THE RELATIONSHIP BETWEEN ROAD ACCIDENTS AND URBAN STRUCTURE

M.D. HODGKINSON BSc

(A Thesis submitted for the degree of Doctor of Philosophy)

Department of Civil Engineering
University of Salford

May, 1974

LIST OF FIGURES

LIST OF FIGURES Contd....

LIST OF TABLES

CONTENTS

Chapter I INTRODUCTION

The development of the motor-vehicle as a means of transport has had a distinctly dichotomous impact upon our present society. Whilst it has helped to remove the natural obstructions of distance and space for any individual, it has also created many problems, especially within the urban environment, for which there seems to be no immediate solution. If such solutions are to be forthcoming in the near future, it necessitates several immediate courses of action, all of which in their own turn demand a greater amount of research and consequently a greater understanding of the problems involved.

1.1 THE GROWTH IN TRAFFIC VOLUMES

The most obvious causal factor in regard to these urban social and economic problems is the growth in the number of motor-vehicles and the consequent increase in the volume of traffic, as measured by the annual vehicle mileage.

Since the first motor-vehicle was introduced into Britain in 1888, the number of vehicles has risen at a phenominal rate of growth. (Fig. 1.1.1). From a figure of approximately 400,000 vehicles in 1914, the number of vehicles increased to 3.0 million in 1946, and 14.5 million in 1970. (HMSO 1972). That represents an increase of almost 500% in under 25 years, (1946-1970).

