

a)



b)

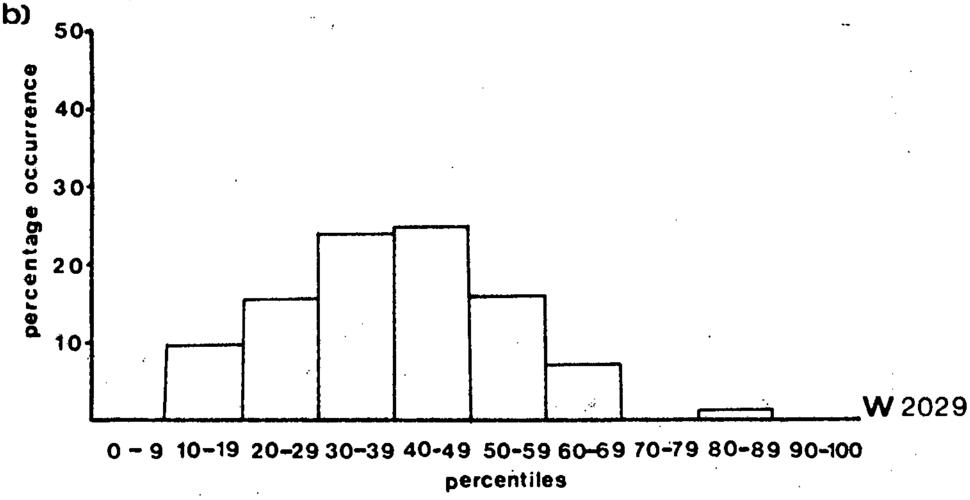


Figure 3.13 Profiles of ESD's in the same modal class

	W9500	W2029
% 30 - 39	25	50
% 40 - 49	25	25
% 50 - 59	50	25

Where the cells join close to one another and a chaining effect exists, similarity within the data set is indicated. A distinctly stepped pattern however, shows that while there are cells closely related to one another, these as a group differ from other individuals or groups of cells. The longer the line representing a cell before it joins with another, the more marked is the difference between it and other cells. An examination of the dendrogram reveals the level at which clearly defined groups emerge and similarities and differences between groups may then be assessed. The cutoff point chosen by the researcher and drawn on the diagram determines the number of groups.

i) Patterns within each residential area

The dendrograms produced for each residential concentration on the basis of full profiles for each ESD are extremely flat (Figures 3.14 - 3.17). Clustering begins at approximately the 95% level of similarity and is complete for the northern and eastern concentrations by the 40% level (Figures 3.14 & 3.16); by 30% for the southern suburbs (Figure 3.17) but only by 5% for the central concentration Figure 3.15). Between group differences for the northern, eastern and southern concentrations is relatively small while greater diversity can be expected in the central area.

Well defined clusters emerge at the 80% level of similarity in each of the residential concentrations. At this level of the clustering process, criteria for individuals joining a group have not become too generalized. Use of the same cut-off level in each case provided uniformity over the whole study area although it resulted in different numbers of groups being defined for each of the residential concentrations. (These will be referred to as "clusters" or "80% level clusters".) These clusters and the values defined for them by the information statistic formed the basis of the patterns to be considered.

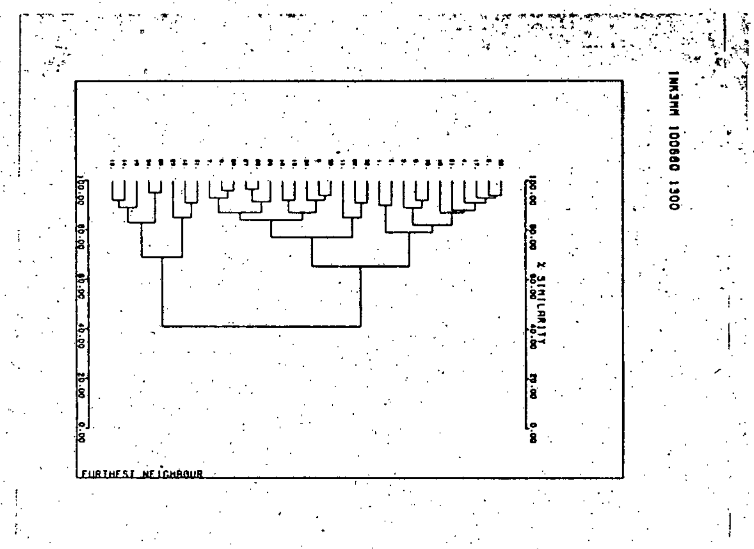


Figure 3.14 Dendrogram: northern suburbs

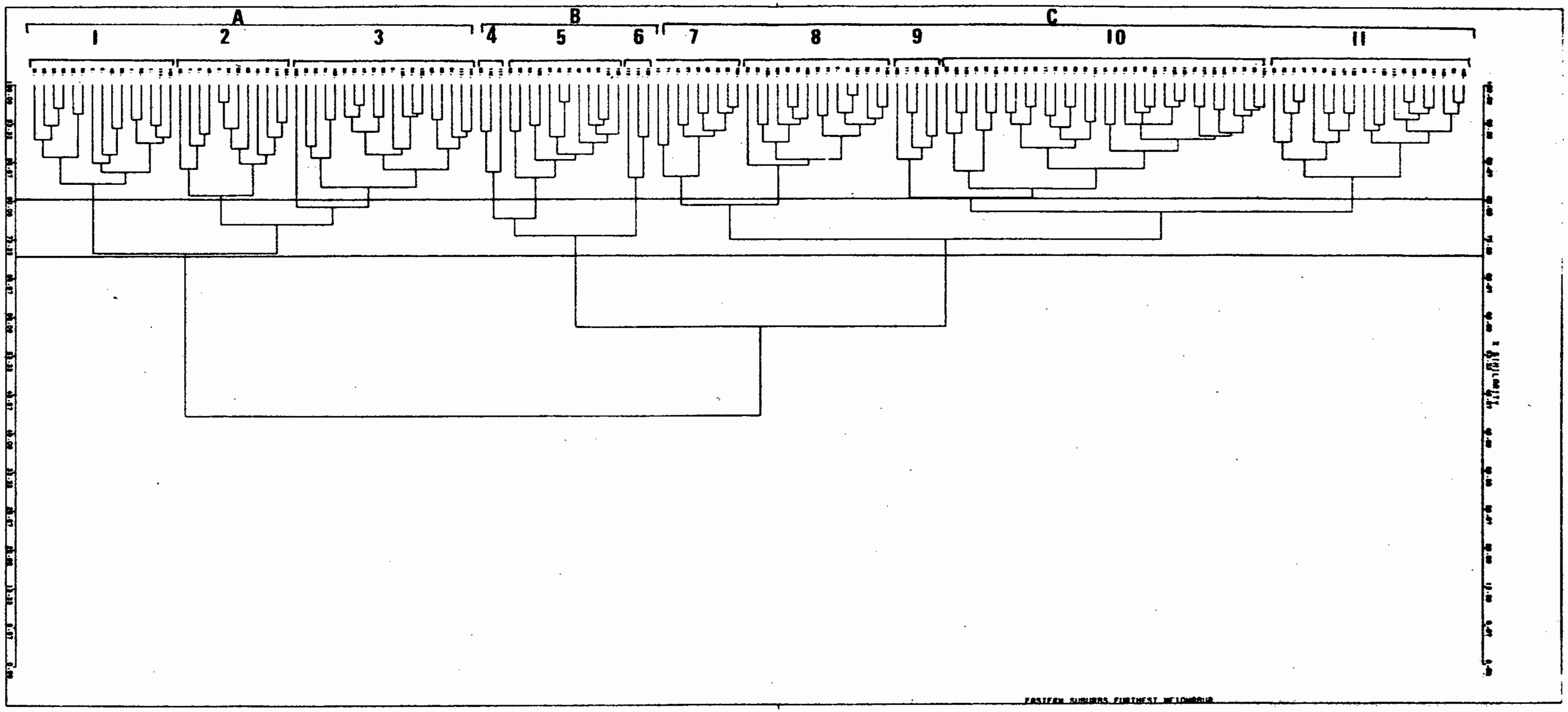


Figure 3.16 Dendrogram: eastern suburbs

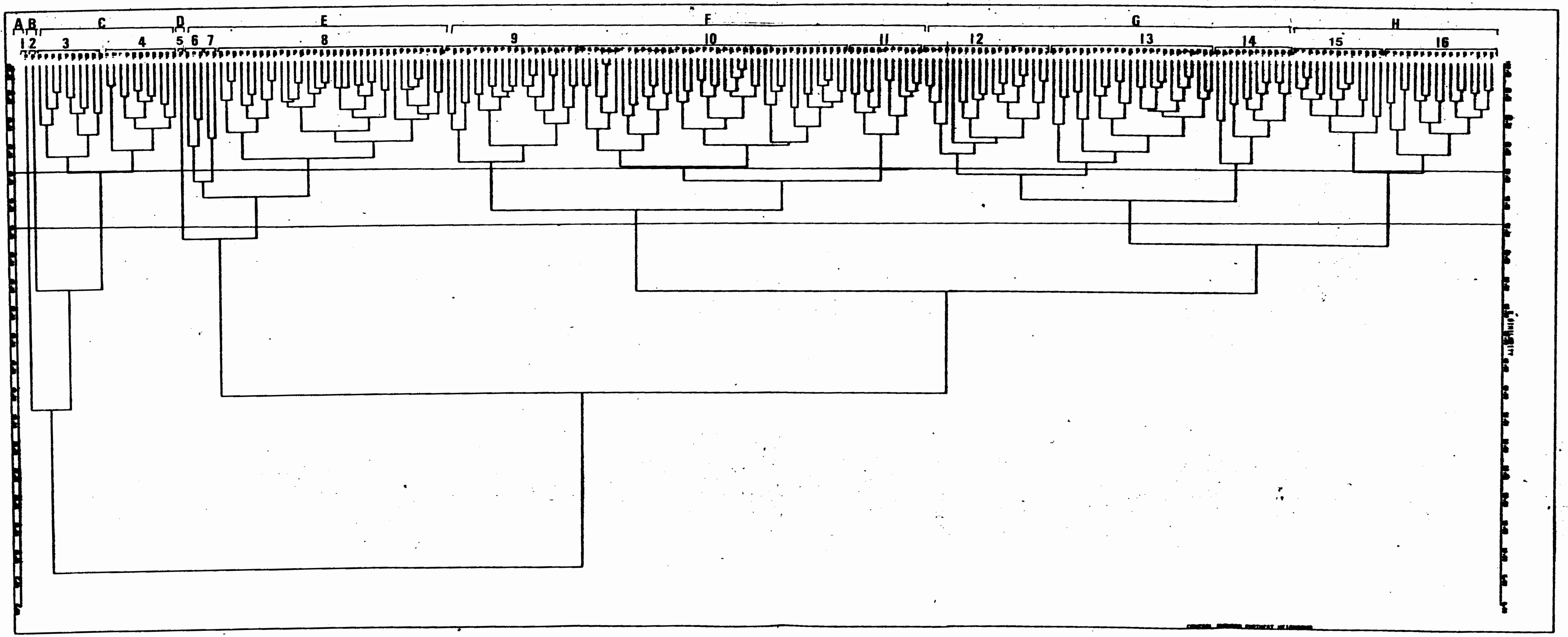


Figure 3.15 Dendrogram: central suburbs

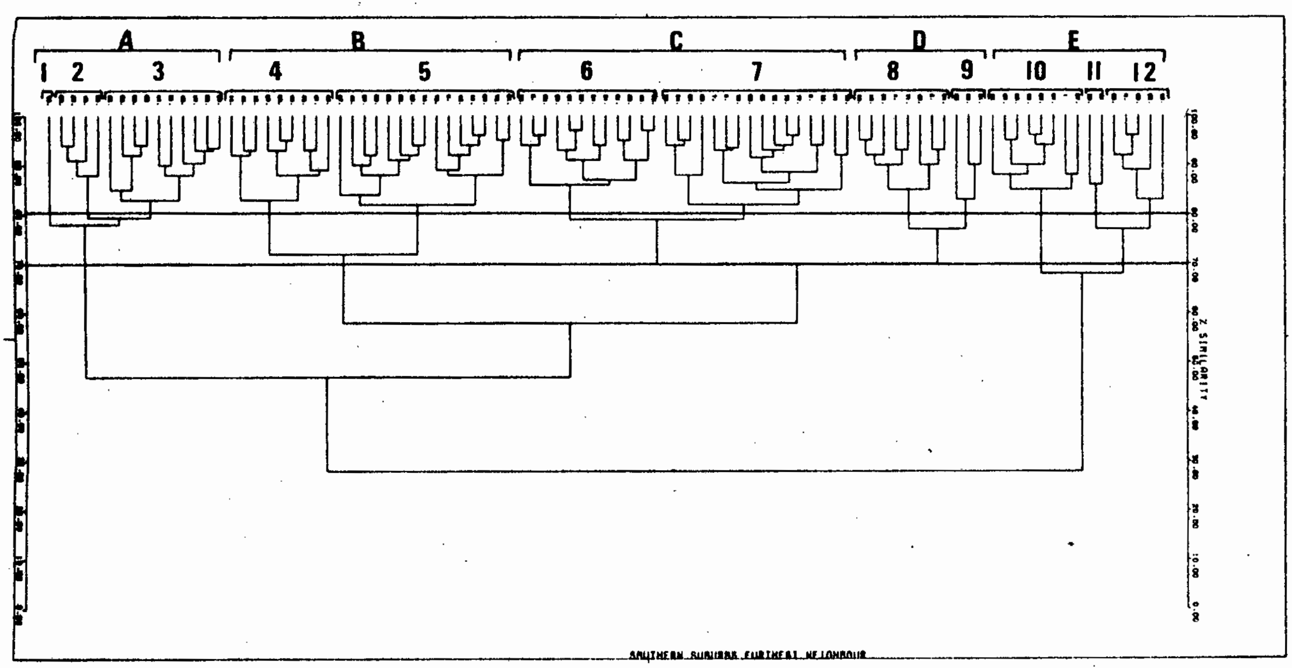


Figure 3.17 Dendrogram: southern suburbs

TABLE 3.3

Area	No. of clusters at 80% level	70% level	Clusters combined	Outlines Single ESD	ESD's brought in adjacent? on map
Northern Suburbs	6	4	5 & 6	-	yes
			4 & 3		
Central Suburbs	16	8	15 & 16	-	yes
			12,13 & 14	-	yes
			9,10 & 11	-	yes
			6,7 & 8	-	yes
			3 & 4	-	yes
				1,2,5	
Eastern Suburbs	11	3	8,9,10 & 11	-	yes
			5,6 & 7	-	yes
			1,2 & 3 & 4	-	yes
Southern Suburbs	12	5	11 & 12		
			6,7,8 & 9	-	yes
			5 & 4	-	yes
			3,2 & 1	-	yes

The number of individuals in each cluster is, however small generally. Although a total of 45 clusters is defined overall, only three contain more than 20% of the total number of cells in the particular area. The information statistic was run for each of the clusters defined at this level of similarity to determine the percentile/s which characterise each cluster and the significance level of the clustering.

Although the level of generalization is still acceptable at the 70% level of similarity, far fewer clusters (20) are defined (Table 3.3). The decrease in the number of clusters at the 70% level of similarity is most marked in the southern and eastern suburbs, where the number of groups was reduced by 58% and 73% respectively. This reflects the high level of similarity in these areas. The combination of a number of clusters into a single one does not appear to alter or influence the character of the clusters. The information statistic shows that groups which combine have the same or very similar distinguishing characteristics. Where these differ, e.g. Figure 3.14, they are only one ten percentile above or below characteristic percentile groups for the other clusters in the combination. Proportion of the total number of cells represented by each cluster is therefore greater (eight clusters now contain over 20% of the total ESD's in the area).

The clusters which are defined once the cut-off point on the dendrogram has been chosen indicate the grouping of cells on the basis of the socio-economic status of the coloured population in social space. The secondary aim of this research is however, to establish whether social differentiation among the coloureds could be identified in physical space. The link between the two can be established by transferring the clusters onto ESD maps and thus locating them in physical space. Clusters generated do not form a complete hierarchy as did modal classes and numbering on the maps are therefore not hierarchical Each character represents only one cluster and does not necessarily bear any relationship to other clusters.

Northern Suburbs

In this area six clusters emerged at the 80% level of similarity (Figure 3.14). Two clusters, 3 and 6 dominate the northern suburbs, accounting for 61% of the total number of ESD's. Cluster 1 contains a further 15%. Thus 76% of the total number of cells in the area fall into three groups. Of the ESD's in the two dominant clusters, 82% and 89% respectively occur in groups which have three or more ESD's contiguous, while cluster 1 has 60% of its ESD's grouped (Table 3.4, Figure 3.18). At the 70% level of similarity, clusters 5 & 6 and 3 & 4 combine, giving a total of four clusters for this concentration (Figure 3.14). As the cells in these cluster groups are adjacent, groups of contiguous ESD's are larger. Groups 3 & 4 now dominate the area (Table 3.5) accounting for 75% of the total number of ESD's. These clusters have 93% and 91% respectively of their ESD's forming contiguous groups (Figure 3.18). The cells of low value in the centre and north east corner identified by modal class mapping, are clearly evident in the cluster mapping at both 70% and 80% levels of similarity (Figure 3.18). Trends indicated by modal class mapping are confirmed by the cluster mapping. In physical space as well as social, then, definite social groups or social areas can be identified in the northern suburbs.

Central suburbs

Cluster analysis of the central suburbs at the 80% level of similarity (Figure 3.15) produces 16 clusters of which three - clusters 1, 2 & 5 are distinct outliers consisting of a single ESD. Cluster 10 contains the highest proportion of the total number of ESD's (18%), closely followed by cluster 8 with 16%. None of the remaining 13 clusters contains over 9% of the total number of cells in the area and none dominates the area. Clusters 3, 4, 6 & 7 together comprise only 12% of the total number of ESD's and the other 84% of the cells fall into nine different clusters. The central suburbs, at the 80%

similarity level are thus divided into a large number of small social groups with slight differences between them. Patterns and trends which emerged from average SES score plots are not at all evident (Figure 3.19).

At the 70% level of similarity, the number of clusters is reduced by half (Figure 3.15). Groups 1, 2 & 4 remain as outliers but clusters 3 & 4; 6, 7 & 7; 9, 10 & 11; 12, 13 & 14 and 15 & 16 combine to form single clusters. Cluster 6 now dominates the concentration with 38% of the total number of ESD's while clusters 5 and 7 account for a further 38% (Table 3.5). When mapped, (Figure 3.19) only 4 of the clusters at the 80% similarity level (i.e. 25%) have over 60% of their ESD's in groups of three or more contiguous cells (Table 3.4). Clusters 1 - 7 consist of scattered, non-contiguous ESD's, with the exception of clusters 3 & 4 in which 80% and 55% respectively of the ESD's are grouped. The remaining clusters (8 - 16) are combinations of scattered and contiguous cells, indicating an overall lack of contiguity of social grouping in physical space. This might have been expected from the large number of social groups identified by the cluster analysis. Trends and patterns suggested by modal class mapping can however, be recognized. When clusters at the 70% level of similarity are mapped (Figure 3.19) contiguity increases markedly as might be expected since the numbers of ESD's in each cluster is larger (Table 3.5). With the exception of clusters 1, 2 & 4 (the outliers, consisting of one ESD each), the other clusters all have over 65% of their ESD's forming groups of three or more contiguous cells. Most groups in fact consist of between 8 and 20 contiguous cells. The patterns which emerge from the 70% level cluster mapping are thus similar to those from the modal class map, although the pattern of high modal class values (70 - 79) which emerged on the western side of the area and north east of the indian residential area, are both split by 70% level cluster mapping into two groups. As in the case of the northern suburbs, distinct groups of ESD's corresponding to the clusters can be identified at the 70% similarity level, indicating the existence of different socio-economic strata in the coloured population group, both in social and in physical space.

Eastern suburbs

Eleven clusters were distinguished in the eastern suburbs at the 80% similarity level (Figure 3.16). Cluster 10 which has the highest proportion of the total number of ESD's (26%) dominates as no other cluster contains more than 14% of the total. As in the case of the central suburbs, the large number of clusters each containing only a small proportion of the total number of ESD's in the area would suggest that with perhaps the exception of cluster 10, the social differences will not be mirrored in physical grouping of the cells. However, Figure 3.20a shows that two distinct groupings emerge. Cluster 1 has 80% of its ESD's contiguous in groups of three or more (Table 3.4). These form the diagonal strip stretching from north west to south east which was identified by modal class mapping. Cluster 10 has 60% of its ESD's contiguous in groups of three or more. As this is the dominant cluster (containing the majority of the total number of ESD's, the proportion of grouped cells is less than expected. Clusters at the 80% level of similarity show little contiguity when mapped (Figure 3.20) and patterns bear resemblance to those obtained from modal class mapping only in respect of the diagonal running across the main section of the eastern residential concentration. The number of clusters at the 70% similarity level is reduced to three (Table 3.3) and each cluster now contains a larger percentage of the total number of ESD's in the area. Cluster 3 contains the majority of cells (56%) followed by cluster 1 with 32% of the total (Table 3.5). Very distinctive differences in social space are thus identifiable. Contiguity is far higher in each of the groups resulting from 70% level cluster mapping (Figure 3.20). This might however, have been expected because of the large number of ESD's and small number of clusters. Only cluster 2 has less than 87% of its in groups of three or more while the other two clusters essentially form one or two large groups of cells. This closely resembles the pattern of mapped modal classes 10 - 19 and 30 - 39. Overall the trends suggested by mean SES score and modal class mapping are intensified by the mapping of clusters resulting from a cut-off at 70% level of similarity on the dendrogram. The grouping in social space at this level is clearly reflected in physical space in the eastern suburbs.

Southern suburbs

Here 12 clusters emerge at the 80% similarity level (Figure 3.17), but no single cluster dominates the concentration. The large number of clusters and the small percentage of the total number of cells in each (Table 3.4) suggests that when mapped, the cells will not form coherent groups reflecting the social grouping. The cluster with the highest proportion of the total ESD's (although this is only 15%) is cluster 7 (Table 3.4) closely followed by cluster 5 with 13%. Cluster 6 represents only 12% and cluster 10 a small 9% of the total number of ESD's in the southern suburbs.62% and 70% respectively of their cells form contiguous groups of three or more. Together these clusters account for 49% of the total number of ESD's in the area. Clusters 1, 9 & 11 consist entirely of scattered ESD's and the remainder of the clusters show a combination of scattered and contiguous cells (Figure 3,21). At the 70% level of similarity however, (Figure 3.17) the number of clusters has decreased to five. Each cluster therefore contains a higher proportion of the total number of cells (Table 3.5). Cluster 3 dominates the area now, accounting for 37% of the cells and these form two large groups (15 and 17 cells) of contiguous ESD's when mapped (Figure 3.21). Cluster 2 also contains a significant proportion of the total number of cells (22%) of which 75% form small groups (3 to 5 cells) of contiguous ESD's. The remaining three clusters make up the balance of 24% of the total number of cells and all have over 45% of their ESD's forming small groups of contiguous cells. Clusters mapped at the 80% level of similarity (Figure 3.12) do not show such clear patterns of socio-economic differentiation as does the mapping of modal classes. The trends of broken topography illustrated

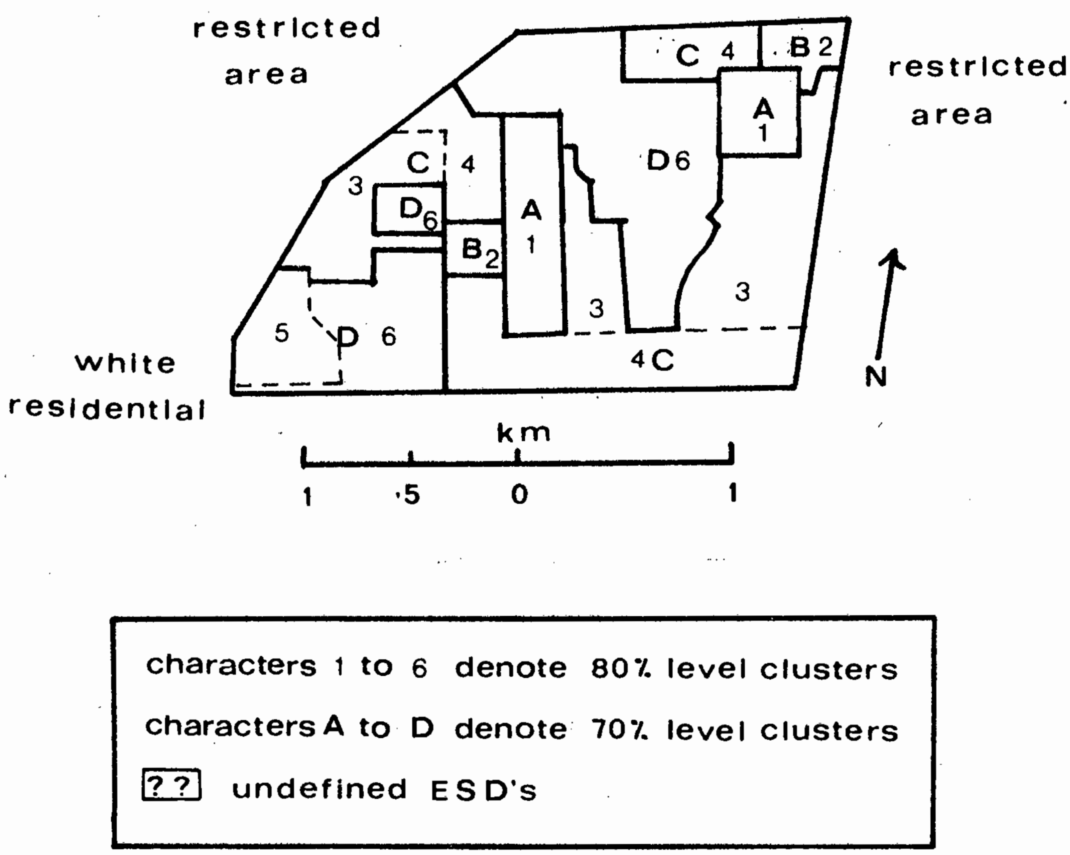


Figure 3.18 Cluster patterns: northern suburbs

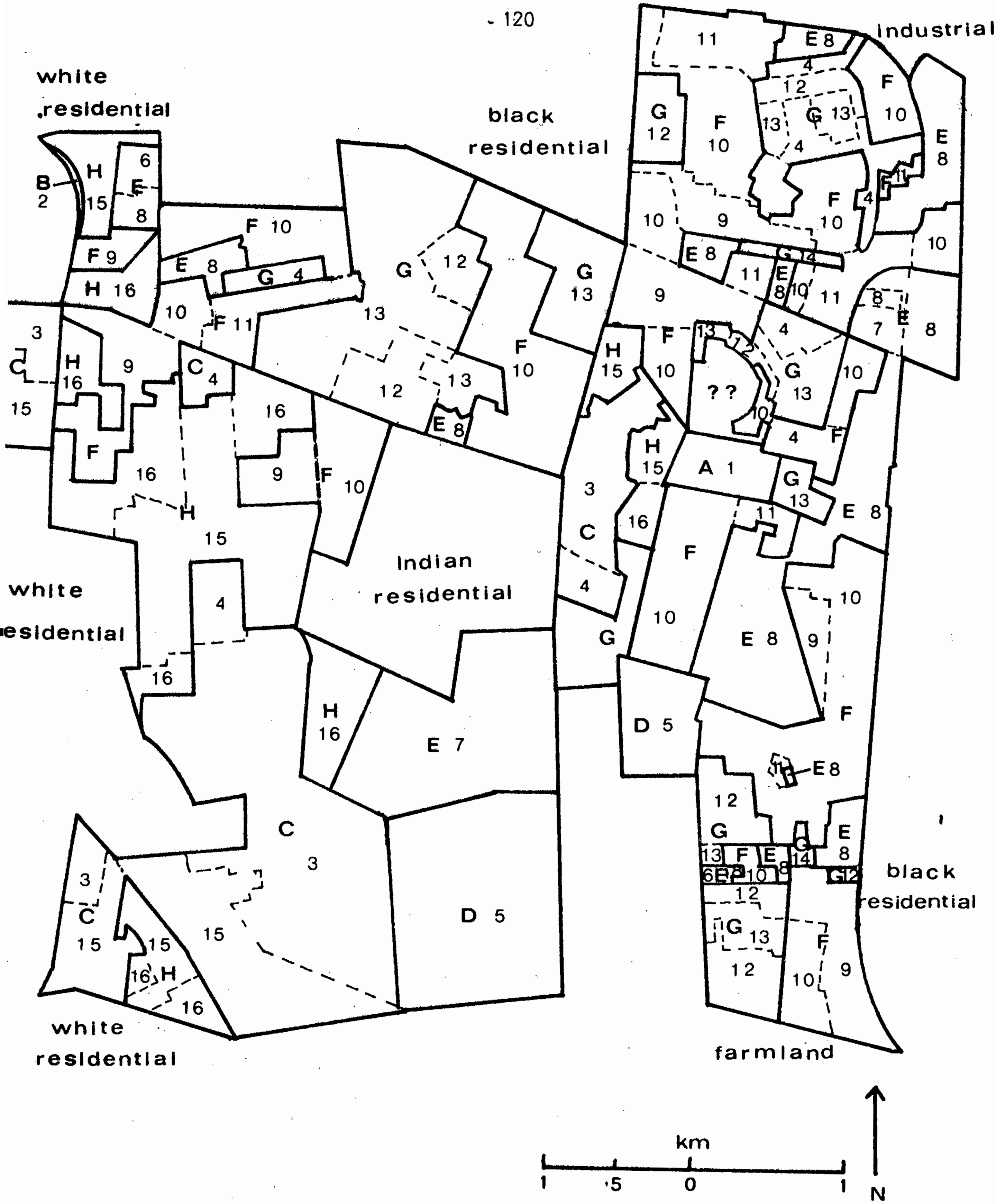


Figure 3.19 Cluster patterns:

central suburbs

characters 1 to 16 denote 80% level clusters

characters A to H denote 70% level clusters

[??] undefined ESD's

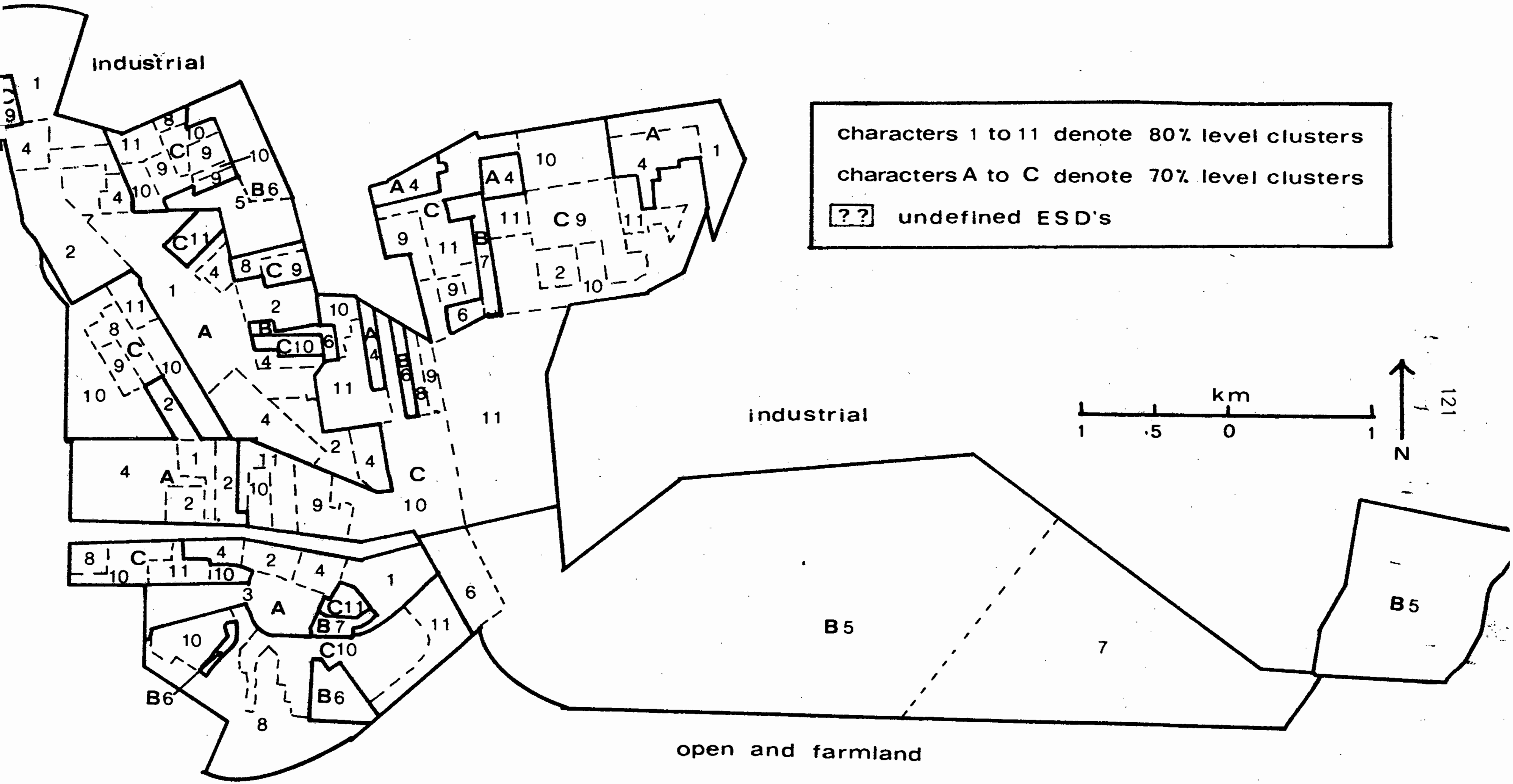


Figure 3.20 Cluster patterns: eastern suburbs

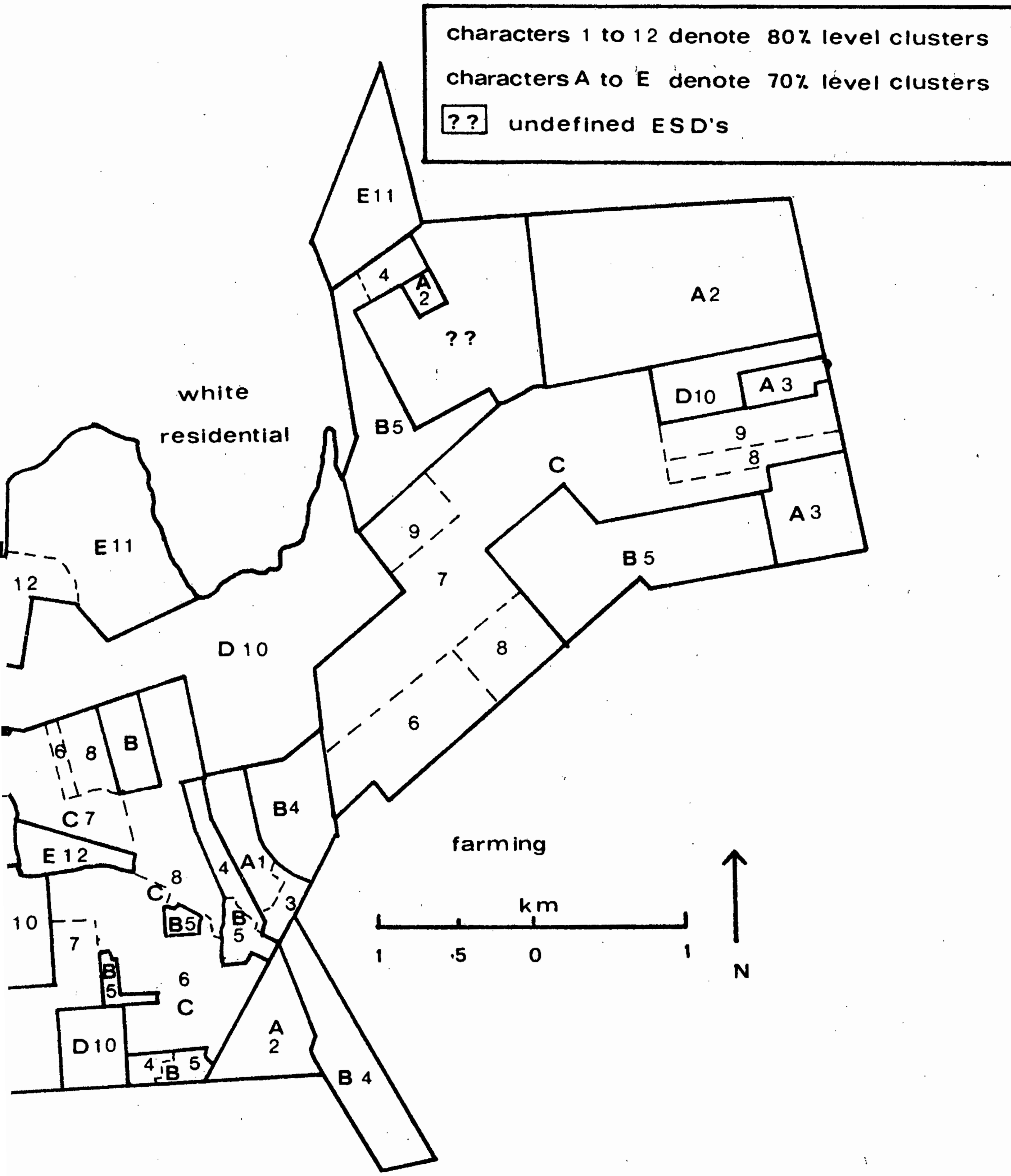


Figure 3.21 Cluster patterns: southern suburbs

TABLE 3.5 Contiguity in residential concentrations: 70% level clusters

Cluster	Group of 1 ESD	Prop. of group	Group of 2 ESD	Prop of group	Group of 3+ ESD	Prop. of group	No.	Prop. of total ESD's
N 1	0	-	1	40%	1	60%	5	15%
N 2	1	33%	1	67%	0	-	3	9%
W 3	1	7%	0	-	2	93%	14	42%
N 4	1	9%	0	-	2	91%	11	33%
C 1	1	100%	0	-	0	-	1	1%
C 2	1	100%	0	-	0	-	1	1%
C 3	1	5%	0	-	3	95%	21	10%
C 4	1	100%	0	-	0	-	1	17%
C 5	7	19%	3	16%	3	65%	37	1%
C 6	2	2%	1	2%	9	96%	83	38%
C 7	5	11%	4	17%	4	72%	46	21%
C 8	3	10%	2	14%	2	76%	29	13%
E 1	4	9%	1	4%	2	87%	46	32%
E 2	3	17%	3	33%	2	50%	18	12%
E 3	4	5%	1	2%	5	93%	82	56%
S 1	3	33%	1	22%	1	45%	9	10%
S 2	3	15%	1	10%	4	75%	20	22%
S 3	0	-	0	-	2	100%	34	37%
S 4	3	38%	0	-	1	62%	8	9%
S 5	2	40%	0	-	1	60%	5	5%

by modal class mapping and average SES score plots are however, well reflected in the 80% level cluster map. The patterns which emerged from modal class mapping are (as in the case of the eastern and central suburbs) intensified and very clearly illustrated by mapping at the 70% level of similarity. Once again therefore, differences in socio-economic status and groupings identified in social space for the southern residential concentration are mirrored in physical space.

In each of the residential concentrations, the trends and general patterns identified from modal class mapping are recognisable in the 80% level cluster mapping. Patterns tend to be fragmented on account of the large number of clusters and the small proportion of the total ESD's in the area contained in each cluster. When the 70% cluster groups are mapped, however, number of clusters is greatly reduced and proportion of the total number of cells in each cluster is consequently increased. Patterns shown by the 80% cluster maps are generalized in the 70% ones, as clusters combine and more adjacent ESD's group together. This is most noticeable in the eastern suburbs where the number of groups decreases from 11 to 3. The 70% cluster patterns very closely resemble the patterns produced by mapping of modal classes although in a limited number of cases of these groups are split and others are amalgamated. Although in the northern suburbs patterns from modal class are generalized by the 70% level clusters, the cells of low value in the centre and north eastern corner remain clearly distinguishable. In the central suburbs the trend is also to a generalization of modal class patterns except in the case of the high modal class (70 - 79) areas in the west and to the east of the Indian residential area which are divided into two groups. The distinct diagonal pattern in the main section and the general trend to grouping elsewhere in the area identified by modal class were, as expected, easily distinguishable from the 70% cluster mapping in the eastern area. In the southern concentration too, modal class and 80% cluster patterns were amalgamated by 70% level clusters with the exception of the plateau of high modal class values (70 - 79) in the north west and the groups of modal class 30 - 39 in the far south and east which are disaggregated.

An explanation for the differences in patterns between modal class and 70% clusters is that while cells are clustered on the basis of *similarity* of one ten percentile group in modal class mapping, 70% clusters distinguish similarity on the basis of all ten percentile groups, thus giving more representative groupings. In terms of the generalization of 80% cluster patterns by the 70% ones, contiguity must increase when number of clusters decreases, as numbers of cells in each cluster is larger.

Clusters of ESD's of similar socio-economic status have thus been identified in social space in each of the coloured residential concentrations in Greater Cape Town. Mapping of these clusters defined at different levels of similarity has indicated the existence of groups of contiguous cells in physical space. Social differentiation among the coloured population in Greater Cape Town thus has a spatial component.

ii) Patterns emerging when residential areas are regarded as a single system

Since the four concentrations form part of a whole, viz. the residential area for a single population group (the coloureds) and there is no reason to suppose that they are socially independent and do not influence each other, an overall cluster analysis had to be run. Computer limitations and the desire to minimise generality meant that the 80% level clusters defined for each residential concentration provided the base from which average profiles could be constructed. Aggregates of percentage occurrence in each of the ten percentiles between zero and one hundred were calculated for the 45 clusters defined at the 80% similarity level, giving one overall profile per cluster. These profiles were then grouped using the Bray-Curtis coefficient.

TABLE 3.6 Overall cluster contiguity (490 ESD's)

Cluster	1 ESD	Prop.	2 ESD	Prop.	3+ ESD	Prop.	No.	Prop. of total
1	1	100%	-	-	0	-	1	1%
2	1	100%	-	-	0	-	1	1%
3	4	8%	2	8%	4	84%	49	10%
4	9	4%	1	1%	8	95%	223	45%
5	11	52%	3	29%	1	19%	22	4%
6	9	5%	5	6%	13	87%	167	34%
							463	94%

27 ESD's not taken into account

The resulting dendrogram is again remarkably flat (Figure 3.22). Clustering begins between the 95% and 100% similarity level. At 45% similarity there are only three clusters and the grouping process is complete by a level of similarity of 25%. At the 80% similarity level, 11 definite clusters can be identified, but each contains only a very small proportion of the total number of cells. Approximately half of these 11 clusters have joined by the 70% similarity level where there are six clearly defined clusters. This 70% level of similarity was used as a cut-off point as it was here that distinct clusters were easily recognisable and the level of generalization was not too great. Six clusters were identified for the study area at the 70% similarity level and these will be referred to as "overall clusters" (Figure 3.22).

The six clusters form a hierarchy of differing socio-economic status. Clusters were plotted on ESD maps (Figures 3.23 - 3.26a) and contour diagrams (Figures 3.23 - 3.26b) and perspective diagrams (Figure 3.27) were constructed using the Saclant Graphics Package (U.C.T. Computing Service, 1979), to show the distribution of the overall clusters over the entire study area. Of these clusters, two are distinctly different from all the rest, each consisting of only one ESD. Both are located in the central suburbs but are not adjacent to one another and were identified as belonging to modal class 60 - 69 and 70 - 79. On the basis of the information statistic both are characterised by values in the 0 - 9 percentile, though cluster 1 is identified by also containing values in the 20 - 29 percentile. Essentially, therefore, only four clusters are distinguished at the 70% similarity level. Cluster 5 is the smallest multicell cluster and contains only 10% of the total number of cells (Table 3.6). It is characterised by a mixture of the whole range of percentiles and with the exception of clusters 1 and 2 has the lowest status. Cluster 5 has 52% of its ESD's non-contiguous or isolated and only 19% forming groups of three or more. Cluster 3 forms only 10. of the total number of cells and although it has a medium status, suggested by the dominance of the percentile 40 - 49 it is also distinguished by values in the 90 - 100 percentile. Cluster 6 accounts for a substantial proportion of the total (34%) and is distinguished by a fairly high status indicated by the dominance of percentiles between 50 and 89. Clusters 3 and 6 together include 44% of the total number of cells and have 84% and 87% of these forming groups of 3 to 25 and 4 to 60 contiguous cells respectively. These are generally small to medium sized ESD's. Cluster 4, containing 45% of the total number of ESD's is the dominant cluster. Characterised by percentiles 0 - 9 and 20 - 29 this cluster has a fairly low social status. 95% of the cells in this cluster form contiguous groups ranging in size from three to over one hundred. These ESD's range in size too, from large to very small. Contiguity in all groups could be expected to be high on account of the number of cells and the very small number of clusters. Thus a type of hierarchy exists, ranging from a low social status (15) through medium low (25) and medium (45) to a high social status (55). The northern suburbs is dominated by cells of value 45, giving it a medium-low social status. Medium-low and medium status characterise the central suburbs while the status of the southern area overall, is medium. Areally the eastern suburbs is dominated by a low social status (i.e. the panhandle) but the main section of this area is divided between high and medium-low status values. Although large portions of each of the four residential concentrations (northern, central, eastern and southern suburbs) are covered by values for one cluster, peaks and valleys of differing clusters are clearly in evidence. The "topography" created by the differentiation of the study area by clusters is well illustrated in the three-dimensional diagram (Figure 3.27). The northern suburbs have a central ridge and peaks of high values in the north east and south west corners (Figure 3.23b) identified earlier by both 80% and 70% clusters. Cells of low value can also again be recognised just off centre and in the north east corner. The patterns which result when the overall clusters are mapped for this area are almost identical to those obtained from mapping by modal class (Figure 3.32a).

The central suburbs shows a marked contrast between medium and low values on the western side and low and high values in the east (Figure 3.24b). The lowest value (cluster 5) is represented by only four scattered cells while the area is dominated by cluster 4, also of fairly low value (Figure 3.24a). Higher values are in evidence on the eastern and western sides of the area. The plateaux of high values which emerged along the western side of the central suburbs and east of the indian residential area from modal class mapping are again clearly in evidence in the overall cluster mapping. The high values located on the eastern side of the area in overall cluster mapping corresponds to a similar pattern of cells in the 40 - 49 class in modal class mapping. The patterns and trends which emerged from modal class mapping in the central suburbs are thus repeated when overall clusters are mapped for this area.

The eastern suburbs are dominated by contrasting high and low values in the main section. The panhandle consists of uniform very low values (Figure 3.25b). A plateau of high values occurs in the north eastern section of the area and extends southwards, intruding into the diagonal of low values which runs from north west to south east (Figure 3.27). The high values extend to the southernmost section of the eastern suburbs and two small enclaves of high values are found on the western side of the diagonal. There are no cells of medium value in the area and as for the central suburbs, the lowest value is confined to a small number of scattered cells (Figure 3.25a). The same pattern (diagonal of the same values running across the area from north east to south west) which emerged in modal class mapping is clearly evident in the overall cluster mapping.

The southern area has a gentle but more broken topography (Figure 3.25b), indicative of alternating high and low values. Three plateaux are in evidence, the north western one having slightly higher value than that of the more southerly one, while the third, at the most northerly point of the area, is of middle value (Figure 3.27). A larger number of cells of the lowest value occur in this area than in the other con-

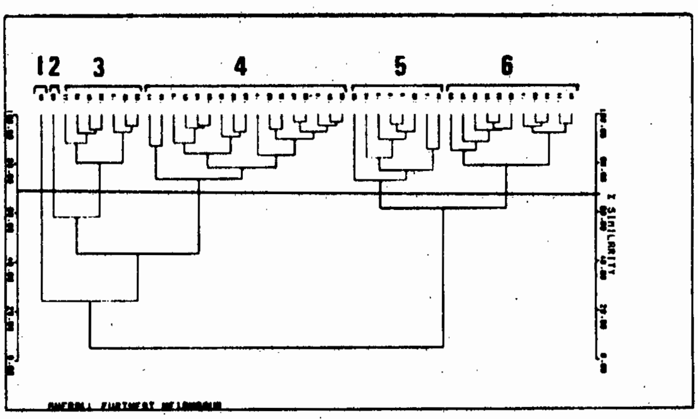
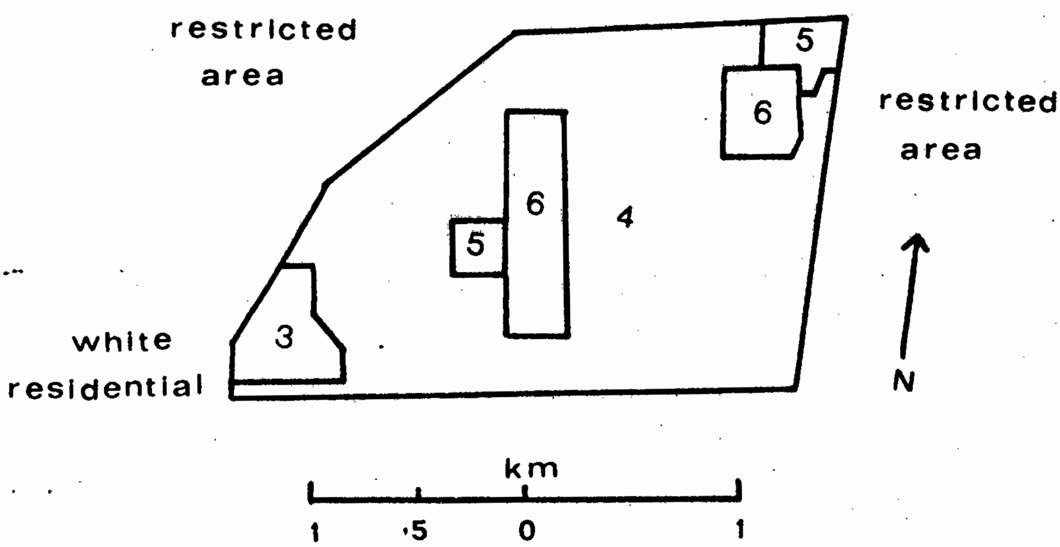
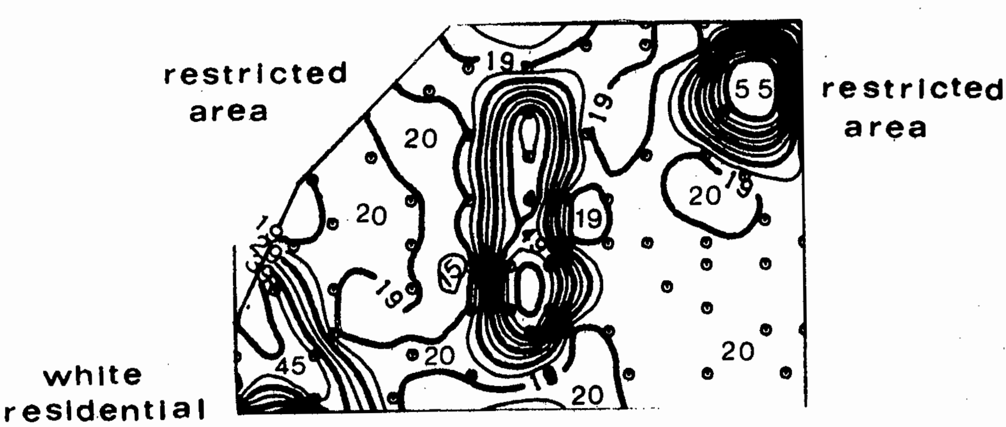


Figure 3.22 Overall dendrogram

a) Location



b) Contours

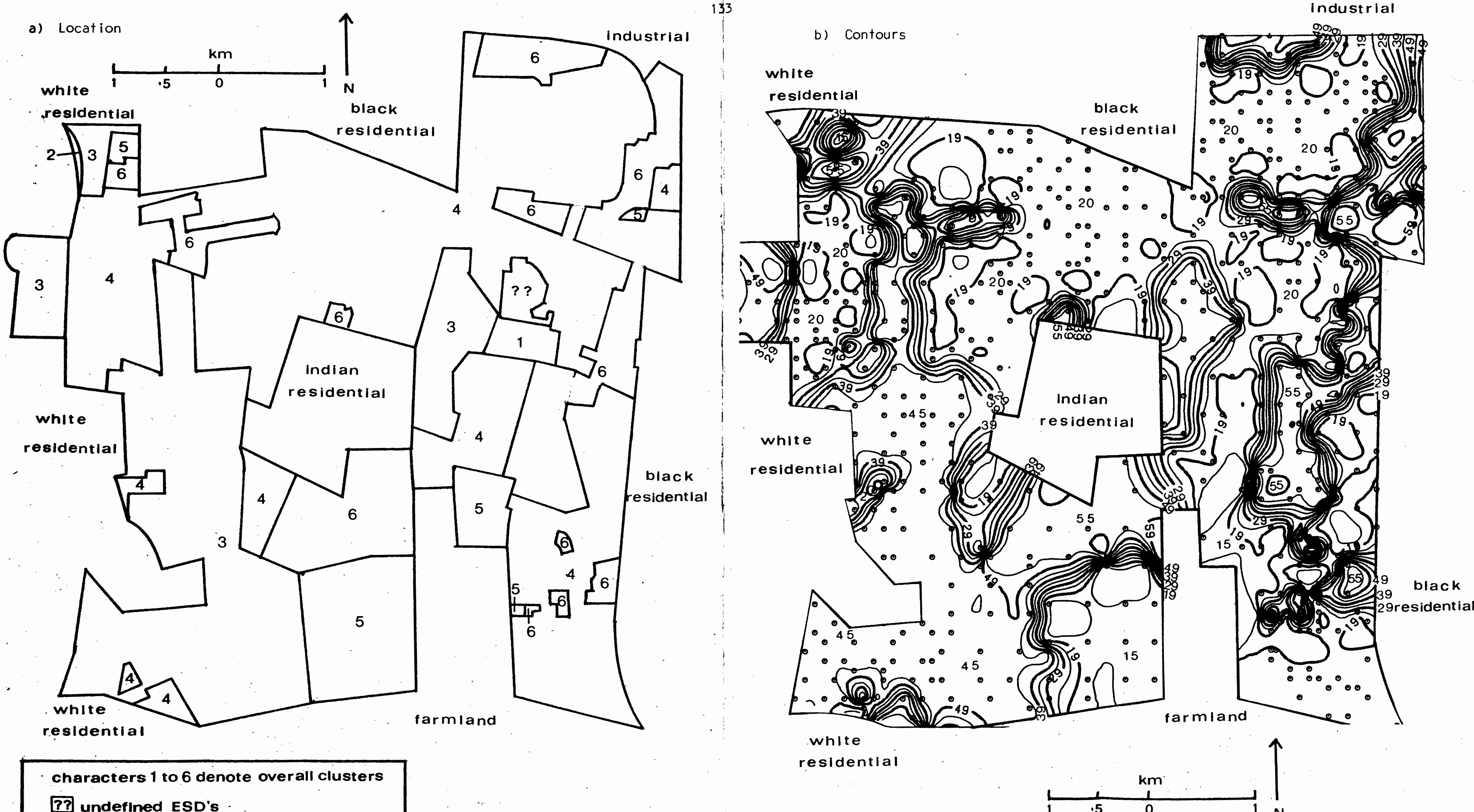


characters 1 to 6 denote overall clusters

[??] undefined ESD's

value	modal class	cluster
5	0 – 9	2
15	10 – 19	5
20	0 – 9 20 – 29	4
25	20 – 29	1
45	40 – 49	3
55	50 – 59	6

Figure 3.23 Overall cluster patterns: northern suburbs



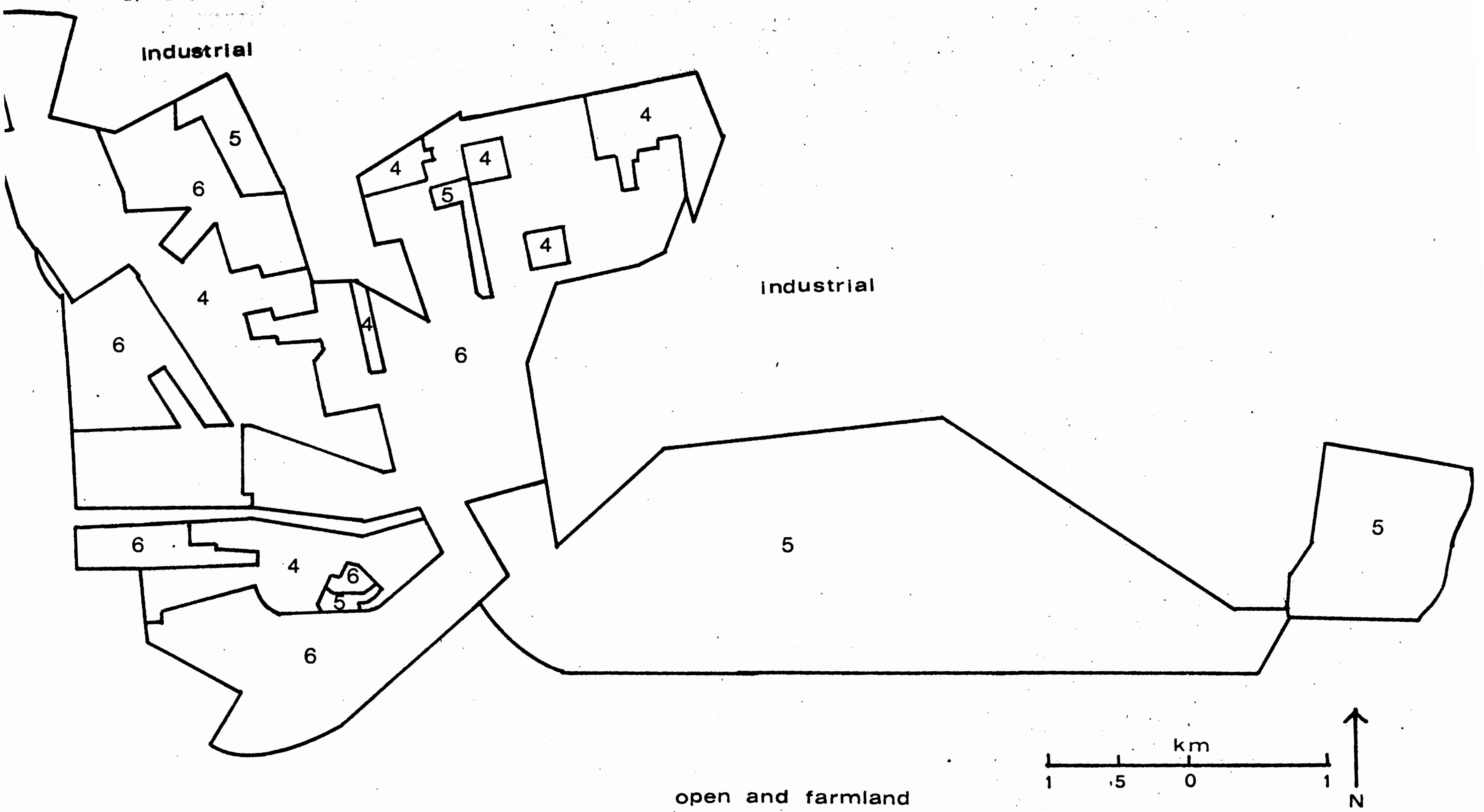
characters 1 to 6 denote overall clusters

?? undefined ESD's

value	modal class	cluster
5	0 – 9	2
15	10 – 19	5
20	0 – 9 20 – 29	4
25	20 – 29	1
45	40 – 49	3
55	50 – 59	6

Figure 3.24 Overall cluster patterns: central suburbs

a) Location



characters 1 to 6 denote overall clusters

?? undefined ESD's

value	modal class	cluster
5	0 – 9	2
15	10 – 19	5
20	0 – 9 20 – 29	4
25	20 – 29	1
45	40 – 49	3
55	50 – 59	6

b) Contours

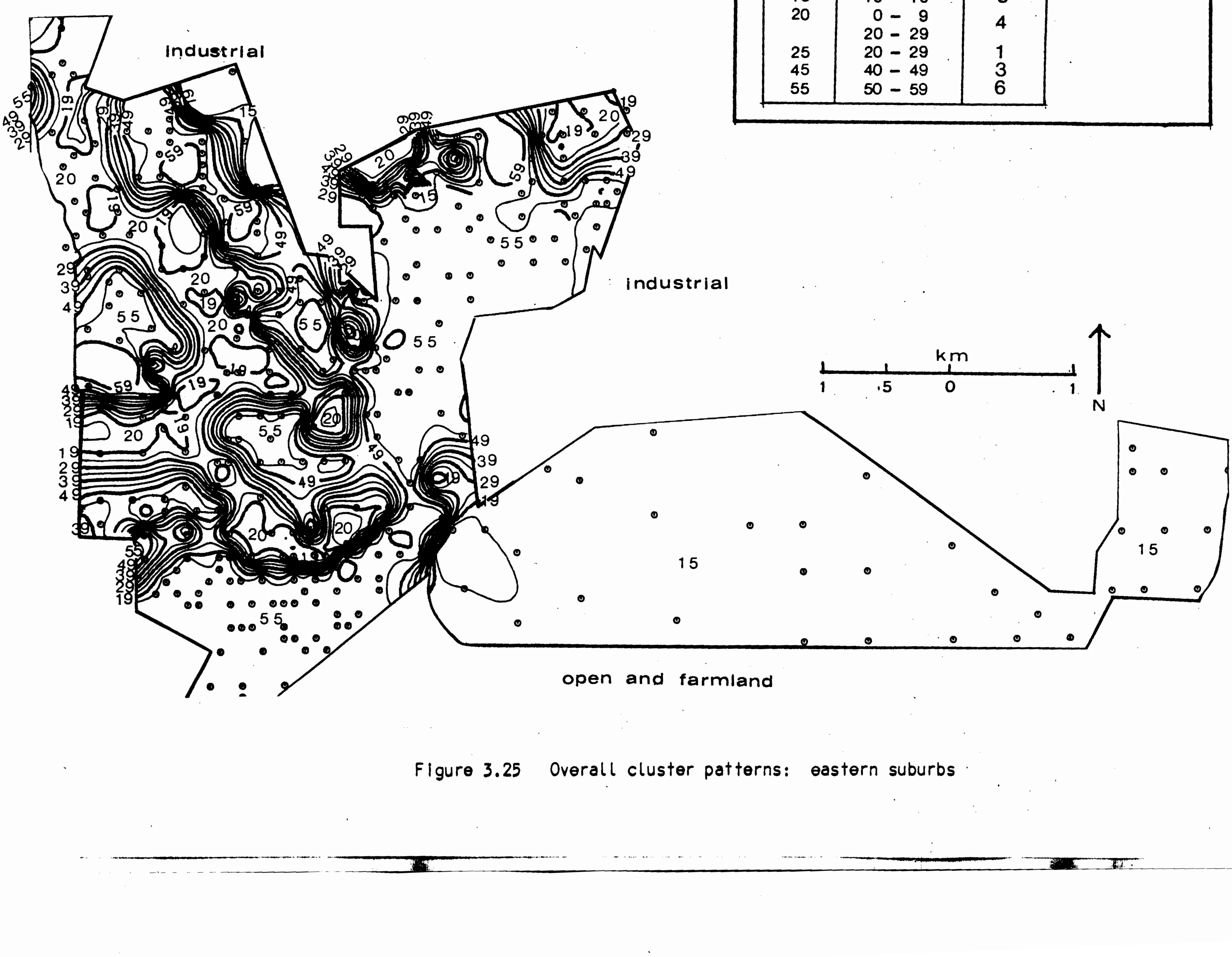
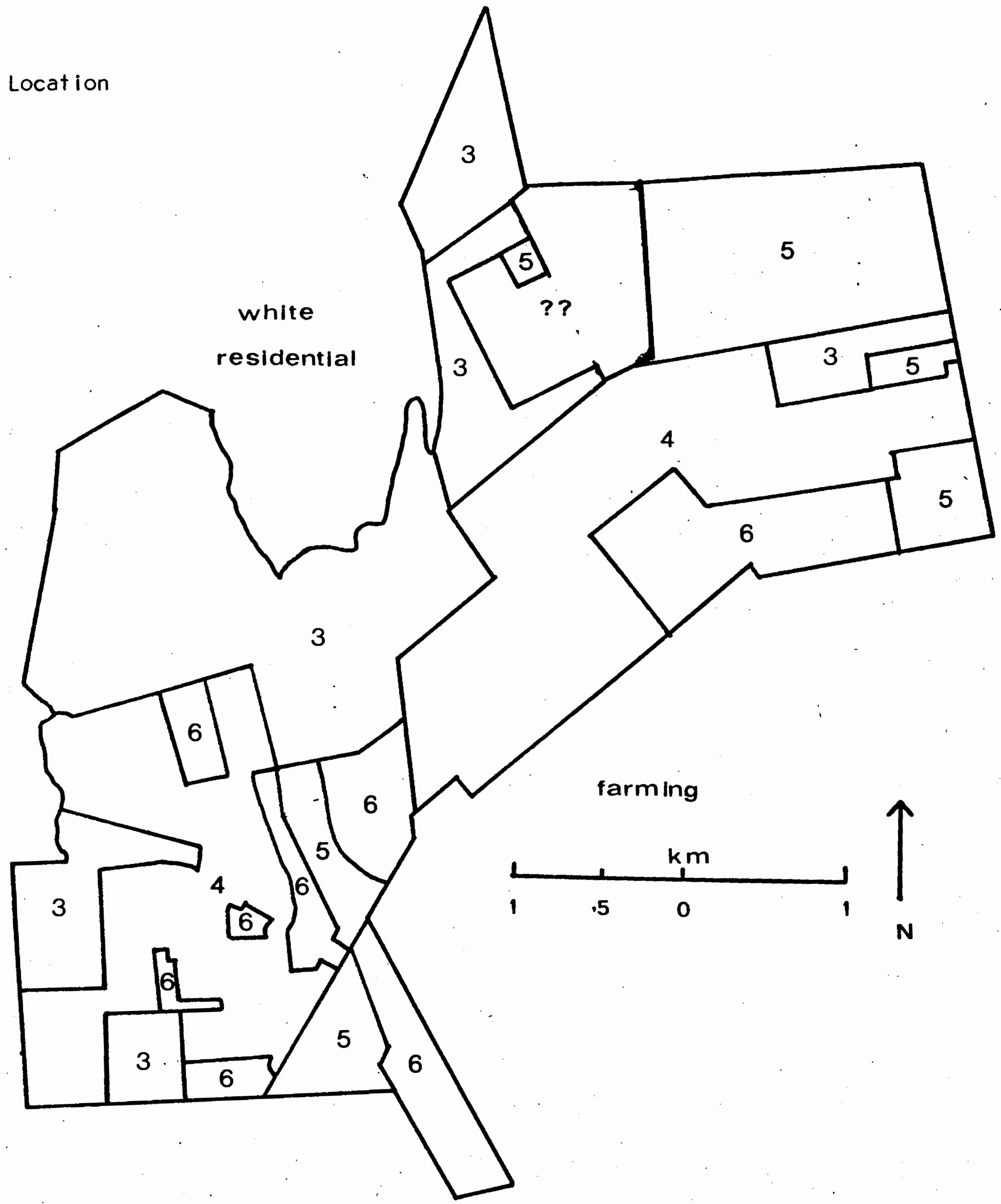


Figure 3.25 Overall cluster patterns: eastern suburbs

a) Location



characters 1 to 6 denote overall clusters

?? undefined ESD's

value	modal class	cluster
5	0 – 9	2
15	10 – 19	5
20	0 – 9 20 – 29	4
25	20 – 29	1
45	40 – 49	3
55	50 – 59	6

b) Contours

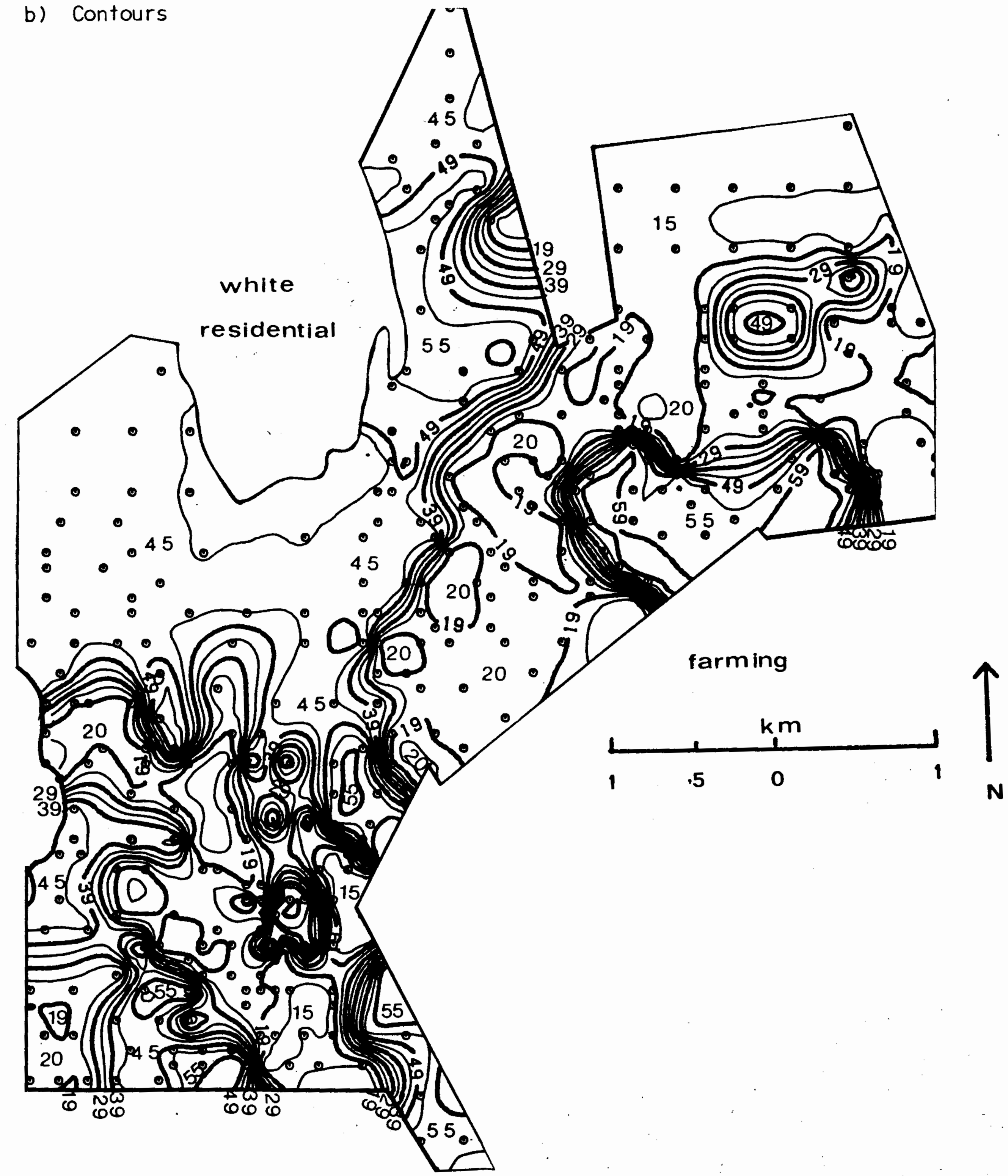


Figure 3.26 Overall cluster patterns: southern suburbs

Figure 3.27 Perspective diagram of overall cluster distribution in the study area

centrations in the study area (Figure 3.26a). They are also not as scattered here as elsewhere. Once again the patterns and trends indicated by modal class mapping are confirmed by overall cluster mapping. This is evident in the occurrence of the plateau of high value in the northwestern section of the area which borders on a white residential area.

Socio-economic differentiation into clearly defined status groups has been carried out for coloureds living in the designated residential areas which form the study area. That social stratification occurs in social space may be inferred as existing in the coloured population from differences in average SES scores and groups of ESD's classed by the cluster analysis as similar on the basis of the socio-economic characteristics of their populations. The question which must be answered, however, is whether the groups defined in social space form genuine groups in physical space, or whether these groups are coincidental. Data is taken from the census of each ESD. As such, the characteristics of the census data can be related to a particular location, i.e. a latitudinal and longitudinal position in physical space. It does not however, necessarily mean that social areas in social space are directly translatable into the physical space from which the data were drawn. Mapping of the cluster information produces patterns in each of the residential concentrations and reveals that groups of ESD's exist. This implies that the clusters identified in social space are translatable into physical space despite the constraints on residential areas and housing in which the racial group functions. The identification of clusters of ESD's and the generally high level of contiguity among ESD's of the same groups points to the existence of distinct social areas among the coloured population of Greater Cape Town. Mapping of both descriptive statistics (modal class) and clusters derived by numerical classification was not by any means identical, although the resulting distinctive patterns of grouping of ESD's and general trends were readily identifiable and very similar for both cases. The locations of groups of high, low and medium values in each case begs the question of the locational relationships between and spatial dependence of, these areas of differing social status on each other. It may be asked whether the existence of a high social status group of ESD's in a particular area determines the social status of adjacent ESD's in terms of either similarity or difference. The means whereby this spatial dependence may be ascertained is the application of the Cliff & Ord (1969) autocorrelation statistic.

CHAPTER IV

SPATIAL DEPENDENCE IN PATTERNS OF SOCIO-ECONOMIC DIFFERENTIATION

I Spatial Dependence

Mapping of the various forms of socio-economic status scores, i.e. mean scores, modal class and full profiles, has facilitated recognition of contiguity between cells of the same social status in the study area. Groups of ESD's occur which have similar socio-economic status and may be said to form distinct social areas. Patterns of social differentiation have thus been identified in the coloured residential concentrations of Greater Cape Town and the primary aim of this study has thus been achieved.

A further aim is to establish the degree of spatial dependence in social differentiation, i.e. whether the status of an area necessarily determines the status of adjacent areas. This can be realised by applying the Cliff & Ord (1969) technique of spatial autocorrelation to data which, when mapped gives rise to groups. The nature of the cells comprising the study area (i.e. irregular shapes and sizes of ESD's) make the Cliff & Ord statistic more appropriate than other correlation methods as it is the only technique for spatial correlation and it has a weighting mechanism which provides for the standardization of the sub-areas relative to one another.

Spatial autocorrelation involves application of simple correlation procedures to assess the relationship between values of a single set of data. It measures whether the presence of some quality in an area makes its presence more or less likely in neighbouring areas. "If high values of a variable in one area are associated with high values of that variable in neighbouring areas... the set of areas exhibits positive autocorrelation with regard to that variable. Conversely when high and low values alternate, the spatial autocorrelation is negative" (Cliff *et al.*, 1975, p.145).

Measuring autocorrelation in spatially distributed data is different from measuring autocorrelation in a time series because dependence extends in all directions rather than in a single direction. The study of spatial autocorrelation therefore is important because an observation is seen to be a function of all neighbouring values and by calculating the autocorrelation coefficient, the degree to which every value on a two dimensional surface is affected by the interaction between itself and neighbouring values is established (Cliff *et al.*, 1975).

a) The Autocorrelation Coefficient

The Cliff & Ord statistic used in this study was developed from earlier statistics put forward by Geary (1968) and Moran (1950). This statistic has advantages over the earlier ones by virtue of its incorporation of a generalised interaction function w_{ij}. The reason for inclusion of the interaction function was firstly to extend the adaptability of the statistic because the concept of a neighbouring value need not necessarily be restricted to interaction between adjacent cells only (as was the case with the Geary and Moran statistics). Secondly, if the surface comprises irregular sub-areas, the relative sizes of these may be standardised by the use of w_{ij} values. The statistic can however, still be applied to the regular subarea case using adjacent interaction only (Ward, 1978) if w_{ij} is set to 1. In Ward's (1978) application of the technique, therefore, only first order interactions will be taken into account, and no standardization will take place between sub-areas.

Realizing the usefulness of the spatial autocorrelation statistic to those geographers concerned with establishing the extent to which space per se plays a part in a distribution pattern, Ward (1978) wrote a computer programme (Appendix 3) to calculate the autocorrelation coefficient given by Cliff & Ord (1973, p.12) as:

$$r = \frac{n \; (2) \; w_{ij} \; z_i \; z_j}{W \sum_{i=j}^{n} z_i^2}$$

where:

$$(2) = \sum_{i=1}^{n} \quad \sum_{j=1}^{n}$$

w_{ij} = matrix of values indicating the nature of the interaction between x_i and x_j.

z_i = $(x_i - \bar{x})$

z_j = $(x_j - \bar{x})$

n = the number of x_i values

W = the total number of joins in the system.

Very simply the method by which the autocorrelation coefficient is computed is as follows: The mean of the data set is calculated. The difference between the mean and the value in a cell (x_i) is obtained. This is the ZDIF score. This score is squared and stored, and each successive ZDIF² score for each x_i cell is added to it. The final value is known as the SUMZ value. At the same time the ZDIF score for an adjacent cell x_j is calculated. These two ZDIF values, viz. $(x_i - \bar{x})$ and $(x_j - \bar{x})$, are then multiplied together. The product is multiplied by the weighting factor w_{ij} for those two cells. This is stored and each successive value so obtained is added to this value giving a total score, known as TOTAL or (2) in the equation. TOTAL is finally multiplied by the number of cells in the data set NNNN (n in the equation). This product gives the numerator of the equation. The total SUMZ value (as calculated above and given by

$$\sum_{i=j}^{N} z_i^2$$

in the equation is then multiplied by the total number of joins in the system (W). This product forms the denominator of the equation. The autocorrelation coefficient r is obtained by dividing the numerator by the denominator. By limiting the summation to specific rows and/or columns the statistic can be modified to give correlations in specific directions instead of an overall correlation. The order of the dependence (i.e. first order (or interaction between adjacent cells) or second and third order (or interaction between next but one or next but two cells) may also be specified by adjusting the matrix for the data input.

Cliff & Ord (1971) prove that as the number of observations increase, so the distribution of their coefficient tends towards normality and as such r calculated can be tested as a standard normal deviate for significance. A test for kurtosis is included, as are calculations of expected values of the coefficient under the assumptions of normality and of randomness. Finally, significance is evaluated according to both the assumptions of a randomly and a normally distributed variable. Comparison of the Cliff & Ord (1969) method with Pearson's method of correlation (Ward, 1978) has established that the manner in which geographers have traditionally interpreted coefficients from the Pearson technique can possibly be applicable to the Cliff & Ord statistic.

Ward (1978) however, identifies Cliff & Ord's failure to strongly emphasize that which is of importance to geographers, viz. the reaction of the statistic in practical situations and how best to interpret the results. By applying a regular square matrix he was able to highlight the effect of the five factors which influence the autocorrelation coefficient. As the distribution of the Cliff & Ord (1971) coefficient is so strongly affected by the five factors (although Cliff & Ord (1971) identified 3 broad factors affecting the distribution of their statistic, these can be subdivided, and given as five separate but interrelated ones), the data for this study will be considered with regard to each of them.

i) The distribution of the variable

The Cliff & Ord correlation test is a parametric test and there is therefore a need to establish whether or not data are normally distributed. The variable to be used for calculation of the co-efficient is the SES score. The method by which these scores are calculated ensures that mean scores are normally distributed over the entire coloured population. Modal class and median values are also normally distributed. In any sub-set of data, however, normality cannot be assumed. According to Silk (1979, p.118) though, the probability of a cell having a high or low value in free sampling should not be less than 0,2 if the data is assumed to be normally distributed. In the case of SES scores calculated probability values for the distribution of mean values is $p = 0,36$; of a modal class value is $p = 0,70$ and for the overall cluster values at 70% similarity level is $p = 0,60$. The limiting conditions therefore hold for all the SES score data i.e. mean values, modal class or cluster data and the Cliff & Ord autocorrelation coefficient can be applied.

ii) The size of the sub-area system

In this project the study area consists of a total of 490 ESD's. An interaction matrix therefore would require the calculation of a single directional weighting matrix for approximately 1960 joins. Such an undertaking would constitute a project on its own in terms of the time required for these calculations. Further, drum capacity on the computer is limited and special changes would have to be made to programs for space allocation. Even when broken down into sub-areas, three of the four residential concentrations, viz. southern, central and eastern suburbs, consisting of 93, 218 and 146 cells respectively, likewise present serious data handling problems. In order therefore to obtain some measure of spatial dependence for the study area as a whole it was decided to draw a sample from the population. On account of having to consider only contiguous ESD's, the sample was stratified re: residential concentrations.

Silk (1979, p.118) states with regard to the size of the sub-area system that approximate normality is ensured as long as the number of cells (n) in the study area is greater than 20 and preferably 30 or more. In order to meet the requirements for the assumption of normality, the northern suburbs, which consist of only 33 ESD's, are regarded as a single sample. In the case of the southern, central and eastern suburbs samples were drawn as follows: The ESD's for each area were numbered consecutively. Using a random number table a cell was selected and located on an ESD map of the area. Adjacent ESD's surrounding this cell were then chosen, the number being greater than 30 (thus sufficiently large for the assumption of normality) but varying for each of the three concentrations in order to include a compact area. Table 4.1 shows the numbers of ESD's considered for each of the residential concentrations as well as the proportions which samples formed relative to the areas from which they were drawn, and to the study area as a whole.

TABLE 4.1 Sample sizes for autocorrelation

Area	No. of cells	% of area	% of total
North	33	100	7
South	37	40	7
Central	35	15	7
East	36	25	7

iii) The shape of the system of sub-areas, i.e. the length of the perimeter. The ratio of perimeter to area increases as a shape becomes less compact and thus means that the influence of the boundary becomes steadily more marked. A circle is the most compact configuration and will have the smallest perimeter in relation to its area. If an area of 4 is assumed for a circle, a square and a rectangle, their perimeters are respectively 7, 8 and 10. Other polygons would have much large perimeters for the same area. On account of their generally non-compact configurations, each of the coloured residential areas to a certain extent, the effect cannot be ignored. The apparent relationship between social status in coloured residential areas and adjacent landuse presents a problem similar in nature to the boundary problem encountered in Nearest Neighbour analysis in which an occurrence outside the boundary may be closer to the point under consideration than any within the limits of the boundary (Dacey, 1963; De Vos, 1973). It must then be established whether the boundary is real or artificial. If it is considered to be real, it is acceptable to ignore occurrences of the variable outside the boundary, as these are not seen to influence the distribution within the boundary in any way. If the boundary is artificial, then within the delimited area, patterns identified are not valid, as inward interaction is emphasized.

In the case of this particular study the concentration is on one single ethnic group and the boundaries are "real" in that they are legally, if artifically defined to delimit the residential areas of that group. However, the problem is compounded by the fact that the coloured residential units form part of an urban system made up of two different socio-economic groups, viz. white and non-white. While the boundaries of the coloured residential areas are officially delimited and may as such be regarded as "real" boundaries the influence of adjacent landuse and relationships with the rest of the system tend to suggest that the legal boundaries are in fact artificial. A further phase of the study of social differentiation in coloured residential areas (which is not within the scope of this project), should therefore be an investigation into the problem of whether social status changes across white-coloured boundaries and whether these are in fact "real" boundaries.

iv) Interaction

Cliff & Ord (1969) derived their statistic so that interaction need not necessarily be restricted to adjacent cells only (i.e. first order relationships) but could also be applied when interaction was with second and third order cells. Correlation could thus be measured between cells, either vertically or horizontally removed from the original cell (rooks case) or diagonally non-adjacent (bishops case).

As social status is assumed to be reflected in housing any influence would not be felt further away than the adjacent ESD's because individuals tend to associate with and be influenced by their immediate neighbourhood. The average number of dwellings in an ESD is 160 and therefore closest neighbours are likely to fall within the same or an adjacent ESD. Only first order interaction between immediately adjacent ESD's is therefore considered.

Patterns which emerged from the mapping had no obvious trends. No directional component is to be included in the first runs as there is no strong directional bias in housing or layout of residential areas owing to the flat topography of the study area. Further, the irregularity of the shape and location of the cells make it almost impossible to recognize directional patterns in cell location and therefore only an overall autocorrelation is applied although this can be expected to give a lower coefficient than that for directional runs (Ward, 1978, p.42).

v) The weighting factor

When a surface comprises irregular sub-areas (as many geographic surfaces do), Cliff & Ord (1969) suggest that the statistic should be adjusted to take this into account. As the cells in the study area are irregular in both shape and size, a weighting factor must be included so that cells are standardised relative to each other for comparative purposes. The relationship between adjacent cells is not necessarily reciprocal because of the relative importance of the degree of contact between cells. Where a large cell is bounded by smaller cells, the proportion of the perimeter of the large cell in contact with each small cell is far smaller than the contact proportion for each small cell relative to the large one. The influence of the large cell is thus greater on the small cell than vice versa. This fact, the character of the U.C.T. computer and the use of Fortran necessitated that a square weighting matrix was constructed, the size depending upon the number of cells under consideration. Where at least one side of an ESD forms part of a boundary that side is not considered as part of the cell perimeter. It is therefore presumed that 100% interaction takes place over contacting sides. As Ward (1978) set the weighting factor at unity, the computer programme had to be modified (Appendix 4), in order to include the weights. The programme was then tested on a regular and an irregular grid by a random number run with w_{ij} set to 1. Satisfied that the programme was operational a series of correlations was run on hypothetical data. The 33 cells of the northern suburbs were used as a base and interaction in an east-west direction was simulated. A set of values graded from high in the western to low in the eastern cells formed the data set for testing the effect of various weights.

Using an unweighted matrix (i.e. w_{ij} set to 1) the correlation was as expected viz. r = 0,77 (Table 4.3). Cliff & Ord (1971) emphasize the importance of using (at least initially) length or proportion of common boundary for adjacent cells for standardizing cells relative to one another. The irregular shape of the cells in the study area suggested that this might make a difference to the correlation. Proportion of the perimeter of an ESD in contact with an adjacent cell was thus used as an initial weight. The correlation dropped considerably (Table 4.3) to r = 0,177. This fourfold decrease in the coefficient, bringing it to a level that was neither significant nor indicative of any association, showed clearly how important it is to test patterning stastically and objectively and not to rely on visual interpretation as it can be misleading. On the other hand the weighting itself may be inappropriate. This needs testing on standardised and directly comparable data sets. The choice of other factors influencing the interaction between cells depends on the nature of the variable being correlated. Since this study deals with socio-economic status and since this is affected by the number of wage earners in each cell it was felt that it would be appropriate to incorporate this variable into the weighting system. The second weight chosen therefore is composite, combining proportion of perimeter in contact and number of wage earners in a cell. When this weighting factor was applied to the hypothetical data the correlation coefficient increased very slightly to 0,182 (Table 4.3 run 3).

TABLE 4.3 Effects of varying weights (w_{ij}) on hypothetical data

Run	Area	Data Input	Weighting factor	Correlation coefficient
1	North as base	Graded	Unweighted	0,760
2	North as base	Graded	Contact Proportion	0,177
3	North as Base	Graded	Contact Proportion and wage earners	0,180

Note: Number of joins in each case = 168

The weighting factor thus considerably reduces the correlation coefficient and a single factor weight produces the lowest correlation within an area. Thus while it appears when calculating the correlation coefficient without a weighting component that there is definite spatial dependence between cells, when this component is included, spatial dependence does not occur and groups of cells may be regarded as discrete. Although Cliff & Ord (1969) indicated that a weighting system should be used when applying their statistic to irregularly shaped areas because the difference in contact between cells would tend to change the correlation, they gave no indication of the order of the change that could be expected. The order of magnitude of the decrease from unweighted correlation to a single factor weighted one is exceptionally high (over 300%). An investigation into the exact operation and effects of weighting is thus necessary. This would require the skills of an applied mathematician or statistician, however, and is beyond the scope of this study which concentrates on patterns of social status.

b) Spatial dependence in social status

Correlations were now sought for each of the four residential concentrations using the sample areas defined earlier. Data input for each area was midpoint of modal class for each ESD. Once again using the northern suburbs as the cell matrix the programme was run three times. The first run was unweighted. Secondly, proportion of the perimeter was used as a weight and finally, once again the composite weighting factor of proportion of perimeter and wage earners, was applied. The correlations which are the results of these experimental runs are given in Table 4.4.

TABLE 4.4 Effects of varying weights (w_{ij}) on real data

Run	Area	Data Input	Weighting factor	Correlation Coefficient
1	North	Midpoint of modal class	Unweighted	0,171
2	North	Midpoint of modal class	Contact proportion	0,053
3	North	Midpoint of modal class	Contact Proportion and wage earners	0,063

Note: Number of joins in each case = 168

The use of real data with only a single weighting factor again shows a marked decrease in the correlation coefficients similar to that observed in the runs using hypothetical data. Testing for autocorrelation in samples from all four components of the study area therefore had to take into account both size of and interaction between cells. The weighting factor used was the combined proportion of perimeter in contact and number of wage earners for each ESD. This composite factor appears to give marginally higher correlations than the single factor weighting system. Results of the correlations are given in Table 4.5 below.

TABLE 4.5 Calculated coefficients of autocorrelation (sample values).

Area	kurtosis	No. of cells	No. of joins	Correlation coefficient
North	3,2	33	168	0,063
South	3,2	37	186	0,012
Central	1,4	35	162	0,069
East	5,0	36	174	0,070

The data for the sample in each area may be regarded as normally distributed. As expected from the sample size and nature, values obtained for the coefficient of kurtosis indicate that while data for the northern and southern suburbs are quite normally distributed, those for the central and eastern areas are respectively slightly positively and negatively skew. Except for the Southern suburbs which shows an exceptionally low value for the correlation coefficient the four areas exhibit almost identical, low values. Since the coefficient here is a measure of the overall spatial dependence in the area under consideration, it must be concluded that, using modal class midpoint as data input values and a combination of proportion of perimeter in contact and number of wage earners in the ESD as the weighting factor, overall spatial dependence does not exist. That is, given that a cell has residents with a particular modal class of SES score there is no guarantee of the modal class SES score to be found in adjacent ESD's: they may be lower, higher or the same. It was therefore decided to test whether the apparent patterns in overall cluster groups were spatially dependent. The highly contiguous groups of cells isolated by the 70% level cut-off on the overall cluster diagram for the northern suburbs was the data input. The correlation coefficients (given in Table 4.6) are lower than those for a data input of midpoint of modal class. The higher the contiguity between cells of the same socio-economic status therefore, the lower the correlation and the less the spatial dependence.

TABLE 4.6 Calculated coefficients of autocorrelation (overall clusters)

Run	Area	Data Input	Weighting factor	Correlation coefficient
1	North	Overall cluster groups	Contact proportion	0,040
2.	North	Overall cluster groups	Contact proportion and wage earners	0,043

The expectation of finding a high degree of spatial dependence in view of the high degree of contiguity of cells of the same modal class in each of the residential concentrations is thus not fulfilled. Some explanation is offered by consideration of two facts. An even lower correlation coefficient is obtained using data which shows a higher degree of contiguity than midpoint of modal class, viz. 70% level cluster groups (Table 4.5 and 4.6). Clearly the higher the degree of contiguity of cells of the same value, the lower the overall spatial dependence. Although groups of cells of the same value occur in each of the areas it is the breaks or differences between these groups of cells which are emphasized. These tend to lower the overall spatial dependence in any area. By implication therefore, low or no correlation coefficients point to the existence of small groups of contiguous ESD's of the same social status. Given that a particular status occurs in one area however, there is no guarantee of the status which will be found in an adjacent area.

Spatial autocorrelation appears to be an extremely effective technique for measuring spatial dependence. Visually, patterns of socio-economic status suggested that spatial dependence between groups of ESD's of the same status would be high, as contiguity between ESD's in each status group was high. It is however, the spatial dependence between adjacent groups which must be measured. On account of the irregular shape of the system of sub-areas, the correlation was expected to be moderate but nevertheless indicative of definite dependence between groups. The calculated value for correlation was however, extremely low and thus showed a complete lack of spatial dependence within the study area.

The nature of the sub-area system also made the consideration of the weighting factor and weighting systems imperative. The primary investigation of this factor and its mode of operation constitutes a contribution to the field. When real data input is unweighted, the correlation coefficient is low, but weighted values reduce the correlation even further. A single factor weighting matrix (proportion of the perimeter in contact with adjacent cell) gives a lower correlation than does a complex weighting system (proportion of perimeter in contact and number of wage earners in a cell). The investigation of the true nature of the weights and their operation in affecting the correlation coefficient, it was felt, required the skills of an applied mathematician or statistician.

2 Evaluation of patterns of social differentiation

When socio-economic status scores for all race groups are considered for Greater Cape Town, the non-white population tends to appear as a single low socio-economic status group, although some measure of differentiation may be seen for the whites. SES scores for the coloured population only, showed distinct socio-economic differentiation (i.e. social areas of differing socio-economic status could be identified) in social space. Results of the application to data of the Cliff & Ord (1969) auto-correlation coefficient indicate that there is no spatial dependence between patterns of differing socio-economic status which emerge as a result of mapping socio-economic status variables (either descriptive statistics or calculated values). Lack of spatial dependence between adjacent groups of cells of the same socio-economic status however, does not deny the existence of distinct small pockets of ESD's of the same socio-economic status within the coloured residential areas. In fact, it shows up the lack of spatial dependence between individual pockets rather than identifying a situation in which the status of each cell is likely to be affected by the status of its immediately adjacent cells, i.e. the western norm. The investigation has shown that in this case areas which have recognised characteristics can be identified but these are limited in extent and may be distinctly different from their surroundings. An investigation into the implications of these patterns is therefore required.

The first question which may be raised is whether the patterns are associated with any officially delimited or conceptualized areas such as suburbs or neighbourhoods. Herbert (1972, p.226) gives "the classic description of a neighbourhood" as "an area in which the residents are personally well acquainted with each other, and are in the habit of visiting one another or exchanging articles and services and, in general, of doing things together". He does however, go on to mention some qualifications which must be made regarding this definition, namely that the concept of "neighbourhood" differs for the population concerned depending on differences between age groups; between locals and non-locals; between social classes; as well as the effect of increasing mobility over time and also the scale at which the neighbourhood is conceived, i.e. area, configuration and population size (Herbert, 1972). On the basis of such characteristics, the groups of ESD's identified in the study area cannot be considered as neighbourhoods. Information is not available on the required scale as the average population size of ESD's in the study area (approximately 900) precludes interaction among all individuals of an ESD; even individual ESD's are in many cases larger than neighbourhoods.

The concept of "suburb", like that of "neighbourhood" has social implications. A suburb is a residential area possessed of a greater or smaller number of facilities or services, and in which certain social characteristics are expressed in various forms. These characteristics usually include such tangibles as housing type and other status symbols e.g. car, television; as well as the intangibles such as lifestyle, standard of living and aspirations (Harvey, 1975). Generally, people of similar social status locate in areas which reflect the characteristics to which they aspire. Populations thus tend to identify with the character of a particular suburb and to attempt to preserve this. In the entire study area, seven suburbs, generally the smaller ones but representative of 25% of the total, correspond with ESD's grouped together on the basis of modal class. When the results of other grouping techniques, i.e. clusters are mapped, the association between a particular social status and the smaller suburbs is evident in all residential concentrations except the northern suburbs.

In the southern residential concentration the adjacent suburbs of Heathfield and Southfield (Figure 1) together comprise five ESD's of modal class 70 - 79. This presents a marked contrast to the surrounding area and both the social area and the suburbs form a distinct island. In direct contrast the three ESD's of Lavendar Hill fall into modal class 30 - 39 and are of lower social status than the surrounding ESD's. Five separate suburbs, Lansdowne, Wetton, Hanover Park and Vanguard in the central concentration each consist of a group of ESD's of modal class 70 - 79. Again these are in marked contrast to the surrounding area. Kalksteenfontein (which consists of only two ESD's) has a modal class of 40 - 49. Finally, Matroosfontein in the eastern suburbs forms a block of five contiguous ESD's of modal class 50 - 59. The latter two suburbs are not of distinctly different socio-economic status from the surrounding ESD's. They are nevertheless all of the same status and are generally at least one percentile higher than their adjacent cells. Social areas defined for coloureds can thus find expression in officially designated areas. Although it is possible to identify different areas, there is a need for a study of attitudes, physical status characteristics and access to facilities in the different areas in order to assess the contrast between upper and lower status groups. As patterns of differentiation among the coloureds can definitely be recognised in social space and may also be reflected in physical space, the hypothesis that in a well defined ethnic group within a society, social differences parallelling those of the core group may be identified (Timms, 1971) tends to be at least partly substantiated.

Where contiguous cells are not necessarily contained by the boundaries of suburbs, they may stretch into adjacent suburbs but not cover them completely, e.g. from the Lansdowne group into the Athlone area. In other cases, e.g. Bonteheuwel, Grassy Park and Manenberg, a suburb may contain several different groups. For the most part then, suburbs and groups of contiguous ESD's are not coincident. As distinct areas of the same socio-economic status can be identified and are in some cases associated with one another, this may be a useful base on which to create a unity within the coloured group and to involve the people with their own environment.

In more general terms patterns which result from the mapping of either descriptive statistics or calculated values appear to show a decrease in scores from west to east in each residential concentration. This trend appears in the mapping of means, modal class values or of cluster values for each residential concentration at the 70% level of similarity and even with overall cluster mapping. The coloured

Group Areas are legally defined and are officially considered to be completely separate from their surrounding areas. Their boundaries were therefore considered as "real" for the purpose of this study. Seen in the light of interaction with and effect on the status of adjacent coloured residential areas however, the boundaries cannot be considered as such. Contact with white, industrial or farming areas appears to have some influence on the social status in adjacent coloured residential areas. In the southern suburbs, average SES scores are higher on the convex or northwestern side of the area, (which is adjacent to a white residential zone, see Fig. 3.6) than in the south eastern section where the adjacent area is one of small-holdings and farms. Modal class and cluster mapping show plateaux of medium to high social status in the same areas (see Figures 3.11, 3.12, 3.21, 3.26 & 3.27). The central suburbs show a similar trend (Figure 3.4). Average SES scores in the western section of the area, bordering on white residential suburbs are high, and decrease eastwards where the area is bordered by farmland and undeveloped property and other coloured residential areas. Modal class patterns and cluster mapping (Figures 3.9, 3.12, 3.19, 3.24 & 3.27) again confirm these trends as a plateau of high to medium status occurs in the west and status then generally decreases eastwards. The eastern suburbs surrounded by industrial land or open space, show a trend ofgenerally-low-average-SES-scores (Figure 3.5) which is confirmed yet again by modal class and cluster mapping (Figures 3.10, 3.12, 3.20, 3.25 & 3.27). The further postulate by Timms (1971) that residential areas of the ethnic group adjacent to residential areas of the core group show higher status than those non-adjacent, is clearly borne out.

It may be concluded that the coloureds do not form a single low status group, but rather a sub group within which social stratification has taken place. This is comparable with Abu Lughod's (1969) findings of social stratification within a non-western city, i.e. Cairo, which clearly mirrored that found in western society but had no spatial dimension. Social areas among the coloured population of Greater Cape Town have been identified by means of socio-economic status scores.

No body of theory exists as a framework either for devising or assessing the representativeness of a social indicator of this type. However it was decided that as social indicators should be seen as tools to be employed for describing and/or evaluating the social state or conditions of a society, they will be defined in different ways, depending on the requirements of the researcher and the specific problem. In selecting SES scores the criteria for the formulation of social indicators were met and the indicators and specifications for their use were defined.

The use of SES scores for this study overcomes several problems related to definition and use of social indicators. Since the scores have been derived separately for whites, coloureds and Asians as well as for all race groups combined, the coloured population can be seen as an entity rather than a single stratum, generally low socio-economic group. Secondly, the scores are a combination of statistics on income, education and occupation. They are thus composite indicators and overcome the inherent bias of a single factor indicator. Thirdly, the manner in which they are derived ensures a normal distribution of individuals in the population over the range of possible scores. The three indexes are available from the census data and are thus easily quantifiable. These SES scores thus conform to requirements for acceptable and valid social indicators. Education, occupation and income are also generally regarded as valid indicators both individually and in combination (Smith, 1973; Knox, 1975). Finally, SES scores are the best available indicators on which to assess social differentiation within the coloured population. Other indicators available (residential or family status indicators) were regarded as inappropriate. In the case of the former, information on housing was inadequate and the acute housing shortage gives rise to artificially high population densities; while for the latter the ceiling on incomes for coloureds, job reservation and limited availability and access to education make these indicators meaningful only when seen in combination with some and related to all of the previously mentioned variables.

A point of concern with regard to the use of SES scores as indicators, however, is the fact that the three indexes are closely related and inextricably linked. This fact leads to the consideration of whether SES scores are therefore truly socio-economic indicators or merely economic ones. Stemming from this is the realization that SES scores do not reveal much about social evaluation among the coloured population group. SES scores are derived on the basis of white evaluation of social status. There is no indication of what the coloured evaluation of social status is. An example: a coloured male who is employed as a cleaner may enjoy a high social status because his wife is a qualified nursing sister. This suggests a different evaluation of social status from the white. Further study of social attitudes among the coloured population is therefore called for as this is not taken into consideration by SES scores (the best available social indicators) and therefore by this study. Although this may be considered as a shortcoming of the research, since obviously patterns of social differentiation will be influenced by it, derivation of a more representative indicator was not within the scope of the research. The aim of the project was to assess patterns of social differentiation among the coloureds using the best available indicator, i.e, SES scores.

This study therefore points to the direction which should be taken for a contribution to social theory in a multi-racial society (viz. the derivation of an essentially socio-economic or social indicator reflecting social attitudes and evaluation from the coloured point of view) and the processes which lead to social differentiation among these people. A possible course of action here would be the application of multi-variate or factor analysis to generate factors which might be of assistance in recognising valid social indicators and point to social processes operative in distinguishing social classes.

The identification of patterns of differing socio-economic status in the coloured residential areas of Greater Cape Town begs the question of their relation to the processes by which differentiation has come about. By means of synthesis of information the extent to which pattern might suggest process can be assessed, i.e. the links of cause and effect which lead things to be as they are may be established (Harvey, 1973b; Smith, 1975). In assessing the extent to which pattern reflects process in social differentiation among the coloureds, the first question which must be raised is that of the similarity of aspirations of coloureds as opposed to whites. Secondly, the social processes which have led to differentiation must be examined. Although the coloureds are generally believed to have similar aspirations and value systems to those of whites (Theron Commission Report, 1976), no comprehensive study or analysis of coloured attitudes and aspirations has been undertaken. The question of similarity to whites (and therefore advanced western society) remains open. Factors influencing the selective processes in social differentiation among coloureds are also relatively unknown.

In order to bridge this gap, an investigation of the past patterns of social differentiation among the coloureds in Greater Cape Town has been made and this has helped to explain the nature of processes which have led to the development of these patterns and cast some light on those operational at the present time. The inquiry has revealed that from the time of Van Riebeeck to the implementation of the Group Areas Act, social differentiation occurred among the coloured population of Greater Cape Town and was largely associated with economic differences. Differences in housing type reflecting these economic and social differences were recognisable too. It is therefore postulated that over and above the factors which divide white and coloured (in fact all race groups) in Greater Cape Town among the coloureds themselves, stratification takes place primarily on the basis of economic differences. Use of SES scores is justified on this basis due to their strong reflection of the economic state of the population.

Social theory (Tumin, 1967) and the theories of city structure (Berry & Horton, 1970) have shown that in advanced western society differences in social space are easily identifiable in physical space. Although residential differentiation is according to Shevky and Williams, (1949) the most easily accessible and visual means of discerning social differences in a society, this may not pertain in the case of non-western societies as Abu Lughod (1969) found in her study of Cairo. Social differentiation was identified in the society, but was not recognisable in differences in type and/or quality of housing. While groups of ESD's identified in social space (indicative of differences in social status among the coloureds in Greater Cape Town) can be translated into physical space by virtue of their origin in census data, this social differentiation is not reflected in housing differences. As in the case of the non-western societies then, social differences among the coloureds may be identified in social space but are not recognisable in physical space.

The lack of a physical dimension to social stratification among the coloureds is a phenomenon most clearly recognisable in the time period from the implementation of the Group Areas Act to the present. As it is in this period that apartheid has been practised, the influences of the system, i.e. job reservation, wage ceilings, residential restriction due to Group Areas legislation and the housing shortage on social differentiation and its manifestation in physical space must be assessed. According to Abu Lughod (1969) if city structure is to be a reflection of social structure the effective ranking system in a city must be related to the operational definition of social status and the ranking system in a city has to be identifiable by the area units of observation used in analysis. To obtain clear distinctions between stages in the family cycle, each stage must be associated with changes of residence. This is patently not the case in the situation of the coloureds in Greater Cape Town. The reasons postulated for this are the shortage of, and lack of choice in housing for coloureds. Families or individuals who cannot afford to purchase their own houses are forced to live in council rented accommodation. Since council rented dwellings are allocated on the basis of a maximum wage of the head of household, this has several implications. Firstly, false incomes may be given in order to secure a dwelling which is in financial reach of a household (i.e. if better accommodation were available this would be chosen, but home ownership, the only other alternative is not within reach). Secondly, it could mean that even though the earnings of the head of household may be within the limits set for eligibility for council housing, there may be a number of earners in the household (as they cannot obtain separate accommodation due to the housing shortage) which pushes up the amount of money available for acquiring higher social status (i.e. education, or physical symbols of social status, e.g. television, car, appliances). Thirdly, it might imply that because of the lack of choice in housing as well as financial losses incurred, the Group Areas removals and resettlements have caused a discrepancy between social status and housing type. Further possibility of removal to housing compatible with social status once financial losses had been recouped over some years is low due to the lack of availability of alternate or slightly improved housing.

The conclusion to be drawn from this then, is that social status may well be much higher or at least different in coloured areas than types of dwellings in these areas would suggest. When the relationship between social areas and suburbs was considered, the majority of grouped ESD's which coincided with suburbs, particularly in the Athlone area were seen to be areas of council owned or rented accommodation. The status of these areas was thus not accurately reflected in housing. In the case of Matroosfontein, although this was originally a council letting scheme, the area became an area of home ownership. The housing here again did not reflect the true socio-economic status of the residents although this was more moderate than that of the Athlone area. Further, the identification (from mapping of socio-economic status information) of discrete social areas and a lack of spatial dependence between them is readily explicable in terms of the role of municipal, divisional council and state housing schemes. Location of council housing and the allocation explain the production of these forms (Castells, 1977). To do so it would have to be related to the rest of the elements structuring the form and rhythms of an urban area. If the social area concept is to be relevant in the South African context therefore, it must be separated into its social and physical components.

This research represents an attempt not only to recognise patterns of socio-economic differentiation and to relate these to residential differentiation in the coloured residential areas of Greater Cape Town, but also to identify the problems associated with the derivation and analysis of such patterns. Suggestions have been put forward for the directions in future research which might lead to the development of more valid social indicators and therefore more accurate patterns of social differentiation in the hopes that these may better explain processes of social stratification. Patterns of social differentiation identified and partly explained within the framework of this research are merely a baseline for more intensive and analytical research on patterns of differentiation among the coloured population group of Greater Cape Town under the apartheid regime.

REFERENCES

ABU LUGHOD, J., 1969: Testing the Theory of Social Area Analysis: The Ecology of Cairo, Egypt. *American Sociological Review*, 34 2 pp.198-211.

AGOCS, C., 1970: Social Indicators: Selected Readings. *The Annals of the American Academy of Political and Social Science.* 388 March pp.127-132.

BAUER, R.A., 1966: Detection and Anticipation of Impact: The Nature of the Task, In: Bauer, R.A., (ed.), 1966: *Social Indicators*, M.I.T. Press, Cambridge, Mass., pp.1-68.

BELL, W., 1955: Economic, Family and Ethnic Status: an empirical test. *Am. Social. Rev.* 20 pp.45-52.

BERRY, B.J.L., 1958: A note concerning methods of classification, *A.A.A.G.* 48 pp.300-303.

BERRY, B.J.L. & HORTON, F.E., 1970: *Geographic Perspectives on Urban Systems*: with integrated readings, Prentice Hall Inc., Englewood Cliffs, New Jersey, 564p.

BERRY, B.J.L. & KASARDA, J.D., 1977: *Contemporary Urban Ecology*, Macmillan Publ. Co., New York, 497p.

BERRY, B.J.L. & REES, P.H., 1969: The Factorial Ecology of Calcutta, *J. Am. Social.* 74. pp.447-91.

BERRY, B.J.L. & SPODEK, H., 1968: Urban Ecology of Shriveport Louisianna, Unpublished paper, Centre for Urban Studies, University of Chicago.

BIEDERMAN, A.D., 1966: Social Indicators and goals. In: Bauer, R.A., (Ed.), 1966, ***Social Indicators***, M.I.T. Press, Cambridge, Mass. pp.68-153.

BRIXHORN, H. & MINDLIN, A., 1972: Composite Social Indicators for small areas - Methodology and Results in Washington, D.C. In: ***Social Indicators for Small Areas.*** Papers Presented at the Conference on Small Area Statistics. American Statistical Assn. Montreal, Canada, 1972. pp.3-17.

BOESEKEN, A.J., 1977: ***Slaves and Free Blacks at the Cape 1658 - 1700,*** Tafelberg Publishers Limited, Cape Town, p.208.

BRAND, C.M., 1966: The Indian Community of the Cape Peninsula. Unpublished M.A. Thesis. Dept of Sociology, Univ. of Stellenbosch.

BRANDEL-SYRIER, M., 1971: ***Reeftown Elite: A Study in Social Mobility in a Modern African Community on the Reef.*** Routledge & Kegan Paul, London, 335p.

BROWN, M.B., (Ed.), 1977: BMDP - 77 ***Biomedical Computer programs P - series.*** Univ. of California Press, Berkley, 880p.

BUREAU OF MARKET RESEARCH, 1976: Regional Population Estimates for 1975, research report, No. 51. UNISA.

CAPE OF GOOD HOPE. Blue Book, 1891.

CAPE OF GOOD HOPE. Blue Book, 1904.

CARLISLE, E., 1972: The Conceptual Structure of Social Indicators. In: Shonfield, A. & Shaw, S., 1972. ***Social Indicators and Social Policy.*** Heineman Educational Books, London. pp.23-33.

CASTELLS, M., 1977: ***The Urban Question?*** A Marxist approach. Translated by Alan Sheridan. Edward Arnold, London, 502p.

CAZES, B., 1972: The Development of Social Indicators: a survey. In: Shonfield, A. & Shaw, S., 1972: *Social Indicators and social policy.* Heineman Educational Books, London, pp.9-23.

CITY OF CAPE TOWN, 1977: Council Housing in Cape Town, City of Cape Town, Town Clerk's Department Housing branch.

CLIFF, A.D. & ORD, J.K., 1969: The Problem of Spatial Autocorrelation. In: Scott, A.J., (Ed.), 1969. *Studies in Regional Science.* Pion, London. pp.25-55.

CLIFF, A.D. & ORD, J.K., 1971: Evaluating the percentage points of a Spatial Autocorrelation Coefficient. *Geog. Analysis.* 3 1 pp.51-62.

CLIFF, A.D. & ORD, J.K., 1973: *Spatial Autocorrelation.* Pion, London. 178p.

CLIFF, A.D., HAGGETT, P., ORD, J.K., BASSETT, K. & DAVIES, R., 1975: *Elements of Spatial Structure a Quantitative Approach.* Cambridge, Univ. Press. London, pp.145-180.

CLIFFORD, H.T. & STEPHENSON, W., 1975: *An introduction to Numerical Classification,* Academic Press, New York, 1975, 229p.

DACEY, M.F., 1963: Order neighbour statistics for a class of random patterns in multi-dimensional space. **A.A.A.G.,** 53 pp.505-15.

DACEY, M.F., 1968: A Review on Measures of Contiguity for two and K-Color maps. in: Berry, B.J.L. & Marble, D.F., 1968: *Statistical Analysis,* A reader in statistical geography. Prentice Hall Inc., Englewood Cliffs., New Jersey, pp.479-490.

DAVENPORT, T.R.H., 1977: *South Africa a Modern History.* Macmillan, Johannesburg, 2nd Edition, 432p.

DAVIES, R.J., 1963: The Growth of the Durban Metropolitan Area. *S.A.G.J.* XLV pp.15-43.

DAVIES, R.J., 1976: *Of Cities and Societies: a Geographer's Viewpoint.* Inaugural Lecture, New Series No. 38. University of Cape Town. 20p.

DE VOS, S., 1973: The Use of Nearest Neighbour Methods. *T.E.S.G.* 64 5 pp.307-19.

DIEDERIKS, R., 1979: *Saclant Graphics Package User Manual,* U.C.T. Computing service. 64p.

DIVISIONAL COUNCIL OF THE CAPE, 1978: Coloured Housing progress report. Circular of the Divisional Council of the Cape.

DUNCAN, O.D., 1967: Discriminating Against Negros. *The Annals of the American Academy of Political and Social Science,* 371, May, pp.85-103.

ELLIS, G., HENDRIE, D., KOOY, A. & MAREE, J., 1977: *The Squatter problem in the Western Cape.* Some causes and remedies. South African Institute of Race Relations, Johannesburg. 119p.

ELPHIC, R. & GILIOMEE, H., (Eds.), 1979: *The Shaping of South African Society 1652 - 1820.* Longman Penguin, Cape Town, 415p.

ETZIONE, A. & LEHMAN, E.W., 1967: Some Dangers in "Valid" Social Measurement. In: *The Annals of the American Academy of Political and Social Science.* 373, September, pp.1-15. Or in: Gross, B.M. (Ed.), 1969: *Social Intelligence for America's Future.* Allan & Bacon, Boston, pp.45-62.

FAIR, T.J.D., 1969: Southern Africa: Bonds and Barriers in a multiracial region. In: Prothero, R.M. (Ed.), 1969: *A Geography of Africa.* Routledge and Kegan Paul, London, pp.325-379.

FERRARIO, F.F., 1979: ***The Cape Town of the Coloured People.*** An excursion organized for the G_{79} Conference of the S.A.G.S., U.C.T. 1979. 20p.

FIELD, J.G., 1969: The use of the information statistic in the numerical classification of heterogeneous systems. ***J. Ecol.*** 57 pp.565-69.

FORREST, J., 1974: Spatial Aspects of Urban Social Travel. ***Urban Studies*** 11 301-313.

GARN, H.A. & FLAX, M.J., 1972: Indicators and statistics: Issues in the generation and use of indicators. In: ***Social Indicators for Small Areas.*** Papers presented at the Conference on Small-Area Statistics. American Statistical Assn. Montreal, Canada. 1972. pp.37-41.

GEARY, R.C., 1968: The Contiguity Ratio and Statistical Mapping. In: Berry, B.J.L. & MARBLE, D.F. (Eds.) 1968: ***Spatial Analysis,*** Prentice Hall, Englewood Cliffs, New Jersey, pp.461-478.

GORDON, I.R. & WHITTAKER, R.M., 1972: Indicators of Local Prosperity in the South West Region, ***Regional Studies,*** 6 pp.299-313.

GREGORY, S., 1973: ***Statistical Methods and the Geographer*** (3rd Edition) Longman, London, 271p.

GROSS, B.M., 1966: The State of the Nation: Social Systems Accounting. In: Bauer, R.A., (Ed.), 1966: ***Social Indicators,*** M.I.T. Press. Cambridge Mass. pp.154-271.

GROSS, B.M. & SPRINGER, M., 1969: Developing Social Intelligence In: Gross, B.M. (Ed.) 1969: ***Social Intelligence for America's Future: Explorations in Societal Problems*** Boston: Allyn & Bacon, 541p.

HART, G.H.T., 1968: An introduction to the Anatomy of Johannesburg's Southern Suburbs. ***S.A.G.J.*** 50 pp.65-72.

HART, R.A., 1973: Economic Expectations and the decision to migrate. An analysis by Socio-economic group. ***Regional Studies*, 7** pp.271-285.

HART, T., 1975: ***The Factorial Ecology of Johannesburg.*** Occasional paper No. 5, Urban and Regional Research Unit, University of the Witwatersrand, Johannesburg, March 1975. 96p.

HART, T., 1976: The Evolving Pattern of Elite White Residential Areas in Johannesburg 1911-1970. ***S.A.G.J.*** 58 1 pp.68-75.

HARVEY, D., 1973a: ***Social Justice and the City.*** Edward Arnold, London, 336p.

HARVEY, D., 1973b: ***Explanation in Geography.*** Edward Arnold, London, 521p.

HARVEY, D., 1975: Class Structure in a Capitalist Society and the theory of Residential Differentiation. In: Peel, R., Chisholm, M., Haggett, P., 1975: ***Processes in Physical and Human Geography,*** Bristol Essays, Heineman Educational Press, London, p.354-369.

HERBERT, D., 1972: ***Urban Geography.*** A social perspective. David & Charles, Newton Abbot. 319p.

HEY, P.D., 1961: ***The Rise of the Natal Indian Elite,*** Natal Witness, Pietermaritzburg.

HOUGHTON, D.H. & DAGUT, J., 1972: ***Source Material on the South African Economy 1860 - 1970.*** Vol. 1, 1860 - 1899, OUP, Cape Town, 371p.

HOUGHTON, D.H. & DAGUT, J., 1972: ***Source Material on the South African Economy 1860 - 1970.*** Vol. 2. 1899 - 1919, OUP, Cape Town, 247p.

HOUGHTON, D.H. & DAGUT, J., 1973: *Source Material on the South African Economy 1860 - 1970.* Vol. 3, 1919-1970, OUP, Cape Town, 263p.

JOHNSTON, R.J., 1968: Choice in Classification: The subjectivity of objective methods. *A.A.A.G.* 58 pp.275-89.

KAHL, J.A. & DAVIS, J., 1955: A Comparison of Indexes of socio-economic status. *American Sociological Review,* 20 3 pp.317-325.

KAMRANY, N.M. & CHRISTAKIS, A.N., 1970: Social indicators in perspective. *Socio-Econ. Plan. Sci.,* 4 pp.207-216.

KNOX, P.L., 1975: *Social Well Being: A Spatial Perspective.* OUP, London, 60p.

KNOX, P.L., 1976: Social Well Being and North Sea Oil: An Application of Subjective Social Indicators. *Regional Studies, 10* pp.423-432.

KUPER, L., 1965: *An African Bourgeoisie: Race, Class & Politics in South Africa.* Yale Univ. Press, New Haven, pp.118-139.

LAIDLER, P.W., 1939: *The Growth and Government of Cape Town.* Unie Volkspers Bpk, Cape Town, 525p.

LAIDLER, P.W., 1952: *A Tavern of the Ocean* (2nd ed.). Maskew Miller, Cape Town. 222p.

LANCE, G.N. & WILLIAMS, W.T., 1967: A General Theory of Classificatory sorting strategies, I Hierarchical Systems. *Computer Journal,* 9 pp.373-80.

LAND, K.C., 1971: On the Definition of Social Indicators, *The American Sociologist,* 6 (November) pp.322-325.

LOUW, R., 1979: Density and Income: A Case Study. Geography III Project report (unpublished) University of Cape Town, Geography Department, 32p.

M.A.C.C., 1976: ***UNISTAT 2: Univariate descriptive statistics and histograms.*** Univ. of Wisconsin-Madison Computing Centre, Madison, 65p.

MARAIS, J.S., 1957: ***The Cape Coloured People 1652 - 1937.*** Witwatersrand University Press, Johannesburg. 296p.

MEER, F., 1969: ***Portrait of Indian South Africans.*** Avon House, Durban, pp.60-134.

MENTZEL, O.F., 1921: ***A Geographical and Topographical Description of the Cape of Good Hope,*** Van Riebeeck Society, Cape Town, Part 1, Trans.: Mandelbrote, H.J., 181p.

MILLS, E.S., 1972: Welfare Aspects of National Policy towards City sizes. ***Urban Studies,*** 9 pp.117-124.

MOLL, E.J., CAMPBELL, B.M. & PROBYN, T.A., 1976: A rapid statistical method of habitat classification using structural and physiognomic characteristics. ***S.A. Journ. of Wilflife Research*** 6 1 pp.45-50.

MORAN, P.A.P., 1950: Notes on Continuous Stochastic Phenomena. ***Biometrica,*** 37 pp.17-23.

MOSER, C., 1970: Measuring the Quality of Life. ***New Society,*** 10 December 1970, No. 428 pp.1042-1043.

NADER, G.A., 1969: Socio-economic Status and Consumer Behaviour. ***Urban Studies,*** 6 pp.235-245.

OLSON, J.M., 1975: ***Autocorrelation and visual map complexity, A.A.A.G.,*** 65 pp.189-204.

PERLE, E.D., 1970: Editor's Introduction *Urban Affairs Quarterly* 6 2 pp.135-143.

PHILLIPS, B.D., 1971: *The Coloured Population of the Port Elizabeth/ Uitenhage Region: A socio-economic study.* Institute for Planning Research, U.P.E. Research Paper, No. 7, Port Elizabeth.

PLESSAS, D.J. & FEIN, R., 1972: An Evaluation of Social Indicators *Jnl. of the Amer. Inst. of Planners.* 38 1 pp.43-51.

POLLOCK, N.C. & AGNEW, S., 1963: *An Historical Geography of Cape Town*; Longmans, London. 242p.

PUTTERILL, M.S. & BLOCH, C., 1978: *Providing for leisure for the city dweller.* UPRU, (Urban Problems Research Unit) U.C.T. 141p.

REES, P.H., 1972: Problems of classifying subareas within cities. In: Berry, B.J.L. (Ed.), 1972: *City Classification Handbook.* Wiley & Sons Inc. New York, pp.265-330.

R.S.A. 1976: Commission of Inquiry into matters relating to the coloured population group (Theron, E., Chairman).

R.S.A. Dept of Statistics, Population of South Africa 1904 - 1970. No. 02-05-12.

R.S.A. Bureau of Statistics, Population Census, 1960, Vol. 7, No. 2.

R.S.A. Dept of Statistics, Population Census, 1970. No. 02-01-07.

R.S.A. Dept of Statistics, Population Census, 1970. No. 02-05-06.

S.A.I.R.R., Survey of Race Relations in South Africa, 1950 - 1978. Johannesburg.

SCOTT, P., 1955: Cape Town: A Multi-Racial City, *The Geographical Journal CXX 2* pp.149-157.

SHELDON, E.B. & FREEMAN, H.B., 1970: Notes on Social Indicators: Promises and Potential. *Policy Sciences*, 1 pp.97-111.

SHEVKY, E. & BELL, W., 1955: *Social Area Analysis: Theory Illustrative Application and Computational Procedures.* Stanford University Press, Stanford.

SHEVKY, E. & WILLIAMS, M., 1949: *The Social Areas of Los Angeles.* California University Press, Berkley, 172p.

SHONFIELD, A. & SHAW, S., 1972: *Social Indicators and Social Policy*, Heineman Educational Books, London, 151p.

SHORTEN, J.R., 1963: *Cape Town*, Shorten & Smith, Cape Town, 654p.

SILK, J., 1979. *Statistical Concepts in Geography.* George Allen & Unwin, London, 276p.

SIMKINS, C., 1978: Socio Economic characteristics of 16 squatter settlements in Cape Town Area in 1975. *Saldru Working Paper*, No. 21, Cape Town. 43p.

SMITH, D.M., 1973: *The Geography of Social Well Being in the United States.* McGraw Hill Book Co., New York, 144p.

SMITH, D.M., 1975: *Patterns in Human Geography.* An introduction to numerical methods. Crane Russak & Co. Inc., New York, 373p.

SMITH, D.M., 1979: The identification of problems in cities: applications of social indicators. In: Herbert, D.T., & Smith, D.M. (Eds.), 1979: *Social problems and the city: Geographical Perspectives.* Oxford University Press, Oxford, 271p.

SPENCE, N.A., 1968: A Multifactor Uniform Regionalization of British Counties on the basis of employment data for 1961. *Regional Studies*, 2 pp.87-104.

SPENCE, N.A. & TAYLOR, P.J., 1970: Quantitative Methods in regional taxonomy. In: Board, C., Chorley, R.J., Haggett, P. & Stoddart, D.R. (Eds.), 1970: *Progress in Geography, 2* pp.1-65.

SPRINGER, M., 1970: Social Indicators, Reports and Accounts: Toward the management of society. *The Annals of the American Academy of Political and Social Science, 388.* Political Intelligence for America's Future, March pp.1-14.

STONE, G.W., 1972: Identity Among Lower Class Cape Coloureds. In: Whisson, M.G. & Van der Merwe, H.W. (Eds.), 1972: *Coloured Citizenship in South Africa.* Report of the second workshop 1972. Abe Bailey Inst. for Interracial Studies, U.C.T., pp.28-47.

THOMAS, W.H., 1976: Socio Economic Development of the Coloured People. In: Van der Merwe, H.W. & Groenewald, C.J., 1976 (Eds.): *Occupational & Social Change among Coloured People in South Africa.* Proceedings of a Workshop of the Centre for Intergroup Studies at UCT. Juta & Co., Cape Town, pp.48-75.

TIMMS, D., 1971: *The Urban Mosaic.* Towards a theory of Residential Differentiation. Cambridge University Press, Cambridge, 277p.

T.M.S., 1976: *Average Family Income in Greater Cape Town.* Technical Management Services, City Engineer's Department, Cape Town. Data Abstract 1 - 76, 30p.

T.M.S., 1976: *Types of Dwelling in Greater Cape Town.* Technical Management Services, City Engineer's Department, Cape Town, Data Abstract 2 - 76, 41p.

T.M.S., 1976: *Socio-Economic Status Indices,* Technical Management Services, City Engineer's Department, Cape Town, Working Paper 6 - 76, 43p.

T.M.S., 1977: *Cape Urban Area Population by Suburb 1970,* Technical Management Services, City Engineer's Department, Cape Town, Data Abstract, 31p.

TUMIN, M.M., 1967: *Social Stratification.* The forms and functions of Inequality. Prentice Hall, Englewood Cliffs, New Jersey, 118p.

VAN ARDSOL, M., CAMMILLERI, S.F. & SCHMIDT, C.F., 1958: The generality of Urban Social Area Indices. *American Sociological Review* XXIII 2 pp.277-84.

VAN DER MERWE, H.W., 1962: Social Stratification in a Cape Coloured Community. *Sociology and Social Research,* 46 3 pp.302-311.

* WALFORD, R., (Ed.), 1973: *New Directions in Geography Teaching.* Longman, London, 197p.

WARD, G.J., 1978: Spatial Autocorrelation and the analysis of Patterns Resulting from Crime Occurrence. Unpublished M.A. Thesis, Rhodes University, Grahamstown. 129p.

WATTS, H.L., 1971: *The Indian community of Wyebank: A socio-economic survey in the borough of Kloof, Natal.* Institute for Social Research, Univ. of Natal, Durban, 19p.

WESTERN, J.C., 1976: Knowing one's place: Group Areas Removals in Cape Town. Unpublished PhD thesis, University of California, Los Angeles, 688p.

WILKINS, W.H., (Ed.), 1910: Barnard, Lady Anne: *South Africa a Century Ago* (Letters written from the Cape of Good Hope 1797-1801). Maskew Miller, Cape Town, 5th ed. 316p.

WILLIAMS, W.T., 1971: Principles of Clustering. *Annual Review of Ecology and Systematics,* 2 pp.303-26.

WILSON, M. & MAFEJE, A., 1963: *Langa: A Study of Social Groups in an African Township.* OUP, Cape Town, 190p.

* VAN DER MERWE, I.J. & ZIETSMAN, H.L., 1977: *Sociale Atlas van die Kaapse Metropolitaanse Gebied*

Instituut vir Kartografie, Universiteit van Stellenbosch, Publikasie Nr.2/1977.

WILSON, A.G. & REES, P.H., 1979: A note on the calculation of a socio-economic status index and the estimation of income distributions for residential areas from occupation data. Preprint, Department of Geography, University of Leeds, 8p.

WILSON, M. & THOMPSON, L. (Ed.), 1969: ***The Oxford History of South Africa***, I, South Africa to 1870, Clarendon Press, Oxford, 502p.

WHISSON, M.G., 1972: ***The Fairest Cape?*** An Account of the Coloured People in the District of Simonstown, SAIRR, Johannesburg 37p.

APPENDIX 1

CLUSTER PROGRAM

ELT,L GEOG*CLUST.DRIVER
ELT ER1 S74QIC 01/30/80 12:39:39 (0)

```
C**************************** DMAIN ****************************************
C
C              VERSION 4/A (OCTOBER 1978)
C
C              PURPOSE:
C              --------
C                 THIS PROGRAM PERFORMS SIMILARITY ANALYSIS
C                 CLASSIFICATIONS BY UTILIZING THE BRAY & CURTIS MEASURE
C                 OF SIMILARITY, OR ALTERNATIVELY, BY USING THE CANBERRA
C                 METRIC OF LANCE & WILLIAMS IF SO SPECIFIED.
C
C              DESCRIPTION OF INPUT PARAMETERS:
C              --------------------------------
C              CARD 1:
C              -------
C                 IREAD   - LOGICAL UNIT FOR CARD READER.
C                           (COLS.1-3 : I3)
C                 IARIT   - LOGICAL UNIT FOR LINE-PRINTER.
C                           (COLS.4-6 : I3)
C                 IARITA  - LOGICAL UNIT FOR ALTERNATE PRINT DEVICE.
C                           (COLS.7-9 : I3)
C                 IFILE   - INPUT DEVICE UNIT FOR DATA.
C                           (COLS.10-12 : I3)
C                 ITMAT   - OUTPUT DEVICE UNIT FOR LOWER TRIANGLE
C                           OF SIMILARITY MATRIX.
C                           (COLS.13-15 : I3)
C              CARD 2:
C              -------
C                 TITLE   - TITLE OF ANALYSIS.
C                           (COLS.1-80 : 13A6,A2)
C              CARD 3:
C              -------
C                 ISPNO   - NUMBER OF SPECIES.
C                           (COLS.1-4 : I4)
C                 ISAMNO  - NUMBER OF SAMPLES.
C                           (COLS.6-9 : I4)
C                 ILOG    - TRANSFORMATION OPTION.
C                              0=NO TRANSFORM OF DATA
C                              1=LOG TRANSFORM
C                              2=ARC-SIN TRANSFORM
C                           (COL.13 : I1)
C                 ICOEF   - METRIC OPTION.
C                              'CURTIS'=BRAY & CURTIS MEASURE OF SIMILARITY USED
C                              'CANBER'=CANBERRA METRIC OF LANCE & WILLIAMS USED
C                           (COLS.15-20 : A6)
C                 IPRINT  - PRINT OPTION.
C                              0=INPUT DATA NOT PRINTED
C                              1=INPUT DATA MATRIX WILL BE PRINTED
C                                ACCORDING TO FORMAT 'VAROUT'
C                           (COL.22 : I1)
C                 IPLOT   - PLOT OPTION.
C                              0=NO DENDROGRAM PLOTTED
C                              1=DENDROGRAM PLOTTED ON CALCOMP (G.P.C.P.)
C                           (COL.24 : I1)
C                 ITIME   - ESTIMATED PLOT TIME IN MINUTES.
C                           (COLS.26-28 : I3)
C              CARD 4:
C              -------
C                 ILINK   - TYPE OF CLUSTERING ANALYSIS DESIRED.
C                              1=NEAREST NEIGHBOUR
```

```
C                                   3=FURTHEST NEIGHBOUR
C                                 (COL.1 : I1)
C                ICONT     - CONTINUATION OPTION.
C                                   0=END OF CURRENT ANALYSIS
C                                   1=ANOTHER CLUSTERING TO BE PERFORMED. PREVIOUS
C                                     DATA SET WILL BE USED. IF THIS IS DESIRED,
C                                     THIS CARD MUST BE FOLLOWED BY ANOTHER 'LAST'
C                                     CARD.
C                                 (COL.3 : I1)
C           CARD 5:
C           ------
C                VARFOR    - VARIABLE DATA FORMAT.
C                                 FOR EXAMPLE: '(30F3.1)'
C                                 (COLS.1-80 : 13A6,A2)
C           CARD 6:
C           ------
C                VAROUT    - VARIABLE FORMAT FOR DATA OUTPUT.
C                                 THIS CARD NEED ONLY BE INCLUDED IF VARIABLE
C                                 IPRINT (CARD 3) HAS A VALUE >0
C                                 EXAMPLE: '(30F4.1)'
C                                 (COLS.1-80 : 13A6,A2)
C           CARD 7 TO (N):
C           -------------
C                X         - DATA.
C                                 SCORES FOR EACH SAMPLE, ONE SPECIES AT A TIME.
C                                 A NEW CARD FOR EACH SPECIES.
C                                 THE INDIVIDUALS (OR OBJECTS) TO BE CLASSIFIED
C                                 ARE STORED IN COLUMNS AND ATTRIBUTES
C                                 (OR CHARACTERS) IN ROWS.
C                                 EACH RECORD (ROW) IS READ IN ACCORDING TO VARIABLE
C                                 FORMAT 'VARFOR'.
C           METHOD:
C           ------
C                AFTER PRINTING THE MATRIX SHOWING THE SIMILARITY BETWEEN
C                INDIVIDUALS, IN PERCENTAGES, CLUSTERING IS PERFORMED.
C                BRAY & CURTIS METHOD IS DESCRIBED IN ECOL.MONOGR.1957, AND
C                IN ZOOL.AFR.1968 BY FIELD AND MCFARLANE.
C                THE CANBERRA METRIC OF LANCE & WILLIAMS IS DESCRIBED IN
C                AUSTR.COMP.J.1967.
C
C           MODIFICATIONS:
C           -------------
C                BY MIKE RIPP (SCHOOL OF ENVIRONMENTAL STUDIES,
C                U.C.T.), AND TONY JACKSON (IMAGE PROCESSING
C                UNIT, U.C.T.). JAN/FEB.1978.
C                AND JOHN NEWMARCH (COMPUTER CENTRE) OCTOBER 1978.
C
C           NOTES:
C           -----
C                X(J)         - DATA.
C                ANUM(IR)     - COEFFICIENT NUMERATOR ACCUMULATOR, THEN
C                               SIMILARITY MATRIX.
C                ADEN(IR)     - DENOMENATOR ACCUMULATOR, THEN DISSIMILARITY
C                               MATRIX.
C                ISORT(I,J)   - INDIVIDUALS IN EACH CLUSTER DURING SORTING.
C                IGPNO(K)     - NUMBER OF INDIVIDUALS IN EACH CLUSTER.
C
C           1. THE PROGRAM USES DYNAMIC CORE ALLOCATION. THE MAXIMUM
C              NUMBER OF SPECIES AND SAMPLES IS INDETERMINATE
C           2. VARIABLES ANUM(IR) AND ADEN(IR) ARE DIMENSIONED ACCORDING
C              TO THE FOLLOWING EQUATION:
```

```
00
00     C---------------------------------------------------------------------------------
00     C
00     355    FORMAT (5I3)
00     360    FORMAT (A80)
00     364    FORMAT(1H1,T2,'PROGRAM CLUSTER.  VERSION 4/A'/
00           1       T2,'------------------------------------'///)
00     365    FORMAT (1H ,T2,'TITLE OF ANALYSIS:',2X,A80/T2,17(1H-)///)
00     370    FORMAT (I4,1X,I4,3X,I1,1X,A6,1X,I1,1X,I1,1X,I3)
00     490    FORMAT (1H1)
00            END

END ELT.  ERRORS: NONE. TIME:  1.295 SEC. IMAGE COUNT: 1°6

@HDG,P  ******  MAP  ******  .L,C
```

```
      SUBROUTINE ONE(X,ANUM,ADEN,ISAM,ISAM3,IDIM)
      INCLUDE OD,LIST
      INCLUDE ALL,LIST
      INCLUDE OT,LIST
      DIMENSION X(ISAM),ANUM(IDIM),ADEN(IDIM)
     &,IFR(11)
      CHARACTER VARFT*96,VARFOR*80(2),VARCUT*84

 1    READ(IREAD,435)  ILINK,ICONT
      READ (IREAD,375) VARFOR(1)
      IF(VARFOR(1)(80:80) .EQ. '&')THEN @ THIS ALLOWS FOR A SECOND
        VARFOR(1)(80:80) = ' '   @ FORMAT CARD
        READ(IREAD,375)VARFOR(2)
      END IF
      IF(IPRINT.EQ.1) THEN
      READ(IREAD,375) VAROUT
      VARFT='((1X,' & VAROUT & '))'
      END IF
      IF (ICOEF.EQ.'CANBER') THEN
      WRITE (IWRIT,385)
      ELSE IF (ICOEF.EQ.'CURTIS') THEN
      WRITE (IWRIT,380)
      ELSE
      STOP 111
      END IF
      ITOTO=ISAMNO*(ISAMNO+1)/2
      DO 60 I=1,ITOTO
      ANUM(I)=0.0
 60   ADEN(I)=0.0
      CONTINUE
C
C        ANALYSIS OF SAMPLES.
C
      IROW=ISPNO
      ICOL=ISAMNO
      WRITE (IWRIT,395) ICOL,IROW
      WRITE(IWRIT,396) ILOG,ICOEF,IPRINT,IPLOT,ITIME,ILINK,ICONT
      IF(IPRINT.EQ.1) WRITE(IWRIT,400)
C
C        LOOP SPECIES (ISPNO) TIMES
C
 35   READ (IFILE,VARFOR,END=999) (X(J),J=1,ISAMNO)
C
C          PRINT INPUT DATA IF REQUESTED.
C
      IF (IPRINT.EQ.1) WRITE (IWRIT,VARFT)(X(J),J=1,ISAMNO)
C
C
C        LOG TRANSFORMATION OF DATA, ONLY IF REQUESTED.
C
      IF (ILOG.NE.0) THEN
      IF (ILOG.EQ.1) WRITE (IWRIT,405)
      IF (ILOG.EQ.2) WRITE (IWRIT,410)
      DO 75 J=1,ICOL
      IF (ILOG.EQ.2) THEN
      X(J)=ASIN(SQRT(X(J)/100))
      ELSE
      X(J)=ALOG(X(J)+1.0)
      END IF
 75   CONTINUE
```

```
      ICOLM1=ICOL-1
C
C         LOOP FOR EACH ATTRIBUTE
C
      DO 95 I=1,ICOLM1
      K=I+1
      DO 90 J=K,ICOL
C
C         LOOPS TO COMPARE EACH COL WITH EVERY OTHER COL
C
      IR=I+(J*J-J)/2
C
C         FORMULA FOR LOWER TRIANGULAR MATRIX
C
      IF (ICOEF.EQ.'CANBER') THEN
      ANUM(IR)=ABS(X(I)-X(J))
      IF (.NOT.(X(I).EQ.0.0.AND.X(J).EQ.0.0))
     &ADEN(IR)=ADEN(IR)+ANUM(IR)/(X(I)+X(J))
      ELSE
      ANUM(IR)=ANUM(IR)+ABS(X(I)-X(J))
      ADEN(IR)=ADEN(IR)+X(I)+X(J)
      END IF
   90 CONTINUE
   95 CONTINUE
      GO TO 35
  999 CONTINUE
      WRITE (IWRIT,415) (I,I=1,ICOLM1)
C
C                   CALCULATE AND PRINT COEFFICIENTS
C
      DO 120 I=1,ICOLM1
      K=I+1
      DO 120 J=K,ICOL
      IR=I+(J*J-J)/2
      IF (ICOEF.EQ.'CANBER') THEN
C
C                     CANBERRA METRIC DISSIMILARITY FUNCTION
C
      ADEN(IR)=(ADEN(IR)*100./IROW)
      ELSE
C
C        CHECK FOR POSSIBLE DIVIDE ERROR
C
      IF (ADEN(IR).LE.0.0) THEN
      WRITE(IWRITA,500) I,J
      ANUM(IR)=0.0
C
C                   BRAY-CURTIS DISSIMILARITY FUNCTION
C
      ELSE
      ADEN(IR)=ANUM(IR)*100./ADEN(IR)
      END IF
      END IF
  120 ANUM(IR)=100-ADEN(IR)
C
C        ADEN = DISSIMILARITY FUNCTION, ANUM = SIMILARITY FUNCTION
C
      DO 125 I=2,ICOL
      K=I-1
      J=1+(I*I-1)/2
      L=K+(I*I-1)/2
```

```
      WRITE (IWRIT,425) I,(ANUM(IR),IR=J,L)
C
C           WRITE SIMILARITIES TO DISK FILE ITMAT.
C
      WRITE (ITMAT,420) (ANUM(IR),IR=J,L)
125   CONTINUE
      DO 130 I=1,11
130   IFR(I)=0
      DO 135 I=1,ICOLMI
      K=I+1
      DO 135 J=K,ICOL
      IR=I+(J*J-J)/2
      IP=IFIX(ANUM(IR)/10)
      IP=IP+1
      IFR(IP)=IFR(IP)+1
135   CONTINUE
      WRITE (IWRIT,430) (1,I=1,11),(IFR(IP),IP=1,11)
      DO 140 I=1,ICOLMI
      K=I+1
      DO 140 J=K,ICOL
      IR=I+(J*J-J)/2
140   ANUM(IR)=ADEN(IR)
      RETURN
375   FORMAT (A80)
380   FORMAT (1X,'BRAY-CURTIS SIMILARITY MEASURE -ABUNDANCE-WEIGHTED')
385   FORMAT (1X,'CANBERRA METRIC (THIS MEASURE IS NOT ABUNDANCE-WEIGHTE
     1D)')
395   FORMAT (1X,'SAMPLE- OR Q-ANALYSIS',///,' NO OF SAMPLES (INDIVIDUAL
     1S)= ',I4,/,' NO OF SPECIES (ATTRIBUTES) = ',I4//)
396   FORMAT(1H ,T2,'ILCG    =',I2/
     2              T2,'ICOEF   =',1X,A6/
     3              T2,'IPRINT  =',I2/
     4              T2,'IPLOT   =',I2/
     5              T2,'ITIME   =',I3/
     6              T2,'ILINK   =',I2/
     7              T2,'ICONT   =',I2//)
400   FORMAT (1H1,//,T5,'INPUT DATA MATRIX:'/T5,17(1H-)/)
405   FORMAT (1X,' LOG-TRANSFORMED DATA,Y= LN (X + 1)')
410   FORMAT (1X,'ARC SINE - TRANSFORMED DATA,  Y=ARCSIN(ROOT(X))')
415   FORMAT (1H1,//,T5,'SIMILARITY MATRIX'/T5,17(1H-)/(9X,12I9))
420   FORMAT (6G13.6)
425   FORMAT (1X,I9,12F9.3,/,(10X,12F9.3))
430   FORMAT (1H1,1X,'FREQUENCY DISTRIBUTION OF COEFFICIENT',/2X,37(1H-)
     1//1X,11I8//1X,11I8///)
435   FORMAT (I1,1X,I1)
500   FORMAT(1H ,T5,'***** E R R O R *****',5X,'DIVIDE CHECK NO =',I4,
     13X,'OTUS =',I4)
      END

END ELT.  ERRORS: NONE.  TIME:  1.124 SEC.  IMAGE COUNT: 170

HCG,P ****** PROCS ****** .L,O
```

ELT,L
ELT 8R1 S74Q1C 01/30/80 12:39:45 (0)

```
      SUBROUTINE TWO(ANUM,ADEN,ISORT,IGPNO,DNODL,ISAM,ISAM3,IDIM)
      INCLUDE ALL,LIST
      INCLUDE OT,LIST
      INCLUDE DT,LIST
      DIMENSION ANUM(IDIM),ADEN(IDIM),ISORT(ISAM,ISAM)
     &,IGPNO(ISAM),DNODE(6,ISAM3)
      DO 155 I=1,ICOL
      IR=I+(I*I-I)/2
  155 ADEN(IR)=8888.
      M=ICOLM1
      GO TO (160,165,170), ILINK
  160 WRITE (IWRIT,440)
      WRITE (IWRIT,445)
      GO TO 175
  165 WRITE (IWRIT,450)
      WRITE (IWRIT,455)
      GO TO 175
  170 WRITE (IWRIT,460)
      WRITE (IWRIT,465)
  175 DO 180 I=1,ICOL
      DO 180 J=1,ICOL
  180 ISORT(I,J)=0
      DO 185 I=1,ICOL
  185 IGPNO(I)=1
      OVALU=998.
      AMIN=9999.
  190 DO 200 I=1,ICOLM1
C
C        LOOP FOR EACH COLUMN OF MATRIX
C
      K=I+1
      DO 200 J=K,ICOL
C
C         LOOP FOR EACH ROW OF MATRIX
C
      IR=I+(J*J-J)/2
C
C        CALCULATE MATRIX LOCALITY
C
      IF (ADEN(IR)-AMIN .LE. 0.0  ) THEN
C
C        SEARCH FOR LOWEST VALUE IN DISSIMILARITY MATRIX
C
      AMIN=ADEN(IR)
      IGRP=J
      IDEL=I
C
C       IDEL IS OTU TO BE DELETED WHEN NEW GROUP IS FORMED
C
      END IF
  200 CONTINUE
      DO 205 MO=1,ICOLM1
      IF (IGRP-ISORT(MO,1))205,210,205
C
C       IS IGRP ON LHS OF ANY ISORT ARRAY?  IS IT A MEMBER OF AN
C       EXISTING GROUP?
C
  205 CONTINUE
      I=1
      ISORT(M,I)=IGRP
```

```
      [illegible]
      GO TO 220
  210 IT=IGPNO(IGRP)
      DO 215 I=1,IT
  215 ISORT(M,I)=ISORT(MO,I)
      I=IT
  220 ITOT=IGPNO(IGRP)+IGPNO(IDEL)
      I=I+1
      DO 225 MO=1,ICOLM1
      IF (IDEL-ISORT(MO,1))225,230,225
C
C        IS IDEL ON LHS OF ANY ISORT ARRAY? IS IDEL IN AN EXISTING GP?
C
  225 CONTINUE
      ISORT(M,I)=IDEL
      GO TO 240
  230 DO 235 J=I,ITOT
C
C        ANSWER YES - PUT OLD ARRAY ON RHS OF NEW ARRAY
C
      K=J-I+1
  235 ISORT(M,J)=ISORT(MO,K)
  240 VALU=100-AMIN
C
C        COMPLEMENT OF MINIMUM VALUE
C
      IF (VALU-OVALU .NE. 0.0) THEN
      WRITE (IWRIT,470) VALU,IGRP,IDEL,ITOT
      ELSE
      WRITE (IWRIT,475) IGRP,IDEL,ITOT
      END IF
      OVALU=VALU
      WRITE (IWRIT,480) (ISORT(M,J),J=1,IT)
      WRITE (IWRIT,490) (ISORT(M,J),J=1,ITOT)
      M=M-1
      SI=SI+1
C
C                 RELEVANT PLOTTING DATA STORED IN VARIABLE DNODE.
C
      DNODE(1,SI)=AMIN/100.0
      DNODE(2,SI)=IGRP
      DNODE(3,SI)=IGRP
      DNODE(4,SI)=IDEL
      DNODE(5,SI)=IGPNO(IGRP)
      DNODE(6,SI)=IGPNO(IDEL)
C
C        CHANGE LOCATION FOR NEXT NEW ROW OF ARRAY ISORT
C
      IR=IDEL+(IGRP*IGRP-IGRP)/2
      ADEN(IR)=8888.
      IF(M)340,340,260
  260 GO TO (265,280,265), ILINK
  265 ALPHI=0.5
      ALPHJ=0.5
      GO TO (270,280,275), ILINK
  270 GAMMA=-0.5
      GO TO 285
C
C        NEAREST NEIGHBOUR SORTING
C
  275 GAMMA=0.5
      GO TO 285
```

```
C
C      FURTHEST NEIGHBOUR SORTING
C
  280 A=FLOAT(IGPNO(IGRP))
      ALPHI=A/(A+IGPNO(IDEL))
      A=FLOAT(IGPNO(IDEL))
      ALPHJ=A/(A+IGPNO(IGRP))
      GAMMA=0.0
C
C            GROUP AVERAGE SORTING
C
  235 DO 335 J=1,ICOL
      IF (IGRP-J .LE. 0) THEN
C
C            LOOP FOR ALL COEFFICIENTS WITH NEW GROUP IGRP
C
      IS=IGRP+(J*J-J)/2
      ELSE
C
C            MATRIX ADDRESS OF IGRP
C
      IS=J+(IGRP*IGRP-IGRP)/2
      END IF
      IF (IDEL-J .LE. 0) THEN
      IR=IDEL+(J*J-J)/2
      ELSE
      IR=J+(IDEL*IDEL-IDEL)/2
C
C                   MATRIX ADDRESS OF IDEL
C
      END IF
      IF (ADEN(IS)-8888.) 320,330,330
  320 IF (ADEN(IR)-8888.) 325,335,335
  325 ADEN(IS)=ALPHI*ADEN(IS)+ALPHJ*ADEN(IR)+GAMMA*(ABS(ADEN(IR)-ADEN(IS
     1)))
  330 ADEN(IR)=8888.
  335 CONTINUE
      IGPNO(IGRP)=ITOT
      IGPNO(IDEL)=7777
      AMIN=9999.
      GO TO 190
  340 WRITE (IWRIT,485)
      WRITE (IWRIT,480) (ISORT(1,I),I=1,ICOL)
      DO 345 I=1,ICOLM1
      K=I+1
      DO 345 J=K,ICOL
      IR=I+(J*J-J)/2
  345 ADEN(IR)=ANUM(IR)
  290 RETURN
  440 FORMAT (1H1,'NEAREST NEIGHBOUR SORTING')
  445 FORMAT (1H ,25(1H-)/)
  450 FORMAT (1H1,'GROUP AVERAGE SORTING')
  455 FORMAT (1H ,21(1H-)/)
  460 FORMAT (1H1,'FURTHEST NEIGHBOUR SORTING')
  465 FORMAT (1H ,26(1H-)/)
  470 FORMAT (/,1X,'SIMILARITY LEVEL = ',F7.3,/,2X,'OBJECTS ',I9,5X,I9,2
     10X,'NO OF OBJECTS IN GROUP ',I4)
  475 FORMAT (2X,'OBJECTS ',I9,5X,I9,10X,'NO OF OBJECTS = ',I4)
  480 FORMAT (1X,13I10)
  485 FORMAT (1H1,1X,'DENDROGRAM SEQUENCE OF OTUS '/2X,27(1H-)//)
  500 FORMAT(1H ,T5,'***** E R R O R *****',5X,'DIVIDE CHECK NO =',I4,
     13X,'OTUS =',I4)
```

```
5590.         00      C
5600.         00      C-----------------------------------------------------------------
5610.         00            END
END ELT.  ERRORS: NONE. TIME:  1.134 SEC. IMAGE COUNT: 187
@HDG,N
@RESUME,F PR
```

LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LIND
LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LIND
LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LINDYM LIND

```
      SUBROUTINE DPLOT(DNODE,S1,IWRIT,TITLE,ICOL,ITIME,ISAM,ISAM3
     &,ICONT,DNDE)
C
C        PURPOSE:
C        -------
C           THIS SUBROUTINE DRAWS THE DENDROGRAM.
C
C        NOTE:
C        ----
C           ALL DATA TO PLOT DENDROGRAM ARE NOW STORED IN DNODE.
C
C-----------------------------------------------------------------
C
C        SET UP PLOTTING AREA AND PLOT HEADERS.
C
      DIMENSION DNODE(6,ISAM3),DNDE(2,ISAM3)
      INTEGER S1,S9
      REAL LMER
      CHARACTER TITLE*80
C
C
C        INITIALIZE PLOT.
C
      DMAX=DNODE(1,S1)
      DO 450 IFACT=1,10
      DMAX=DMAX*10.0
      IF (DMAX.GT.1.0) GOTO 460
 450  CONTINUE
 460  IFACT=IFACT-1
      AXMLT=10.0**IFACT
C
C        SET HEIGHT OF PLOT (INCHES).
C
      IF(ICOL.LT.50)   AXLEN=5.0
      IF(ICOL.GE.50)   AXLEN=10.0
      IF(ICOL.GE.100)  AXLEN=15.0
      IF(ICOL.GE.200)  AXLEN=20.0
      AXMLT=AXMLT*AXLEN
      DMAX=AINT(DMAX+1.0)/10.0
      DELTAV=-(100.0/AXLEN)
      CALL PLOTS(0,0,0)
      CALL PLTIME(ITIME)
      CALL NEWPAG
      XLEN=.25*ISAM+3.0
      CALL PAGSIZ(XLEN*2.54,(AXLEN+3.0)*2.54)
      CALL FACTOR(2.54)
C
C
C        PLOT LEFT HAND AXIS AND INITIALIZE COUNTER.
C
      CALL PLOT(2.0,2.0,-3)
      CALL AXIS(0.0,0.0,12H% SIMILARITY,12,AXLEN,90.0,100.0,DELTAV)
      POSN=1.0+AXLEN
      CALL SYMBOL(11.5,POSN,0.14,TITLE,180.0,80)
      X=0.5
      S9=S1
      DNDE(1,S1)=-1.0
C
C
C        SEARCH DOWN LEFT HAND SIDE OF TREE.
```

```
  490 IF(DNODE(5,SI).EQ.1.0) GOTO 510
      DNDE(1,SI)=-1.0
      LMER=DNODE(3,SI)
  500 SI=SI-1
      IF(DNODE(2,SI).NE.LMER) GOTO 500
      GOTO 490
C
C        PLOT A LEFT POINT.
C
  510 Y=DNODE(1,SI)*AXMLT
      XST=X+0.035
      CALL NUMBER(XST,-0.42,0.07,DNODE(3,SI),90.0,0)
      CALL PLOT(X,0.0,3)
      CALL PLOT(X,Y,2)
      DNDE(2,SI)=X
      X=X+0.25
C
C        SEARCH DOWN RIGHT HAND SIDE OF TREE.
C
  520 IF(DNODE(6,SI).EQ.1.0) GOTO 540
      DNDE(1,SI)=-2.0
      RMER=DNODE(4,SI)
  530 SI=SI-1
      IF(DNODE(2,SI).NE.RMER) GOTO 530
      GOTO 490
C
C        PLOT A RIGHT HAND POINT AND LINK TO LEFT POINT.
C
  540 XST=X+0.035
      CALL NUMBER(XST,-0.42,0.07,DNODE(4,SI),90.0,0)
      CALL PLOT(X,0.0,3)
      Y=DNODE(1,SI)*AXMLT
      CALL PLOT(X,Y,2)
      I=SI
      IF(DNODE(5,SI).EQ.1.0) GOTO 560
  550 I=I-1
      IF(DNODE(2,I).NE.DNODE(3,SI)) GOTO 550
  560 CALL PLOT(DNDE(2,I),Y,2)
      Y=DNODE(1,I)*AXMLT
      CALL PLOT(DNDE(2,I),Y,2)
      DNDE(2,SI)=(X+DNDE(2,I))/2.0
      X=X+0.25
      DNDE(1,SI)=-999.0
C
C        SEARCH UP TREE FOR TAGS.  DNDE(1,I)=-1.0 FOR LEFT TAG,
C        -2.0 FOR RIGHT TAG, AND -999.0 TO KILL CLUSTER.
C
  570 SI=SI+1
      IF (SI.GT.S9) GOTO 650
      IF(DNDE(1,SI).EQ.-1.0) GOTO 520
      IF(DNDE(1,SI).EQ.-2.0) GOTO 580
      GOTO 570
C
C        DRAW LINK BETWEEN CLUSTERS.
C
  580 I=SI
      IF(DNODE(5,SI).EQ.1.0) GOTO 600
  590 I=I-1
      IF(DNODE(2,I).NE.DNODE(3,SI)) GOTO 590
  600 XI=DNDE(2,I)
      Y=DNODE(1,I)*AXMLT
```

```
      Y=DNODE(1,S1)*AXMLT
      CALL PLOT(X1,Y,2)
      I=S1
      IF(DNODE(6,S1).EQ.1.0) GOTO 620
  610 I=I-1
      IF(DNODE(2,I).NE.DNODE(4,S1)) GOTO 610
  620 X2=DNDE(2,I)
      CALL PLOT(X2,Y,2)
      Y=DNODE(1,I)*AXMLT
      CALL PLOT(X2,Y,2)
      DNDE(2,S1)=(X1+X2)/2.0
  630 DNDE(1,S1)=-999.0
C
C        SEARCH UP TREE FOR NEXT TAG.
C
  640 S1=S1+1
      IF(S1.GT.S9) GOTO 650
      IF(DNDE(1,S1).EQ.-1.0) GOTO 520
      IF(DNDE(1,S1).EQ.-2.0) GOTO 550
      GOTO 640
C
C
C        PLOT RIGHT HAND AXIS AND TERMINATE PLOT.
C
  650 Y=AXLEN-0.16
      X=X+0.25
      CALL AXIS(X,0.0,12H% SIMILARITY,-12,AXLEN,90.0,100.0,DELTAV)
      X=X+1.0
      X=X+8
      Y=-2.0
      IF(ICONT.EQ.1)RETURN
      STOP 'NORMAL TERMINATION AFTER PLOTTING'
      END

END ELT.  ERRORS: NONE.  TIME:  1.116 SEC.  IMAGE COUNT:  156

@HDG,P   ******   DRIVER   ******   .L,C
```

APPENDIX 2

INFORMATION STATISTIC PROGRAM

```
GELT.L CROWE*INFOSTAT.ALL
ELT 8R1 S74QIC 05/21/80 16:17:08 (0)
        00
        00      C          THIS PROGRAM PERFORMS INFORMATION STATISTIC TESTS BETWEEN
        00      C       GROUPS OF INDIVIDUALS.
        00
        00      C       THE ALGORITHMS AND CODE TO CALCULATE THE 2 DELTA I VALUES WAS
        00      C       PRODUCED BY JOHN FIELD IN 1969.
        00
        00      C       THE REST OF THE PROGRAM WAS WRITTEN BY IAN BLEACH
        00      C                                              101 CALEDONIAN COURT
        00      C                                              DURBAN ROAD
        00      C                                              MOWBRAY
        00      C                                              CAPE.
        00      C                                             TELEPHONE 654-855
        00      C                                             7/11/79
        00
        00
        00
        00      C          PROGRAM INPUTS
        00      C          ==============
        00
        00
        00      C       CARD 1
        00
        00      C       COLUMNS 1-80 THE TITLE OF THE ANALYSIS
        00
        00
        00      C       CARD 2
        00
        00      C       COLUMNS 1-4  THE NUMBER OF SAMPLES
        00      C       COLUMNS 6-8  THE NUMBER OF SPECIES
        00      C       COLUMNS 10   INDICATIOR      = 0  ANALYSIS OF SPECIES
        00      C                                    = 1  ANALYSIS OF SAMPLE
        00      C       COLUMNS 12   FREQUENCY OR PRESENCE/ABSENCE    = 0  PRESENCE/ABSENCE
        00      C                                                     = 1  FREQUENCY SCORE
        00      C       COLUMN  14      = 0  MAXIMUM PRINTOUT
        00      C                       = 1  MINIMUM PRINTOUT
        00      C       COLUMN 16    THE NUMBER OF LINES IN THE INPUT FORMAT
        00      C                           (0 IS TAKEN TO BE 1 LINE)
        00      C                           MAXIMUM OF 4
        00
        00
        00      C       CARD 3
        00
        00      C       THIS CARD PLUS AS MANY AS INDICATED FOR THE NUMBER OF LINES
        00      C       OF FORMAT ARE TO CONTAIN THE FORMAT USE TO READ IN THE DATA.
        00
        00
        00      C       REST OF CARDS
        00
        00      C       THESE CARDS GIVE THE SETS
        00
```

```
00
00  C      THE FIRST SET IS ENDED WITH A -1
00
00  C      THE SECOND SET STARTS ON THE NEXT CARD
00  C      THE SECOND SET IS ENDED BY A -2
00
00  C      THE SECOND SET IS ENDED BY A -3 IF THERE ARE NO MORE SETS
00
00
00  C      INPUTS FROM FILE
00
00  C      FILE 11.
00
00  C         THIS FILE CONTAINS THE CENSIS DATA.
00  C         IT IS READ IN UNDER THE USER SUPPLIED FORMAT.
00
00
00  C      FILE 12.
00
00  C         THIS FILE CONTAINS THE NAMES OF THE SPECIES
00  C         IT IS READ IN UNDER A FORMAT:
00  C                                  (I4,1X,18A4)
00
00
00
00         CHARACTER    TITLE*4(20),   FORMAT*80(4)
00         INTEGER   FMT
00         DIMENSION   DUMMY(1)
00         DATA   IREAD/8/, IWRIT/5/, IFILE1/11/, IFILE2/12/
00     100 FORMAT(20A4)
00     200 FORMAT(1H1,T30,30HINFORMATION STATISTICS PROGRAM,/,T30,30H********
00        1**********************)
00     300 FORMAT(1H0,T30,20A4,///,T30,30HNUMBER OF SPECIES ............,T60,
00        1I5,/,T30,30HNUMBER OF SAMPLES ............,T60,I5)
00     400 FORMAT(1H ,T30,19HANALYSIS OF SAMPLES)
00     500 FORMAT(1H ,T30,19HANALYSIS OF SPECIES)
00     600 FORMAT(1H ,T30,30HFREQUENCY SCORE IN EACH OBJECT)
00     700 FORMAT(1H ,T30,31HPRESENCE/ABSENCE IN EACH OBJECT)
00     800 FORMAT(1H ,T30,30HNUMBER OF LINES OF FORMAT ....,T60,I5)
00     900 FORMAT(1H ,T30,25HFORMAT STATMENT FOR INPUT,4(/,T30,1H',A80,1H')/)
00    1000 FORMAT(A80)
00    1100 FORMAT(I4,1X,I3,4(1X,I1))
00    1200 FORMAT(1H ,T30,16HMAXIMUM PRINTOUT)
00    1300 FORMAT(1H ,T30,16HMINIMUM PRINTOUT)
00
00
00         DEFINE   X(I,J)     =   DUMMY(I + JX*(J-1) + ID1OFF)
00         DEFINE   ISET1(I)   =   DUMMY(I + ID2OFF)
00         DEFINE   ISET2(I)   =   DUMMY(I + ID3OFF)
00         DEFINE   IZ1(I)     =   DUMMY(I + ID4OFF)
00         DEFINE   IZ2(I)     =   DUMMY(I + ID5OFF)
00         DEFINE   ANAME(I)   =   DUMMY(I + ID6OFF)
00         DEFINE   G(I)       =   DUMMY(I + ID7OFF)
00         DEFINE   RANK1(I)   =   DUMMY(I + ID8OFF)
00         DEFINE   RANK2(I)   =   DUMMY(I + ID9OFF)
00         DEFINE   FLAG(I)    =   DUMMY(I + ID10FF)
00
00
00  C      START OF THE PROGRAM
00
00         REWIND 11
```

```
      WRITE (IWRIT,200)

C     ---- GET ALL THE INPUT PARAMETERS ----

      READ (IREAD,100)  (TITLE(I),I=1,20)
      READ (IREAD,1100)  ISPNO, ISAMNO, IND, IFREQ, IPR, FMT

C     ---- WRITE OUT THE TITLE AND THE # OF SAMPLE & SPECIES ----

      WRITE (IWRIT,300)  (TITLE(I),I=1,20), ISPNO, ISAMNO

C     ---- WRITE THE SAMPLE/SPECIES ANALYSIS HEADINGS ----

      IF (IND .EQ. 0)  WRITE (IWRIT,400)
      IF (IND .EQ. 1)  WRITE (IWRIT,500)

C     ---- WRITE THE FREQUENCY OR PRESENCE/ABUNDANCE INDICATOR ----

      IF (IFREQ .EQ. 0)  WRITE (IWRIT,700)
      IF (IFREQ .EQ. 1)  WRITE (IWRIT,600)

C     ---- WRITE OUT IF MAXIMUM OR MINIMUM PRINTOUT ----

      IF (IPR .EQ. 0)  WRITE(IWRIT,1200)
      IF (IPR .EQ. 1)  WRITE(IWRIT,1300)

C     ---- WRITE OUT THE # OF FORMAT LINES ----

      IF (FMT .EQ. 0)  FMT = 1

      WRITE (IWRIT,800)  FMT

C     ---- GET FORMAT OF INPUT DATA ----

      READ (IREAD,1000)  (FORMAT(I),I=1,FMT)

C     ---- WRITE OUT THE FORMAT ----

      WRITE (IWRIT,900)  (FORMAT(I),I=1,FMT)

C     ---- GET THE DYNAMIC CORE ----

      ID1OFF = ISPNO * ISAMNO + 25 * ISPNO + 10

      CALL  ALLOC (DUMMY, ID1OFF)

      ID2OFF = ISPNO * ISAMNO + ID1OFF + 1
      ID3OFF = ISPNO + ID2OFF + 1
      ID4OFF = ISPNO + ID3OFF + 1
      ID5OFF = ISPNO + ID4OFF + 1
      ID6OFF = ISPNO + ID5OFF + 1
      ID7OFF = ISPNO * 18 + ID6OFF + 1
      ID8OFF = ISPNO + ID7OFF + 1
      ID9OFF = ISPNO + ID8OFF + 1
      ID1OFF = ISPNO + ID9OFF + 1

      IF(IND .EQ. 1)  THEN

        IX = ISPNO
        JX = ISAMNO
```

```
      ELSE

         IX = ISAMNO
         JX = ISPNO

      END IF

C     ---- GET THE INFO SUBROUTINE TO DO THE WORK ----

      CALL  INFO (IX, JX, ISPNO, ISAMNO, IND, IFREQ, IPR, FORMAT,
     1   X(1,1), ISET1(1), ISET2(1), IZ1(1), IZ2(1), ANAME(1), G(1),
     2  RANK1(2), RANK2(1), FLAG(1), IREAD, IWRIT, IFILE1, IFILE2)
      STOP
      END
      START EDIT PAGE @  ALLOC
      SUBROUTINE  ALLOC(DUM, ISZ)
C      THIS ROUTINE ALLOCATES THE REQUIRED CORE AREA

C        INPUTS
C        ======

C     DUM          THE ARRAY ONTO WHICH THE CORE IS TO BE ADDED

C     ISZ          THE NUMBER OF WORDS REQUIRED


C        OUTPUT
C        ======


C     ISZ          HAS THE OFSET NUMBER TO INDICATE THE START OF THE AREA


      ISZ = MCORFS(ISZ) - LOC(DUM)
      RETURN
      END
      START EDIT PAGE @  INFO
      SUBROUTINE  INFO (IX, JX, ISPNO, ISAMNO, IND, IFREQ, IPR, FORMAT,
     IX, ISET1, ISET2, IZ1, IZ2, ANAME, G, RANK1, RANK2, FLAG, IREAD,
     2IWRIT, IFILE1, IFILE2)
      CHARACTER    FORMAT*80(4), ANAME*4(IX*18), CH*4, LIT*4(2)
      REAL  X, G
      INTEGER  ISET1, ISET2, IZ1, IZ2, RANK1, RANK2, FLAG
      LOGICAL  NOTEND  /.TRUE./
      DIMENSION  RANK1(IX), RANK2(IX), FLAG(IX), ISET1(IX), ISET2(IX)
      DIMENSION  IZ1(IX), IZ2(IX), X(IX,JX), G(IX)
      DATA  LIT / 4H   **, 4H    * /
  200 FORMAT(I4,1X,18A4)
  300 FORMAT(3X,25I3)
  400 FORMAT(///,1X,7HSET1 - ,(22I5),10(/,8X,22I5))
  500 FORMAT(/,1X,7HSET2 - ,(22I5),10(/,8X,22I5))
  600 FORMAT(1H1,13HATTRIBUTE #  ,10X,'NAME',60X,9HFREQUENCY,8X,'RANK1',
     &   4X,'2 DELTA 1',4X,'RANK2',/,84X,8HSUBSET 1,2X,8HSUBSET 2,/,84X,
     &   4HN1= ,I3,3X,4HN2= ,I3)
  700 FORMAT(1X,I4,A4,18A4,4X,I4,3X,I6,6X,I4.0,4X,F10.3,3X,I6.0)
  800 FORMAT(/,1X,11H2 DELTA-1= ,F12.3,5X,I3,20H  DEGRESS OF FREEDOM)
  900 FORMAT(1X,5HERROR,10I7)
```

```
      IF (IND .EQ. 1)  GO TO 10

      IROW = ISAMNO
      ICOL = ISPNO

      READ (IFILE1,FORMAT)   ((X(I,J),I=1,ISAMNO),J=1,ISPNO)
      GO TO 40

   10 IROW = ISPNO
      ICOL = ISAMNO

C     ---- GET THE DATA SET ----

      READ (IFILE1,FORMAT)   ((X(I,J),J=1,ISAMNO),I=1,ISPNO)
C     ---- READ IN THE BIRDS NAMES ----

      DO 20 I = 1, ISPNO
      KZ = (I - 1) * 18 + 1
      JZ = KZ + 17
      READ (IFILE2,200,END=30)   INO, (ANAME(LZ),LZ=KZ,JZ)
   20 CONTINUE

C     ---- INPUT THE NEXT TWO SETS ----

C     ---- FIRST SET ----

   30 K = 1
      L = 25
   40 READ (IREAD,300)   (ISET1(I),I=K,L)

      DO 50 IN = K, L
      IF (ISET1(IN) .NE. -1)   GO TO 50
      N1 = IN - 1
      GO TO 60
   50 CONTINUE
      K = K + 25
      L = L + 25
      GO TO 40



C     ---- SECOND SET ----

   60 K = 1
      L = 25
   70 READ (IREAD,300)   (ISET2(I),I=K,L)

      DO 80 IN = K, L
      IF (ISET2(IN) .NE. -2 .AND. ISET2(IN) .NE. -3)   GO TO 80
      N2 = IN - 1
      IF(ISET2(IN) .NE. -2)   NOTEND = .FALSE.
      GO TO 90
```

```
      K = K + 25
      L = L + 25
      GO TO 70



C     ---- WRITE OUT THE SETS AS THEY WERE READ IN ----

   90 WRITE (IWRIT,400)  (ISET1(I),I=1,N1)
      WRITE (IWRIT,500)  (ISET2(I),I=1,N2)



      TDELT1 = 0.0
      NDF  = 0

      DO 160 J = 1, IROW
      IZ1(J)  = 0
      IZ2(J)  = 0
      DO 150 IN = 1, ICOL

C    IS OTU IN 1ST SUBSET?

      DO 120 I = 1, N1
      IF (ISET1(I) .NE. IN)   GO TO 120
      IF (X(J,IN) .LE. 0.0)   GO TO 150
      IF (IFREQ .EQ. 1)  GO TO 110
      IZ1(J) = IZ1(J) + 1
      GO TO 150
  110 IZ1(J) = IZ1(J) + IFIX(X(J,IN))
      GO TO 150
  120 CONTINUE

C    IS OTU IN 2ND SUBSET?

      DO 140 I = 1, N2
      IF (ISET2(I) .NE. IN)   GO TO 140
      IF (X(J,IN) .LE. 0.0)   GO TO 150
      IF (IFREQ .EQ. 1)  GO TO 130
      IZ2(J) = IZ2(J) + 1
      GO TO 150
  130 IZ2(J) = IZ2(J) + IFIX(X(J,IN))
      GO TO 150
  140 CONTINUE
  150 CONTINUE
  160 CONTINUE




      NT = N1 + N2
      G(1) = 0.0
      D = NT * ALOG(FLOAT(NT))
      IF (N1)  530, 530, 170
  170 D1 = N1 * ALOG(FLOAT(N1))
      IF (N2)  530, 530, 180
  180 D2 = N2 * ALOG(FLOAT(N2))
```

```
      IF (IFREQ .EQ. 1)  GO TO 380
C     ---- PRESENCE/ABSENCE 2 DELTA I CALCULATIONS ----

      IC = IZ1(J) + IZ2(J)
      IF (IC)  530, 190, 210
  190 E = 0.0
      GO TO 230

  210 E = IC * ALOG(FLOAT(IC))
      NDF = NDF + 1
      IF (NT-IC)  530, 220, 230
  220 F = 0.0
      GO TO 240

  230 F = (NT - IC) * ALOG(FLOAT(NT - IC))
  240 TOTI = D - E - F
C    TOTAL INF. FOR 2 SUBSETS TOGETHER FOR JTH ATTRIBUTE
      IF (IZ1(J))  530, 250, 260
  250 E = 0.0
      GO TO 270

  260 E = IZ1(J) * ALOG(FLOAT(IZ1(J)))
  270 IX = N1 - IZ1(J)
      IF (IX)  530, 280, 290
  280 F = 0.0
      GO TO 310

  290 F = IX * ALOG(FLOAT(IX))
  310 DELT1 = D1 - E - F
      IF (IZ2(J))  530, 320, 330
  320 E = 0.0
      GO TO 340

  330 E = IZ2(J) * ALOG(FLOAT(IZ2(J)))
  340 IX = N2 - IZ2(J)
      IF (IX)  530, 350, 360
  350 F = 0.0
      GO TO 370

  360 F = IX * ALOG(FLOAT(IX))
  370 DELT2 = D2 - E - F

C    TOTAL INF-(SUBSET 1 INF+ SUBSET 2 INF)

      G(J) = 2 * (TOTI - DELT1 - DELT2)

      GO TO 440



C     ---- FREQUENCY CALCULATIONS ----

  380 IZTOT = IZ1(J) + IZ2(J)
      D = 0.
      E = 0.
      F = 0.
      G(J) = 0.
      IF (IZTOT)  530, 440, 390
  390 NDF = NDF + 1
      DO 430 I = 1, ICOL
```

```
          IF (ISET1(I [illegible]
          D = D + X(J,I) * ALOG(X(J,I) * FLOAT(NT)/IZTOT)
          E = E + X(J,I) * ALOG(X(J,I) * FLOAT(N1)/IZ1(J))
          GO TO 430
      410 CONTINUE
          DO 420 IP = 1, N2
          IF (ISET2(IP) .NE. I)  GO TO 420
          D = D + X(J,I) * ALOG(X(J,I) * FLOAT(NT)/IZTOT)
          F = F + X(J,I) * ALOG(X(J,I) * FLOAT(N2)/IZ2(J))
          GO TO 430
      420 CONTINUE
      430 CONTINUE
          G(J) = 2 * (D - E - F)
      440 TDELTI = TDELTI + G(J)

    C 2 DELTA 1 SUMMED OVER ALL ATTRIBUTES

          IF (IPR .EQ. 1)  GO TO 480
          IF (G(J) .LT. 3.84)  GO TO 450
          IF (G(J)-6.64)  460, 470, 470

    C     ---- NO SIGNIFICANCE ----

      450 FLAG(J) = 0
          GO TO 480

    C     ---- 1% SIGNIFICANCE LEVEL ----

      460 FLAG(J) = 1
          GO TO 480


    C     ---- 5% SIGNIFICANCE LEVEL ----

      470 FLAG(J) = 2

      480 CONTINUE

    C     ---- DO THE OUTPUT OF THE RESULTS ----

          IF (IPR .EQ. 1)  GO TO 520

    C        ---- GET THE RANK1 AND RANK2 ARRAYS SET ----

             CALL SORTIT(IX, IZ1, IZ2, FLAG, G, N1, N2, RANK1, RANK2,
         +                ISET1, ISET2)

          WRITE (IWRIT,600)  N1, N2

          DO 520 J = 1, ISPNO
          K = (J - 1) * 18 + 1
          L = K + 17
          CH = ' '
          IF (FLAG(J) .EQ. 0)  GO TO 510
          IF (FLAG(J) .EQ. 1)  GO TO 490
          CH = LIT(1)
          GO TO 510
      490 CH = LIT(2)
      510 WRITE (IWRIT,700)  J, CH, (ANAME(M),M=K,L), IZ1(J), IZ2(J),
         1  RANK1(J), G(J), RANK2(J)
```

```
      IN = NDF - 1
      WRITE (IWRIT,800)  TDELTI, IN
      IF (NOTEND)  GO TO 30

      RETURN

C     ---- ERROR ----

  530 WRITE (IWRIT,900)  J, IZ1(J), IZ2(J), N1, N2, IC, IX
      STOP
      END
      START EDIT PAGE  ω  SORTIT
      SUBROUTINE  SORTIT(IX, IZ1, IZ2, FLAG, G, N1, N2, RANK1,
     &      RANK2, LINK1, LINK2)

C        THIS SUBROUTINE SORTS THE 5% AND 1% SIGNIFICANCE
C     2 DELTA I VALUES & GIVES THEN THIER RANKING.
C     THE RANKINGS ARE IN 2 PARTS AS RANKINGS FOR SET 1
C     SIGNIFICANCE & RANKINGS FOR SET 2 SIGNIFICANCE.

C        INPUTS
C        ======

C     IX            THE SIZE OF ALL THE ARRAYS

C     IZ1           THE SPECIES NUMBERS OF SET 1

C     IZ2           THE SPECIES NUMBERS OF SET 2

C     FLAG          A FLAG ARRAY INDICATING WHICH OF THE VALUES ARE
C                   SIGNIFICANT.  FLAG = 0   NO SIGNIFICANCE
C                                 FLAG = 1   1%  SIGNIFICANCE
C                                 FLAG = 2   5%  SIGNIFICANCE

C     G             2 DELTA I VALUES

C     N1            THE NUMBER OF MEMBERS IN SET 1

C     N2            THE NUMBER OF MEMBERS IN SET 2


C        OUTPUTS
C        =======

C     RANK1         THE RANKINGS OF SET1 SIGNIFICANCE

C     RANK2         THE RANKINGS OF SET 2 SIGNIFICANCE


C        TEMPORY STORAGE
C        ===============


C     LINK1,LINK2  THE ARRAYS TO HOLD THE LINKS
```

```
      REAL          G(IX)
      INTEGER       IZ1(IX), IZ2(IX), FLAG(IX), RANK1(IX), RANK2(IX)
      INTEGER       LINK1(IX), LINK2(IX)

C     ---- GET THE MEMBERS SIGNIFICANT TO THE FIRST SET ----

      RATIO = FLOAT(N1)/FLOAT(N2)
      J = 0
      K = 0

      DO 10  I = 1,IX

C        ---- IS THIS SPECIES SIGNIFICANT ----

      LINK1(I) = 0
      LINK2(I) = 0
         IF(FLAG(I) .NE. 0)  THEN

C           ----SIGNIFICANT BUT TO WHICH SET ----

         IF(IZ1(I) .EQ. 0) THEN
C           ---- SIGNIFICANT TO THE SECOND SET ----

            K = K + 1
            LINK2(K) = I

         ELSE  IF(IZ2(I) .EQ. 0) THEN

C           ---- SIGNIFICANT TO THE FIRST SET ----

            J = J + 1
            LINK1(J) = I

            ELSE  IF(FLOAT(IZ1(I))/FLOAT(IZ2(I)) .GT. RATIO)  THEN
C              ---- SIGNIFICANT TO SET 1 ----

               J = J + 1
               LINK1(J) = I

            ELSE

C              ---- SIGNIFICANT TO SET 2 ----

               K = K + 1
               LINK2(K) = I

            END IF
         END IF
 10   CONTINUE

C     ---- NOW SORT THE ARRAYS ----
```

```
00            IF(J .GT. 1)  CALL BUBBLE(J, G, LINK1)
00            IF(K .GT. 1)  CALL BUBBLE(K, G, LINK2)
00
00      C     ---- NOW DO THE RANKINGS ----
00
00      C     ---- FIRST CLEAR THE RANK ARRAYS ----
00
00            DO 20 I=1,IX
00              RANK1(I) = 0
00              RANK2(I) = 0
00      20    CONTINUE
00
00      C     ---- NOW GET THE RANKINGS FOR THE FIRST SET ----
00
00            IF(J .NE. 0)  THEN
00              DO 30 I = 1,J
00                RANK1(LINK1(I)) = I
00      30      CONTINUE
00
00            END IF
00
00      C     ---- NOW DO THE RANKINGS FOR THE SECOND SET ----
00
00            IF(K .NE. 0) THEN
00              DO 40 I = 1,K
00                RANK2(LINK2(I)) = I
00      40      CONTINUE
00
00            END IF
00
00      C     ---- ALL DONE ----
00
00            RETURN
00            END
00            START EDIT PAGE @ BUBBLE
00            SUBROUTINE  BUBBLE(NPOINT, RESULT, LINK)
00
00
00      C       THIS ROUTINE PRODUCES THE POINT ARRAY LINK WHICH ORDERED SO THAT
00      C       LINK(I) POINTS TO THE HIGHEST RESULT & SO ON, FOR ALL THE N
00      C     VALUES.
00
00
00      C        INPUTS
00      C        ======
00
00      C     NPOINT        THE NUMBER OF ELEMENTS IN THE INPUT ARRAYS
00
00      C     RESULT        THE ARRAY TO BE SORTED
00
00      C     LINK          ON ENTRY THIS ROUTINE MUST HAVE EACH ELEMENT
00      C                   POINTING TO ONE OF THE RESULTS TO BE SORTED.
00
00
00
00      C        OUTPUTS
00      C        =======
00
00      C     LINK          AN ARRAY THAT POINTS TO EACH OF THE MEMBERS OF
00      C                   THE RESULT ARRAY.  THE FIRST OF THE LINK
```

```
C                ELEMENTS, THE SECOND TO THE SECOND SMALLEST ELT.

C          DECLARATIONS
C          ============

         INTEGER     LINK(NPOINT), SWAP

         REAL        RESULT(NPOINT)

         LOGICAL   NOCHNG

         JSTART = 2
         JEND = NPOINT

         DO  20  I  =  2,NPOINT,2

           NOCHNG = .TRUE.

           DO  10  J  =  JSTART, JEND

             IF(RESULT(LINK(J-1)) .LT.  RESULT(LINK(J)))  THEN
C            ---- BUBBLE DOWN THE SMALL VALUES ----

               NOCHNG = .FALSE.
               SWAP   =  LINK(J)
               LINK(J)   =  LINK(J-1)
               LINK(J-1)   =   SWAP

             END IF

    10     CONTINUE

C          ---- TEST FOR ORDERED LIST ----

           IF(NOCHNG) RETURN

C           ---- SET UP THE LIMITS FOR THE NEXT SCAN ----


           JSWAP = JSTART
           JSTART = JEND - 2
           JEND = JSWAP - 1
           IF(JSTART .LT. JEND)  RETURN
           NOCHNG = .TRUE.

           NOCHNG = .TRUE.


           DO  15  J  =  JSTART, JEND, -1

             IF(RESULT(LINK(J+1)) .GT.  RESULT(LINK(J)))  THEN
C            ---- BUBBLE UP THE SMALL VALUES ----

               NOCHNG = .FALSE.
               SWAP   =  LINK(J)
               LINK(J)   =  LINK(J+1)
               LINK(J+1)   =   SWAP
```

```
   15   CONTINUE
C       ---- TEST FOR ORDERED LIST ----
        IF(NOCHNG) RETURN
C        ---- SET UP THE LIMITS FOR THE NEXT SCAN ----

        JSWAP = JSTART
        JSTART = JEND + 2
        JEND = JSWAP + 1
        IF(JSTART .GT. JEND)  RETURN
   20 CONTINUE
      RETURN
      END
```

END ELT. ERRORS: NONE. TIME: 1.385 SEC. IMAGE COUNT: 753

@RESUME,P PR

MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV
MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV
MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV MIKENV

APPENDIX 3

WARD'S AUTO-CORRELATION PROGRAM

```
10        DIMENSION IX(35,35),ZDIF(1226),TOT(1226),WW(1226),DEE(1226),
10       &          ISEL(5),TITLE(14)
14        DIMENSION KX(5000),KY(5000),JX(5000),JY(5000),FMTIN(14),FMTOUT(14)
10   C    ##################################################################
10   C    THIS PROGRAM CALCULATES THE DEGREE OF PAIRWISE SPATIAL
10   C    AUTOCORRELATION FOR VARIOUS CARDINAL DIRECTIONS ACROSS A SURFACE
10   C    THE PROGRAM CALCULATES A CORRELATION COEFFICIENT FIRSTLY ACCORDING
10   C    TO CLIFF AND ORD'S METHOD AND SECONDLY ACCORDING TO PEARSON'S
10   C    PRODUCT MOMENT METHOD. IN THE FORMER CASE SPATIAL DEPENDENCE IN
10   C    ALL DIRECTIONS IS CONSIDERED WHEREAS IN THE LATTER IT IS A SINGLE
10   C    DIRECTIONAL DEPENDENCE WHICH IS OF IMPORTANCE.
10   C    ##################################################################
10   C
10   C    ******************************************************************
10   C    IX REFERS TO THE OBSERVED VALUES
12   C   DIMENSIONED AS :(IROW,ICOL)
10   C    ZDIF REFERS TO THE DIFFERENCE BETWEEN THE OBSERVED AND THE MEAN
10   C    DIMENSIONED AS:((IROW*ICOL)+1)
10   C    TOT REFERS TO THE PRODUCT OF CURRENT ZDIF AND ADJACENT ZDIF
10   C    WW REFERS TO THE NUMBER OF 'JOINS' IN THE SYSTEM
10   C    DEE REFERS TO THE CALCULATION OF D, USED IN THE TEST FOR SIGNIFICANCE
10   C    TOT,WW,DEE,DIMENSIONED AS:(IROW*ICOL)
10   C    DIMENSIONS KX,KY,JX,JY, = THE TOTAL NUMBER OF POSSIBLE
10   C    X AND Y PAIRS IN THE SYSTEM. CALCULATED AS FOLLOWS:
10   C    (IROW*(ICOL-1))+((IROW-1)*ICOL)+(((IROW-1)*ICOL-1))*2)
10   C
10   C    THIS SEGMENT READS THE DATA
10   C    CARD 1:TITLE OF ANALYSIS
10   C           THIS IS PRINTED
10   C           (COLS 1-80 :(13A6,A2))
10   C    CARD 2:THE NUMBER OF ROWS OF DATA
10   C           THE NUMBER OF COLUMNS OF DATA
10   C           THE NUMBER OF DIRECTIONAL CHOICES(NUMM)
10   C           (COLS 1-15 :3I5)
10   C    CARD 3:THE DIRECTIONAL POSSIBILITIES(ISEL)
10   C           ANY COMBINATION OF 5 DIRECTIONS MAY BE CHOSEN,
10   C           IF <5,CHOICES MUST BE LEFT ALIGNED
10   C           NUMM VALUES OF ISEL ARE GIVEN BY:
10   C                ISEL(1)=A NORTH/SOUTH DIRECTION
10   C                    (2)=A WEST/EAST DIRECTION
10   C                    (3)=A NORTHEAST/SOUTHWEST DIRECTION
10   C                    (4)=A NORTHWEST/SOUTHEAST DIRECTION
10   C                    (5)=ALL DIRECTIONS,IE.THE WHOLE SURFACE
10   C          (COLS 1-25 :5I5)
10   C    CARD 4:VARIABLE DATA INPUT FORMAT
10   C           MUST BE INTEGER VALUES
10   C           E.G. :(10I5)
10   C           (COLS 1-80 :13A6,A2)
10   C    CARD 5:VARIABLE FORMAT FOR PRINTING INPUT MATRIX
10   C           E.G. :(10I5)
10   C           (COLS 1-80 :13A6,A2)
10   C    CARD 6 TO N :DATA
10   C                PUNCHED BY ROW
13   C
13   C
13   C    METHOD:CLIFF & ORD IN SCOTT,A.J.1969.STUDIES IN
13   C           REGIONAL SCIENCE.
13   C    WRITTEN BY:GARY WARD,RHODES UNIV.,1977
10   C    ******************************************************************
10        READ (7,3) (TITLE(J),J=1,14)
```

```
10          3 FORMAT(13A6,A2)
10            WRITE (5,5) (TITLE(J),J=1,14)
10          5 FORMAT(1H1 ,T10,'TITLE OF ANALYSIS:',13A6,A2///)
10            READ (8,1) IROW,ICOL,NUMM
10          1 FORMAT (3I5)
10            READ (8,2) (ISEL(J),J=1,NUMM)
10          2 FORMAT (5I5)
10            WRITE(5,4) IROW,ICOL,NUMM,(ISEL(K),K=1,5)
10          4 FORMAT (1H  ,T5,'IROW =',I3/
10           1            T5,'ICOL =',I3/
10           2            T5,'NUMM =',I3/
10           3            T5,'ISEL =',5I2////)
10            READ (7,6) (FMTIN(J),J=1,14)
10          6 FORMAT (13A6,A2)
10            READ (7,6) (FMTOUT(K),K=1,14)
10            DO 10 I=1,IROW
10         10 READ (8,FMTIN)(IX(I,J),J=1,ICOL)
10            WRITE(5,8)
10            DO 7 K=1,IROW
10          7 WRITE (5,FMTOUT) (IX(K,L),L=1,ICOL)
10          8 FORMAT (1H ,T5,'INPUT DATA MATRIX'/T5,17(1H-)///)
10      C     ------------------------------------------------------------------
10      C     THIS SEGMENT CALCULATES THE MEAN VALUE
10      C     USE ISUM AS AN ACCUMULATOR. NNNN IS A COUNTER WHICH COUNTS ALL
10      C     THOSE OBSERVATIONS THAT HAVE A VARIATE GREATER THAN ZERO.
10      C     A VALUE OF ZERO IN ANY GRID SQUARE INDICATES THAT THAT PARTICULAR
10      C     GRID SQUARE IS NOT INCLUDED IN CALCULATION OF THE COEFFICIENT.
10      C     NNNN, ICON, AND ISUM ARE SET TO ZERO.
10      C     ICON IS A COUNTER WHICH IS USED IN THE CALCULATION OF PEARSON'S R
10      C     MM IS A DIRECTIONAL COUNTER. SPATIAL AUTOCORRELATION IS
10      C     CALCULATED 5 TIMES FOR BOTH CLIFF AND ORD'S R AND PEARSON'S R
10      C     ------------------------------------------------------------------
10            MM=1
10            ICON=0
10         13 ISUM=0
10            NNNN=0
10      C     ------------------------------------------------------------------
10      C     SUM ALL THE VALUES GREATER THAN ZERO
10      C     ------------------------------------------------------------------
10            DO 20 J=1,ICOL
10            DO 20 I=1,IROW
10            IF(IX(I,J).EQ.0) GO TO 20
10            ISUM=ISUM+IX(I,J)
10            NNNN=NNNN+1
10         20 CONTINUE
10            MSUM=ISUM-NNNN
10      C     ------------------------------------------------------------------
10      C     DIVIDE ISUM BY THE NUMBER OF ELEMENTS IN THE MATRIX GREATER THAN
10      C     ZERO, VIZ. NNNN, TO GET THE MEAN.
10      C     ------------------------------------------------------------------
10            AVER=FLOAT(ISUM)/FLOAT(NNNN)
10      C     ------------------------------------------------------------------
10      C     THE RESULTS FOR COMPUTATION OF AUTOCORRELATION IN THE VARIOUS
10      C     DIRECTIONS FOR BOTH CLIFF AND ORD'S COEFFICIENT AND PEARSON'S
10      C     COEFFICIENT ARE PRINTED WITH TITLES.
10      C     ------------------------------------------------------------------
10            IF(ISEL(MM).GT.1) GO TO 211
10            WRITE (5,22)
10         22 FORMAT (1H1,1X,45HTHIS IS THE FIRST RUN. DIRECTION NORTH/SOUTH ,/,
10           12X,44(1H-),//)
10            GO TO 27
10        211 IF(ISEL(MM).GT.2) GO TO 212
```

```
10           WRITE (5,23)
10        23 FORMAT (1H1,1X,44HTHIS IS THE SECOND RUN.DIRECTION WEST/EAST ,/,2
10          1X,43(1H-),//)
10           GO TO 27
10       212 IF(ISEL(MM).GT.3) GO TO 213
10           WRITE (5,24)
10        24 FORMAT (1H1,1X,53HTHIS IS THE THIRD RUN. DIRECTION NORTHWEST/SOUTH
10          1EAST ,/,2X,52(1H-),//)
10           GO TO 27
10       213 IF(ISEL(MM).GT.4) GO TO 214
10           WRITE (5,25)
10        25 FORMAT (1H1,1X,54HTHIS IS THE FOURTH RUN. DIRECTION NORHTEAST/SOUT
10          1HWEST ,/,2X,53(1H-),//)
10           GO TO 27
10       214 WRITE (5,26)
10        26 FORMAT (1H1,1X,52HTHIS IS THE FIFTH RUN. WHOLE SURFACE,ALL DIRECTI
10          1ONS ,/,2X,51(1H-),//)
10        27 WRITE (5,28)
10        28 FORMAT (1H0,27H(A) CLIFF AND ORD'S METHOD ,/,1X,26(1H-),//)
10           WRITE (5,14)
10        14 FORMAT (1H0,3X,5HISUM ,5X,5HMEAN ,5X,15HCORRECTED ISUM ,/)
10           WRITE (5,15) ISUM,AVER,MSUM
10        15 FORMAT (2X,I6,2X,F8.2,12X,I8,//)
10     C     ------------------------------------------------------------------
10     C     THIS SEGMENT REFERS TO THE CLIFF AND ORD METHOD ONLY,
10     C     AND CALCULATES THE DIFFERENCE BETWEEN ALL THE
10     C     OBSERVED VALUES IX AND THE MEAN, THE ANSWER IN EACH CASE
10     C     BEING LABELLED ZDIF. IF THE OBSERVED VALUE IX IS ZERO.,
10     C     THEN THE ZDIF VALUE FOR THAT GRID SQUARE IS SET TO -999.0 FOR
10     C     IDENTIFICATION PURPOSES. IF THE ZDIF VALUE IS CALCULATED
10     C     AND THE RESULT IS ZERO,IE. ZDIF IS NATURALLY ZERO BECAUSE
10     C     IX=MEAN, THEN ZDIF IS SET TO +999.0, ALSO FOR IDENTIFICATION
10     C     PURPOSES. A COUNTER K IS INTRODUCED FOR THE CALCULATION OF ZDIF.
10     C     ------------------------------------------------------------------
10           K=0
10           SUMZ=0.0
10           SUMB=0.0
10           DO 30 I=1,IROW
10           DO 30 J=1,ICOL
10           IF(IX(I,J).GT.0) GO TO 33
10           K=K+1
10           ZDIF(K)=-999.0
10           GO TO 34
10        33 K=K+1
10           ZDIF(K)=FLOAT(IX(I,J))-AVER
10        34 IF(ZDIF(K).NE.-999.0) SUMZ=SUMZ+ZDIF(K)*ZDIF(K)
10           IF(ZDIF(K).NE.-999.0) SUMB=SUMB+ZDIF(K)*ZDIF(K)*ZDIF(K)*ZDIF(K)
10           IF(ZDIF(K).EQ.0.0) ZDIF(K)=999.0
10        30 CONTINUE
10     C     ------------------------------------------------------------------
10     C     SUMB IS THE SUM OF ZDIF TO THE POWER FOUR.USED IN CALCULATION OF
10     C     SKEWNESS
10     C     SUMZ IS THE TOTAL OF (IX-XMEAN) SQUARED.--REFER VARIANCE.
10     C     PRINT THE RESULTS WITH TITLE
10     C     ------------------------------------------------------------------
10           WRITE (5,31)
10        31 FORMAT (1H0,20HTHE ZDIF SCORES ARE ,/)
10           WRITE (5,32)(ZDIF(N),N=1,K)
10        32 FORMAT (16(1X,F6.1),//)
10     C     ------------------------------------------------------------------
10     C     THIS SEGMENT CREATES A GRID SQUARE OUTSIDE THE LIMITS OF THE MATRIX
10     C     NEND. REFER COMMENT STATEMENT BELOW
```

```
10    C    ------------------------------------------------------------------
10          NEND=IROW*ICOL
10          IC=NEND+1
10          ZDIF(IC)=0.0
10    C    ------------------------------------------------------------------
10    C    THIS SEGMENT ISOLATES ALL THE ADJACENT GRID SQUARES TO EVERY SINGLE
10    C    GRID SQUARE IN THE MATRIX. EVERY GRID SQUARE HAS 8 ADJACENT SQUARES
10    C    WHICH ARE CODED I1 TO I8. ZDIF SCORES IN ADJACENT GRID SQUARES ARE
10    C    MULTIPLIED.
10    C    ALL RESULTING VALUES ARE SUMMED.
10    C    WAIT IS THE TOTAL NUMBER OF 'JOINS' IN THE SYSTEM
10    C    DED IS A COUNTER FOR D, USED IN SIGNIFICANCE.
10    C    BOTH ARE SET TO ZERO. FOR EXPLANATION REFER BELOW.
10    C    TOTAL IS SET TO ZERO. FOR EXPLANATION OF TOTAL REFER BELOW.
10    C    ------------------------------------------------------------------
10          WAIT=0.0
10          DED=0.0
10          TOTAL=0.0
10          DO 50 J=1,NEND
10          I1=(J-ICOL)-1
10          I8=(J+ICOL)+1
10          I3=(J-ICOL)+1
10          I6=(J+ICOL)-1
10          I2=J-ICOL
10          I7=J+ICOL
10          I4=J-1
10          I5=J+1
10    C    ------------------------------------------------------------------
10    C    IF I1 TO I8 FALL OUTSIDE THE LIMITS OF THE MATRIX NEND THEY ARE ROUTED
10    C    TO IC, WHICH READS 0.0
10    C    ------------------------------------------------------------------
10          IF(J.GT.ICOL) GO TO 60
10          I1=IC
10          I2=IC
10          I3=IC
10       60 NB=NEND-ICOL
10          IF(J.LE.NB) GO TO 61
10          I6=IC
10          I7=IC
10          I8=IC
10       61 IF(J-ICOL*(J/ICOL).NE.0) GO TO 62
10          I3=IC
10          I5=IC
10          I8=IC
10       62 NL=(J/ICOL)*ICOL+1
10          IF(J.NE.NL) GO TO 63
10          I1=IC
10          I4=IC
10          I6=IC
10    C    ------------------------------------------------------------------
10    C    THE DIRECTIONAL LIMITS ARE NOW SET
10    C    FOR THE FIRST RUN N/S, FOR THE SECOND RUN E/W, FOR THE THIRD AND
10    C    FOURTH RUNS BOTH DIAGONALS, AND FOR THE FIFTH RUN THE WHOLE SURFACE.
10    C    I1 REFERS TO THE GRID SQUARE NORTHWEST OF THE CURRENT GRID
10    C    SQUARE. FURTHER, I2=N, I3=NE, I4=W, I5=E, I6=SW, I7=S, I8=SE,
10    C    OF THE CURRENT GRID SQUARE.
10    C    ------------------------------------------------------------------
10       63 IF(ISEL(MM).GT.1) GO TO 64
10          I1=IC
10          I4=IC
```

```
10          I6=IC
10          I3=IC
10          I5=IC
10          I8=IC
10          GO TO 67
10       64 IF(ISEL(MM).GT.2) GO TO 65
10          I1=IC
10          I2=IC
10          I3=IC
10          I6=IC
10          I7=IC
10          I8=IC
10          GO TO 67
10       65 IF(ISEL(MM).GT.3) GO TO 66
10          I2=IC
10          I3=IC
10          I4=IC
10          I5=IC
10          I6=IC
10          I7=IC
10          GO TO 67
10       66 IF(ISEL(MM).GT.4) GO TO 67
10          I1=IC
10          I2=IC
10          I4=IC
10          I5=IC
10          I7=IC
10          I8=IC
10    C     ------------------------------------------------------------------
10    C     T1 IS THE PRODUCT OF THE GRID SQUARE, ZDIF(J), MULTIPLIED
10    C     BY ADJACENT ZDIF(I1), ZDIF(I2), ZDIF(I3),... ZDIF(I8).
10    C     TOT(J) IS THE SUM OF THESE 8 PRODUCTS. THERE WILL BE
10    C     NNNN  TOT(J) VALUES.
10    C     TOTAL IS THE SUM OF ALL THE TOT(J) VALUES
10    C     THE RESULTS ARE PRINTED WITH TITLES,TOT VALUES, AND TOTAL.
10    C     BETWEEN STATEMENT NUMBERS 67 AND 626 GRID SQUARES HAVING A
10    C     NATURAL VALUE OF ZERO, AND THOSE WHICH WERE SET TO ZERO ARE
10    C     DIFFERENTIATED. JOINS BETWEEN ADJACENT GRID SQUARES ARE ALSO NOTED.
10    C     ------------------------------------------------------------------
10       67 ZZ=ZDIF(J)
10          IF(ZDIF(J).EQ.-999.0) ZDIF(J)=0.0
10          T1=0.0
10          S1=0.0
10          IF(ZDIF(J).NE.999.0.AND.ZDIF(I1).NE.999.0) GO TO 611
10          IF(ZDIF(I1).NE.0.0) S1=1.0
10          GO TO 612
10      611 IF(ZDIF(I1).NE.-999.0) T1=ZDIF(J)*ZDIF(I1)
10      612 T2=0.0
10          S2=0.0
10          IF(ZDIF(J).NE.999.0.AND.ZDIF(I2).NE.999.0) GO TO 613
10          IF(ZDIF(I2).NE.0.0) S2=1.0
10          GO TO 614
10      613 IF(ZDIF(I2).NE.-999.0) T2=ZDIF(J)*ZDIF(I2)
10      614 T3=0.0
10          S3=0.0
10          IF(ZDIF(J).NE.999.0.AND.ZDIF(I3).NE.999.0) GO TO 615
10          IF(ZDIF(I3).NE.0.0) S3=1.0
10          GO TO 616
10      615 IF(ZDIF(I3).NE.-999.0) T3=ZDIF(J)*ZDIF(I3)
10      616 T4=0.0
10          S4=0.0
10          IF(ZDIF(J).NE.999.0.AND.ZDIF(I4).NE.999.0) GO TO 617
```

```
10           IF(ZDIF(I4).NE.0.0) S4=1.0
10           GO TO 618
10       617 IF(ZDIF(I4).NE.-999.0) T4=ZDIF(J)*ZDIF(I4)
10       618 T5=0.0
10           S5=0.0
10           IF(ZDIF(J).NE.999.0.AND.ZDIF(I5).NE.999.0) GO TO 619
10           IF(ZDIF(I5).NE.0.0) S5=1.0
10           GO TO 620
10       619 IF(ZDIF(I5).NE.-999.0) T5=ZDIF(J)*ZDIF(I5)
10       620 T6=0.0
10           S6=0.0
10           IF(ZDIF(J).NE.999.0.AND.ZDIF(I6).NE.999.0) GO TO 621
10           IF(ZDIF(I6).NE.0.0) S6=1.0
10           GO TO 622
10       621 IF(ZDIF(I6).NE.-999.0) T6=ZDIF(J)*ZDIF(I6)
10       622 T7=0.0
10           S7=0.0
10           IF(ZDIF(J).NE.999.0.AND.ZDIF(I7).NE.999.0) GO TO 623
10           IF(ZDIF(I7).NE.0.0) S7=1.0
10           GO TO 624
10       623 IF(ZDIF(I7).NE.-999.0) T7=ZDIF(J)*ZDIF(I7)
10       624 T8=0.0
10           S8=0.0
10           IF(ZDIF(J).NE.999.0.AND.ZDIF(I8).NE.999.0) GO TO 625
10           IF(ZDIF(I8).NE.0.0) S8=1.0
10           GO TO 626
10       625 IF(ZDIF(I8).NE.-999.0) T8=ZDIF(J)*ZDIF(I8)
10       626 IF(ZDIF(J).EQ.0.0) ZDIF(J)=ZZ
10           TOT(J)=T1+T2+T3+T4+T5+T6+T7+T8
10           TOTAL=TOTAL+TOT(J)
10     C     ------------------------------------------------------------
10     C     FOR PURPOSES OF CALCULATION INSIDE LOOP NUMBER 50 WAIT=WW
10     C     CLEARLY WW WILL CHANGE ACCORDING TO DIRECTION
10     C     WW=WAIT ONCE AGAIN OUTSIDE OF THE LOOP
10     C     WW(J) IS A COUNTER WHICH IS SET TO ZERO.
10     C     IF EITHER THE CURRENT GRID SQUARE OR AN ADJACENT GRID SQUARE
10     C     REFLECTS A PRESENT VALUE OF ZERO THEN THERE IS NO 'JOIN' AND THAT
10     C     PARTICULAR X AND Y PAIR IS NOT INCLUDED IN THE CALCULATION OF THE
10     C     COEFFICIENT. DEE(J) IS A COUNTER FOR D.
10     C     ------------------------------------------------------------
10           WW(J)=S1+S2+S3+S4+S5+S6+S7+S8
10           IF(T1.EQ.0.0) GO TO 601
10           WW(J)=WW(J)+1.0
10       601 IF(T2.EQ.0.0) GO TO 602
10           WW(J)=WW(J)+1.0
10       602 IF(T3.EQ.0.0) GO TO 603
10           WW(J)=WW(J)+1.0
10       603 IF(T4.EQ.0.0) GO TO 604
10           WW(J)=WW(J)+1.0
10       604 IF(T5.EQ.0.0) GO TO 605
10           WW(J)=WW(J)+1.0
10       605 IF(T6.EQ.0.0) GO TO 606
10           WW(J)=WW(J)+1.0
10       606 IF(T7.EQ.0.0) GO TO 607
10           WW(J)=WW(J)+1.0
10       607 IF(T8.EQ.0.0) GO TO 608
10           WW(J)=WW(J)+1.0
10       608 WAIT=WAIT+WW(J)
10           DEE(J)=WW(J)*(WW(J)-1.0)
10           DED=DED+DEE(J)
10           D=DED/2.0
10        50 CONTINUE
```

```
      WRITE (5,51)
   51 FORMAT(1H0,19HTHE TOT VALUES ARE ,/)
      WRITE (5,32)(TOT(K),K=1,NEND)
      WRITE (5,52) TOTAL
   52 FORMAT (1H0,6HTOTAL ,F12.4,//)
C     ------------------------------------------------------------------
C     THIS SEGMENT CALCULATES THE AUTOCORRELATION COEFFICIENT ACCORDING
C     TO CLIFF AND ORD'S METHOD.
C     IF THE VARIANCE, SUMZ, IS ZERO, CORRELATION IS SET TO ZERO
C     THE VALUES FOR NEND, WAIT, SUMZ, SUMB, AND CORRELATION ARE
C     PRINTED WITH TITLES
C     --------------------------------------------------------------------
      IF (SUMZ.NE.0.0) GO TO 56
      CORR=0.0
      GO TO 57
   56 CORR=(FLOAT(NNNN)*TOTAL)/(WAIT*SUMZ)
   57 WRITE (5,70)
   70 FORMAT (2X,5HNNNN ,2X,5HNEND ,9X,5HWAIT ,15X,5HSUMZ ,15X,5HSUMB ,/
     1)
      WRITE (5,71) NNNN,NEND,WAIT,SUMZ,SUMB
   71 FORMAT (1X,I5,2X,I5,F12.2,2(2X,F30.2),//)
      WRITE (5,72) CORR
   72 FORMAT (1H0,31HTHE CORRELATION COEFFICIENT IS ,F12.6,/,1X,30(1H-),
     1//)
C     --------------------------------------------------------------------
C     THIS SEGMENT CALCULATES THE SIGNIFICANCE OF CLIFF AND ORD'S
C     COEFFICIENT.
C     THERE ARE NO FIXED LIMITS FOR THIS COEFFICIENT ANF THEREFORE
C     SIGNIFICANCE IS EVALUATED FROM THE VALUE FOR R, PLUS THE MEAN
C      AND VARIANCE OF THE DATA.  H1 IS ACCEPTED ONLY IF H0 (HYPOTHESIS
C     OF NO CORRELATION) IS REJECTED AT THE 95% LEVEL. SIGNIFICANCE
C     IS EVALUATED ACCORDING TO BOTH THE ASSUMPTIONS OF A RANDOMLY
C     DISTRIBUTED XI AND A NORMALLY DISTRIBUTED XI.
C     WHEN JOINS BETWEEN ADJACENT GRID SQUARES ARE NOT WEIGHTED, A
C     SYSTEM OF BINARY WEIGHTS CAN BE USED WHERE A=2A, S1=4A, AND
C     S2=8(A+D). (REFER CLIFF AND ORD, 1973, PAGE 16.)
C     A=HALF THE SUM OF LI WHERE LI=THE NUMBER OF GRID SQUARES
C     ADJIONING GRID SQUARE I. 'A' IS  THEREFORE W/2.
C     D=HALF THE SUM OF LI(LI-1)
C     SONE=4A,   STWO=8(A+D)
C     SKEW=SAMPLE COEFFICIENT OF KURTOSIS, IE. THE FOURTH SAMPLE MOMENT
C     OF X DIVIDED BY (THE SECOND SAMPLE MOMENT OF X) SQUARED.
C     EXPS=EXPECTED VALUE OF THE COEFFICIENT UNDER ASSUMPTION OF RANDOMNESS
C     EXSP=EXPECTED VALUE OF THE COEFFICIENT UNDER ASSUMPTION OF NORMALITY
C     PONE,PTWO, AND PTHR ARE SECTIONS OF THE FORMULA FOR EXPS
C     FRED=ABSOLUTE VALUE OF (CORRELATION PLUS THE FIRST MOMENT OF X
C     DIVIDED BY THE SQUARE ROOT OF EXPS-FIRST MOMENT SQUARED)
C     TOOK AND TALK ARE SECTIONS OF THE FORMULA FOR EXSP
C     FINA=ABSOLUTE VALUE OF THE FORMULA USED IN FRED, USING EXSP INSTEAD
C     OF EXPS.
C     IF FRED OR FINA ARE GREATER THAN 1.96, WHICH IS THE VALUE OF TWO
C     STANDARD DEVIATIONS FROM THE MEAN UNDER A STANDARD NORMAL CURVE,
C     THEN THE H0 IS REJECTED IN FAVOUR OF H1
C     RESULTS FOR D, A, S1, S2, B2, PONE, PTWO, PTHR, PLUS BOTH EXPECTED
C     AND ABSOLUTE VALUES ARE PRINTED WITH TITLES.
C     --------------------------------------------------------------------
   77 XN=FLOAT(NNNN)
      XW=WAIT
      A=WAIT/2.0
      SONE=4.0*A
      STWO=(8.0*(A+D))
      SKEW=((SUMB/XN)/((SUMZ/XN)**2.0))
```

```
10      771 FORMAT (1X,15HTHE D-VALUE IS ,F20.4,/)
10          WRITE (5,772) A
10      772 FORMAT (1X,15HTHE A-VALUE IS ,F20.4,/)
10          WRITE (5,773) SONE
10      773 FORMAT (1X,18HTHE SONE-VALUE IS ,F20.4,/)
10          WRITE (5,774) STWO
10      774 FORMAT (1X,18HTHE STWO-VALUE IS ,F20.4,/)
10          WRITE (5,775) SKEW
10      775 FORMAT (1X,31HTHE VALUE FOR SKEWNESS (B2) IS ,F10.4,//)
10          PONE=(XN*((((XN*XN)-(3.0*XN)+3.0)*SONE)-(XN*STWO)+(3.0*(XW*XW))))
10          PTWO=(SKEW*((((XN*XN)-XN)*SONE)-((2.0*XN)*STWO)+(6.0*(XW*XW))))
10          PTHR=((XN-1.0)*(XN-2.0)*(XN-3.0)*(XW*XW))
10          EXPS=((PONE-PTWO)/PTHR)
10          WRITE (5,111) PONE
10      111 FORMAT (1H0,8HPONE IS ,F28.4,/)
10          WRITE (5,112) PTWO
10      112 FORMAT (1X,8HPTWO IS ,F28.4,/)
10          WRITE (5,113) PTHR
10      113 FORMAT (1X,8HPTHR IS ,F28.4,//)
10          IF(EXPS.GT.0.0) GO TO 777
10          WRITE (5,776)
10      776 FORMAT (1H0,46HTHE EXPECTED VALUE IS NEGATIVE.RUN TERMINATED ,//)
10          GO TO 83
10      777 WRITE (5,778) EXPS
10      778 FORMAT (2X,37HTHE EXPECTED VALUE ASSUMING IX TO BE ,/,1X,
10         138HDRAWN FROM A NON-NORMAL POPULATION IS ,F16.10,//)
10          ROSE=(EXPS-((1.0/(XN-1.0))**2.0))
10          IF(ROSE.GT.0.0) GO TO 800
10          WRITE (5,779)
10      779 FORMAT (1X,44HVALUE UNDER ROOT IS NEGATIVE.RUN TERMINATED ,//)
10          GO TO 83
10      800 ROOT=(ROSE**0.5)
10          FRED=ABS((CORR+(1.0/(XN-1.0)))/ROOT)
10          WRITE (5,90) FRED
10       90 FORMAT (1X,22HTHE ABSOLUTE VALUE IS ,F16.10,/,1X,21(1H-),//)
10          TOOK=(((XN*XN)*SONE)-(XN*STWO)+(3.0*(XW*XW)))
10          TALK=(((XN*XN)-1.0)*(XW*XW))
10          EXSP=TOOK/TALK
10          IF(EXSP.GT.0.0) GO TO 91
10          WRITE (5,776)
10          GO TO 83
10       91 WRITE (5,92) EXSP
10       92 FORMAT (1X,37HTHE EXPECTED VALUE ASSUMING IX TO BE ,/,1X,
10         134HDRAWN FROM A NORMAL POPULATION IS ,F16.10,//)
10          SAMM=(EXSP-((1.0/(XN-1.0))**2.0))
10          IF(SAMM.GT.0.0) GO TO 93
10          WRITE (5,779)
10          GO TO 83
10       93 RUUT=(SAMM**0.5)
10          FINA=ABS((CORR+(1.0/(XN-1.0)))/RUUT)
10          WRITE (5,94) FINA
10       94 FORMAT (1X,22HTHE ABSOLUTE VALUE IS ,F16.10,/,1X,21(1H-),//)
10    C     -----------------------------------------------------------------
10    C     PEARSON'S METHOD
10    C     THIS SEGMENT CALCULATES AN AUTOCORRELATION COEFFICIENT IMPLEMENTING
10    C     PEARSON'S PRODUCT MOMENT METHOD. THE DIRECTIONS REMAIN THE SAME,
10    C     IE. 1=N/S, 2=W/E, AND SO ON.
10    C     PAIRS OF ADJOINING GRID SQUARES ARE ISOLATED. THE VARIATE VALUE IN
10    C     GRID SQUARE J BECOMES THE X-VALUE, WHILE THE VARIATE VALUE IN
10    C     GRID SQUARES 17,15,18, AND 16 RESPECTIVELY BECOME THE Y-VALUE.
10    C     DEPENDENCE IS THEREFORE MEASURED IN ONE DIRECTION SPECIFICALLY,
```

```
10  C        IE. I=N TO S, 2=E TO W, 3=NW TO SE, AND 4=NE TO SW.
10  C        X AND Y PAIRS FOR EACH OF THESE RUNS ARE ACCUMULATED UNDER COUNTER
10  C        ICON AND ARE ALL USED FOR THE FINAL OPTION OF ALL DIRECTIONS.
10  C        HOAEVER, IF EITHER X OR Y = ZERO THEN THAT PAIR IS NOT USED IN
10  C        THE CALCULATION OF THE COEFFICIENT.
10  C        SOME INDICATION OF AUTOCORRELATION OVER THE ENTIRE SURFACE IS THUS
10  C        GAINED. ONCE THE APPROPRIATE X AND Y  PAIRS HAVE BEEN ISOLATED
10  C        FOR EACH DIRECTIONAL OPTION A SUBROUTINE IS CALLED FOR THE
10  C        CALCULATION OF PEARSON'S R.
10  C        ------------------------------------------------------------------
10           IPER=0
10        83 IF(ISEL(MM).GT.1) GO TC 799
10           K=1
10           DO 911 J=1,ICOL
10           K=K-1
10           L=0
10           DO 911 I=1,IROW
10           K=K+1
10           L=I+1
10           IF(L.GT.IROW) GO TO 911
10           IF(IX(I,J).GT.0) GO TO 901
10           K=K-1
10           GO TO 911
10       901 IF(IX(L,J).GT.0) GO TO 902
10           K=K-1
10           GO TO 911
10       902 IPER=IPER+1
10           KX(K)=IX(I,J)
10           KY(K)=IX(L,J)
10           ICON=ICON+1
10           JX(ICON)=KX(K)
10           JY(ICON)=KY(K)
10       911 CONTINUE
10           CALL PROD(KX,KY,IPER)
10           GO TO 999
10       799 IF(ISEL(MM).GT.2) GO TO 801
10           K=1
10           DO 912 I=1,IROW
10           K=K-1
10           L=0
10           DO 912 J=1,ICOL
10           K=K+1
10           L=J+1
10           IF(L.GT.ICOL) GO TO 912
10           IF(IX(I,J).GT.0) GO TO 903
10           K=K-1
10           GO TO 912
10       903 IF(IX(I,L).GT.0) GO TO 904
10           K=K-1
10           GO TO 912
10       904 IPER=IPER+1
10           KX(K)=IX(I,J)
10           KY(K)=IX(I,L)
10           ICON=ICON+1
10           JX(ICON)=KX(K)
10           JY(ICON)=KY(K)
10       912 CONTINUE
10           CALL PROD(KX,KY,IPER)
10           GO TO 999
10       801 IF(ISEL(MM).GT.3) GO TO 802
10           K=1
10           KEND=IROW-1
```

```
10          K=K-I
10          DO 913 J=I,ICOL
10          K=K+I
10          II=I+1
10          JJ=J+1
10          IF(JJ.GT.ICOL) GO TO 913
10          IF(IX(I,J).GT.0) GO TO 905
10          K=K-1
10          GO TO 913
10      905 IF(IX(II,JJ).GT.0) GO TO 906
10          K=K-1
10          GO TO 913
10      906 IPER=IPER+1
10          KX(K)=IX(I,J)
10          KY(K)=IX(II,JJ)
10          ICON=ICON+1
10          JX(ICON)=KX(K)
10          JY(ICON)=KY(K)
10      913 CONTINUE
10          CALL PROD(KX,KY,IPER)
10          GO TO 999
10      802 IF(ISEL(MM).GT.4) GO TO 803
10          K=0
10          LEND=IROW-1
10          DO 914 I=1,LEND
11          DO 915 J=2,ICOL
10          K=K+1
10          IN=I+1
10          JN=J-1
10          IF(IX(I,J).GT.0) GO TO 907
10          K=K-1
10          GO TO 914
10      907 IF(IX(IN,JN).GT.0) GO TO 908
10          K=K-1
10          GO TO 914
10      908 IPER=IPER+1
10          KX(K)=IX(I,J)
10          KY(K)=IX(IN,JN)
10          ICON=ICON+I
10          JX(ICON)=KX(K)
10          JY(ICON)=KY(K)
11      915 CONTINUE
10      914 CONTINUE
10          WRITE (5,940) IPER
10      940 FORMAT (1H ,T5,'IPER IS:',I5/)
10          CALL PROD(KX,KY,IPER)
10          GO TO 999
10      803 WRITE(5,939) ICON
10      939 FORMAT(1H ,T5,'ICON IS:',I5/)
10          CALL PROD(JX,JY,ICON)
10      999 IF(MM.EQ.NUMM) GO TO 79
10          MM=MM+1
10          GO TO 13
10       79 STOP
10          END
```

ID ELT. ERRORS: NONE. TIME: 1.582 SEC. IMAGE COUNT: 612

IDG,P ****** AUTO-PROD ****** .L,0

```
@ELT,L GEOG*PROGS.AUTO-PROD
ELT 8R1 S74Q1C 03/28/80 09:32:16 (8)
 04             SUBROUTINE PROD(LX,LY,IPER)
 08             DIMENSION LX(5000),LY(5000),X(5000),Y(5000)
 04      C      ##############################################################
 04      C      SUBROUTINE CALCULATES R ACCORDING TO THE PRODUCT MOMENT METHOD
 04      C      DIMENSIONS REFLECT THE POSSIBLE TOTAL NUMBER OF X AND Y PAIRS IN
 04      C      THE GRID. IPER IS THE NUMBER OF POSSIBLE Z AND Y PAIRS IN ONE
 07      C      DIRECTION WHILE ICON REFLECTS ALL JOINS.
 04      C      TOPP, VARX, ANDVARY ARE SECTIONS OF THE EQUATION. SIGNIFICANCE
 04      C      IS EVALUATED ACCORDING TO STUDENTS T TEST.
 04      C      RESULTS ARE PRINTED WITH TITLES
 04      C      ##############################################################
 04             DO 921 I=1,IPER
 04             X(I)=FLOAT(LX(I))
 04         921 Y(I)=FLOAT(LY(I))
 04             SXTY=0.0
 04             SXSQ=0.0
 04             SYSQ=0.0
 04             XSUM=0.0
 04             YSUM=0.0
 04             DO 922 I=1,IPER
 04             XVAL=X(I)
 04             YVAL=Y(I)
 04             XSUM=XSUM+XVAL
 04             YSUM=YSUM+YVAL
 04             XSQ=XVAL*XVAL
 04             YSQ=YVAL*YVAL
 04             SXSQ=SXSQ+XSQ
 04             SYSQ=SYSQ+YSQ
 04             XTY=XVAL*YVAL
 04             SXTY=SXTY+XTY
 04         922 CONTINUE
 04             TOPP=SXTY-((XSUM*YSUM)/FLOAT(IPER))
 04             VARX=SXSQ-((XSUM*XSUM)/FLOAT(IPER))
 04             VARY=SYSQ-((YSUM*YSUM)/FLOAT(IPER))
 04             RUTE=(VARX*VARY)**0.5
 04             COCO=TOPP/RUTE
 04             COEF=ABS(COCO)
 04             NODF=IPER-2
 04             AWON=FLOAT(NODF)
 04             ONE=COEF*(AWON**0.5)
 04             IF(COEF.LE.1.0) GO TO 940
 04             TWO=0.0
 04             GO TO 941
 04         940 TWO=((1.0-(COEF*COEF))**0.5)
 04         941 IF(TWO.NE.0.0) GO TO 928
 04             TEST=-99.0
 04             GO TO 929
 04         928 TEST=ONE/TWO
 04         929 WRITE (5,930)
 04         930 FORMAT (1H0,36H(B) PEARSON'S PRODUCT MOMENT METHOD ,/,1X,35(1H-),/
 04            1/)
 04             WRITE (5,931) IPER
 04         931 FORMAT (1X,31HTHE NUMBER OF X AND Y PAIRS IS ,I12,/)
 04             WRITE (5,932) TOPP
 04         932 FORMAT (1X,31HPART ONE OF THE COMPUTATION IS ,F20.4,/)
 04             WRITE (5,933) RUTE
 04         933 FORMAT (1X,31HPART TWO OF THE COMPUTATION IS ,F20.4,//)
 04             WRITE (5,934) COCO
 04         934 FORMAT (1H0,31HTHE CORRELATION COEFFICIENT IS ,F18.6,/,1X,30(1H-),
 04            1//)
```

```
04          WRITE (5,937)
04      937 FORMAT (1X,31HTHE T-VALUE APPROACHES INFINITY ,/)
04          GO TO 938
04      936 WRITE (5,935) TEST
04      935 FORMAT (1X,24HTHE STUDENTS T VALUE IS ,F20.4,/)
04      938 RETURN
04          END

END ELT.  ERRORS: NONE. TIME:  0.350 SEC. IMAGE COUNT: 68

@HDG,P  ******  BEALE-ABS  ******  .L,0
```

```
         1SFL15)
      DIMENSION KX(5000),KY(5000),IX(5000),JY(5000),FMTIN(14)
      DIMENSION FMAT(14),FMATI(14)
      CHARACTER TITLE*50
C     ###################################################################
C     THIS PROGRAM CALCULATES THE DEGREE OF PAIRWISE SPATIAL AUTOCORRELATION
C     ###################################################################
C
C     *******************************************************************
C     IX REFERS TO THE OBSERVED VALUES
C     DIMENSIONED AS :(IROW,ICOL)
C     ZDIF REFERS TO THE DIFFERENCE BETWEEN THE OBSERVED AND THE MEAN
C     TOT REFERS TO THE PRODUCT OF CURRENT ZDIF AND ADJACENT ZDIF
C      VALUES MULTIPLIED BY THE WEIGHT VALUE FOR (I,J).
C     WN REFERS TO THE NUMBER OF 'JOINS' IN THE SYSTEM
C      D IS USED IN THE TEST FOR SIGNIFICANCE
C
C     THIS SEGMENT READS THE DATA
C     CARD 1:TITLE OF ANALYSIS
C               THIS IS PRINTED
C               (COLS 1-80 :(13A6,A2))
C     CARD 2:THE NUMBER OF ROWS OF DATA
C     CARD 4:VARIABLE DATA INPUT FORMAT
C               MUST BE INTEGER VALUES
C               E.G. :(10I5)
C               (COLS 1-80 :13A6,A2)
C     CARD 5:VARIABLE FORMAT FOR PRINTING INPUT MATRIX
C               E.G. :(10I5)
C               (COLS 1-80 :13A6,A2)
C     CARD 6 TO N :DATA
C                    PUNCHED BY ROW
C
C
C     METHOD:CLIFF & ORD IN SCOTT,A.J.1969.STUDIES IN
C               REGIONAL SCIENCE.
C     WRITTEN BY:GARY WARD,RHODES UNIV.,1977
C     ADAPTED TO INCLUDE THE WEIGHTING FACTOR
C     (CLIFF & ORD,1973) BY IAN COOK ACCORDING TO SPECIFICATIONS
C    BY ROS MORRIS, U.C.T.,1980.
C     *******************************************************************
      READ(7,3)TITLE
    3 FORMAT(A50)
      WRITE(5,5)TITLE
      READ (7,1) IROW
      WRITE(5,1)IROW
    5 FORMAT(1H1 ,T10,'TITLE OF ANALYSIS:',A50///)
      READ(8,888)(FMAT(I),I=1,13)
  888 FORMAT(I2,2X,12A6,A2)
      READ(11,FMAT) (IX(I),I=1,IROW)
    1 FORMAT (1H ,5X,'NUMBER OF ROWS IS',I5)
      WRITE(5,8)
      WRITE(5,FMAT)(IX(I),I=1,IROW)
    8 FORMAT (1H ,T5,'INPUT DATA MATRIX'/T5,17(1H-)//)
C     -------------------------------------------------------------------
C     THIS SEGMENT CALCULATES THE MEAN VALUE
C     USE ISUM AS AN ACCUMULATOR. NNNN IS A COUNTER WHICH COUNTS ALL
C     THOSE OBSERVATIONS THAT HAVE A VARIATE GREATER THAN ZERO.
C     A VALUE OF ZERO IN ANY GRID SQUARE INDICATES THAT THAT PARTICULAR
C     GRID SQUARE IS NOT INCLUDED IN CALCULATION OF THE COEFFICIENT.
C     NNNN AND ISUM ARE SET TO ZERO.
   13 ISUM=0
```

```
C     ------------------------------------------------------------------------
C      SUM ALL THE VALUES GREATER THAN ZERO
C     ------------------------------------------------------------------------
       DO 20 I=1,IROW
       IF(IX(I).EQ.0.0) GO TO 20
       IF(IX(I).EQ.-1.0) GOTO 20
       ISUM=ISUM+IX(I)
       NNNN=NNNN+1
    20 CONTINUE
       MSUM=ISUM-NNNN
       READ(8,888)(FMAT1(I),I=1,13)
       DO 12 I=1,IROW
    12 READ(12,FMAT1) (WTVAL(I,J),J=1,IROW)
       WRITE(5,9)
     9 FORMAT(1H ,I5,'TABLE OF WEIGHTED VALUES'/T5,25(1H-)//)
       DO 24 I=1,IROW
    24 WRITE(5,104) (WTVAL(I,J),J=1,IROW)
  104  FORMAT(26F5.3/7F5.3)
C     ------------------------------------------------------------------------
C      DIVIDE ISUM BY THE NUMBER OF ELEMENTS IN THE MATRIX GREATER THAN
C      ZERO, VIZ. NNNN, TO GET THE MEAN (AVER).
C     ------------------------------------------------------------------------
       AVER=FLOAT(ISUM)/FLOAT(NNNN)
       SUMZ=0.0
       SUMB=0.0

       WRITE(5,444)
  444  FORMAT(1H ,//5X,'TABLE OF TOT VALUES',/
      &5X,'*********************',//)
        DO 11 I=1,IROW
        AGGREG(I)=FLOAT(IX(I))-AVER
        DO 23 J=1,IROW
        IF (WTVAL(I,J).EQ.0.0)GOTO 23
        IF (WTVAL(I,J).EQ.-1.0) GOTO 23
        AGG(J)=FLOAT(IX(J))-AVER
        TOT=AGGREG(I)*AGG(J)*(WTVAL(I,J))
        TOTAL=TOTAL+TOT
        WRITE(5,156)I,J,TOT
   156 FORMAT(1H ,'TOT VALUE CELL',I2,'   WITH CELL',I2,3X,'=',F11.2)
    23 CONTINUE
    11 CONTINUE
        WRITE(5,234) TOTAL
  234   FORMAT(F15.4)
C     ------------------------------------------------------------------------
C       THIS SEGMENT CALCULATES THE AUTOCORRELATION COEFFICIENT ACCORDING
C       TO CLIFF AND ORD'S METHOD.
C       SUMB IS THE SUM OF ZDIF TO THE POWER FOUR.USED IN CALCULATION OF
C       SKEWNESS
C       SUMZ IS THE TOTAL OF (IX-XMEAN) SQUARED.--REFER VARIANCE.
C       PRINT THE RESULTS WITH TITLE
C      RYT IS AN ACCUMULATOR FOR THE CALCULATION OF SUMZ AND SUMB
C      JOINS REFERS TO THE NUMBER OF JOINS IN THE SYSTEM. ALSO
C      REFERRED TO AS WAIT FOR TEST OF SIGNIFICANCE.
C     ------------------------------------------------------------------------
C       IF THE VARIANCE, SUMZ, IS ZERO, CORRELATION IS SET TO ZERO
C       THE VALUES FOR NEND, WAIT, SUMZ, SUMB, AND CORRELATION ARE
C       PRINTED WITH TITLES
        DO 46 I=1,IROW
        RYT=(FLOAT(IX(I))-AVER)
        SUMZ=SUMZ+RYT**2
        SUMB=SUMB+RYT**4
```

```
      JOINS=0
      DO 56 I=1,IROW
      DO 57 J=1,IROW
      IF(WTVAL(I,J).EQ.0) GOTO 57
      IF(WTVAL(I,J).EQ.-1.)GOTO 57
      JOINS=JOINS+1
   57 CONTINUE
   56 CONTINUE
      CORR=0.0
      CORR=(FLOAT(NNNN)*TOTAL)/(FLOAT(JOINS)*SUMZ)
      WRITE(5,70)
   70 FORMAT (2X,5HNNNN ,9X,15X,5HSUMZ ,15X,5HSUME ,5HJOINS/
     1)
      WRITE(5,71) NNNN,SUMZ,SUMP,JOINS
   71 FORMAT (1X,I5,2(2X,F15.2),I5,//)
      WRITE(5,72) CORR
   72 FORMAT (1H0,31HTHE CORRELATION COEFFICIENT IS ,F12.6,/,1X,30(1H-),
     1//)
C     ------------------------------------------------------------------
C     THIS SEGMENT CALCULATES THE SIGNIFICANCE OF CLIFF AND ORD'S
C     COEFFICIENT.
C     THERE ARE NO FIXED LIMITS FOR THIS COEFFICIENT AND THEREFORE
C     SIGNIFICANCE IS EVALUATED FROM THE VALUE FOR R, PLUS THE MEAN
C      AND VARIANCE OF THE DATA.  H1 IS ACCEPTED ONLY IF H0 (HYPOTHESIS
C     OF NO CORRELATION) IS REJECTED AT THE 95% LEVEL. SIGNIFICANCE
C     IS EVALUATED ACCORDING TO BOTH THE ASSUMPTIONS OF A RANDOMLY
C     DISTRIBUTED XI AND A NORMALLY DISTRIBUTED XI.
C     WHEN JOINS BETWEEN ADJACENT GRID SQUARES ARE NOT WEIGHTED, A
C     SYSTEM OF BINARY WEIGHTS CAN BE USED WHERE W=2A, S1=4A, AND
C     S2=8(A+D). (REFER CLIFF AND ORD, 1973, PAGE 16.)
C     A=HALF THE SUM OF LI WHERE LI=THE NUMBER OF GRID SQUARES
C     ADJOINING GRID SQUARE I. 'A' IS  THEREFORE W/2.
C     D=HALF THE SUM OF LI(LI-1)
C     SONE=4A,   STWO=8(A+D)
C     SKEW=SAMPLE COEFFICIENT OF KURTOSIS, IE. THE FOURTH SAMPLE MOMENT
C     OF X DIVIDED BY (THE SECOND SAMPLE MOMENT OF X) SQUARED.
C     EXPS=EXPECTED VALUE OF THE COEFFICIENT UNDER ASSUMPTION OF RANDOMNESS
C     EXSP=EXPECTED VALUE OF THE COEFFICIENT UNDER ASSUMPTION OF NORMALITY
C     PONE,PTWO, AND PTHR ARE SECTIONS OF THE FORMULA FOR EXPS
C     FRED=ABSOLUTE VALUE OF (CORRELATION PLUS THE FIRST MOMENT OF X
C     DIVIDED BY THE SQUARE ROOT OF EXPS-FIRST MOMENT SQUARED).
C     TOOK AND TALK ARE SECTIONS OF THE FORMULA FOR EXSP
C     FINA=ABSOLUTE VALUE OF THE FORMULA USED IN FRED, USING EXSP INSTEAD
C     OF EXPS.
C     IF FRED OR FINA ARE GREATER THAN 1.96, WHICH IS THE VALUE OF TWO
C     STANDARD DEVIATIONS FROM THE MEAN UNDER A STANDARD NORMAL CURVE,
C     THEN THE H0 IS REJECTED IN FAVOUR OF H1
C     RESULTS FOR D, A, S1, S2, B2, PONE, PTWO, PTHR, PLUS BOTH EXPECTED
C     AND ABSOLUTE VALUES ARE PRINTED WITH TITLES.
C     ------------------------------------------------------------------
      WAIT=FLOAT(JOINS)
   77 XN=FLOAT(NNNN)
      XW=WAIT
      A=WAIT/2.0
      D=(WAIT*(WAIT-1))/2.
      SONE=4.0*A
      STWO=(8.0*(A+D))
      SKEW=((SUMB/XN)/((SUMZ/XN)**2.0))
      WRITE(5,771) D
  771 FORMAT (1X,15HTHE D-VALUE IS ,F20.4,/)
      WRITE(5,772) A
```

```
  772 FORMAT (1X,15HTHE A-VALUE IS ,F20.4,/)
      WRITE(5,773) SONE
  773 FORMAT (1X,16HTHE SONE-VALUE IS ,F20.4,/)
      WRITE(5,774) STWO
  774 FORMAT (1X,16HTHE STWO-VALUE IS ,F20.4,/)
      WRITE(5,775) SKEW
  775 FORMAT (1X,31HTHE VALUE FOR SKEWNESS (B2) IS ,F10.4,//)
      PONE=(XN*((((XN*XN)-(3.0*XN)+3.0)*SONE)-(XN*STWO)+(3.0*(XW*XW))))
      PTWO=(SKEW*((((XN*XN)-XN)*SONE)-((2.0*XN)*STWO)+(6.0*(XW*XW))))
      PTHR=((XN-1.0)*(XN-2.0)*(XN-3.0)*(XW*XW))
      EXPS=((PONE-PTWO)/PTHR)
      WRITE(5,111) PONE
  111 FORMAT (1H0,8HPONE IS ,F28.4,/)
      WRITE(5,112) PTWO
  112 FORMAT (1X,8HPTWO IS ,F28.4,/)
      WRITE(5,113) PTHR
  113 FORMAT (1X,8HPTHR IS ,F28.4,//)
      IF(EXPS.GT.0.0) GO TO 777
      WRITE(5,776)
  776 FORMAT (1H0,46HTHE EXPECTED VALUE IS NEGATIVE.RUN TERMINATED ,//)
  777 WRITE(5,778) EXPS
  778 FORMAT (2X,37HTHE EXPECTED VALUE ASSUMING IX TO BE ,/,1X,
     138HDRAWN FROM A NON-NORMAL POPULATION IS ,F16.10,//)
      ROSE=(EXPS-((1.0/(XN-1.0))**2.0))
      IF(ROSE.GT.0.0) GO TO 800
      WRITE(5,779)
  779 FORMAT (1X,44HVALUE UNDER ROOT IS NEGATIVE.RUN TERMINATED ,//)
  800 ROOT=(ROSE**0.5)
      FRED=ABS((CORR+(1.0/(XN-1.0)))/ROOT)
      WRITE(5,90) FRED
   90 FORMAT (1X,22HTHE ABSOLUTE VALUE IS ,F16.10,/,1X,21(1H-),//)
      TOOK=((((XN*XN)*SONE)-(XN*STWO)+(3.0*(XW*XW)))
      TALK=(((XN*XN)-1.0)*(XW*XW))
      EXSP=TOOK/TALK
      IF(EXSP.GT.0.0) GO TO 91
      WRITE(5,776)
   91 WRITE(5,92) EXSP
   92 FORMAT (1X,37HTHE EXPECTED VALUE ASSUMING IX TO BE ,/,1X,
     134HDRAWN FROM A NORMAL POPULATION IS ,F16.10,//)
      SAMM=(EXSP-((1.0/(XN-1.0))**2.0))
      IF(SAMM.GT.0.0) GO TO 93
      WRITE(5,779)
   93 RUUT=(SAMM**0.5)
      FINA=ABS((CORR+(1.0/(XN-1.0)))/RUUT)
      WRITE(5,94) FINA
   94 FORMAT (1X,22HTHE ABSOLUTE VALUE IS ,F16.10,/,1X,21(1H-),//)
      GO TO 13
   79 STOP
      END
```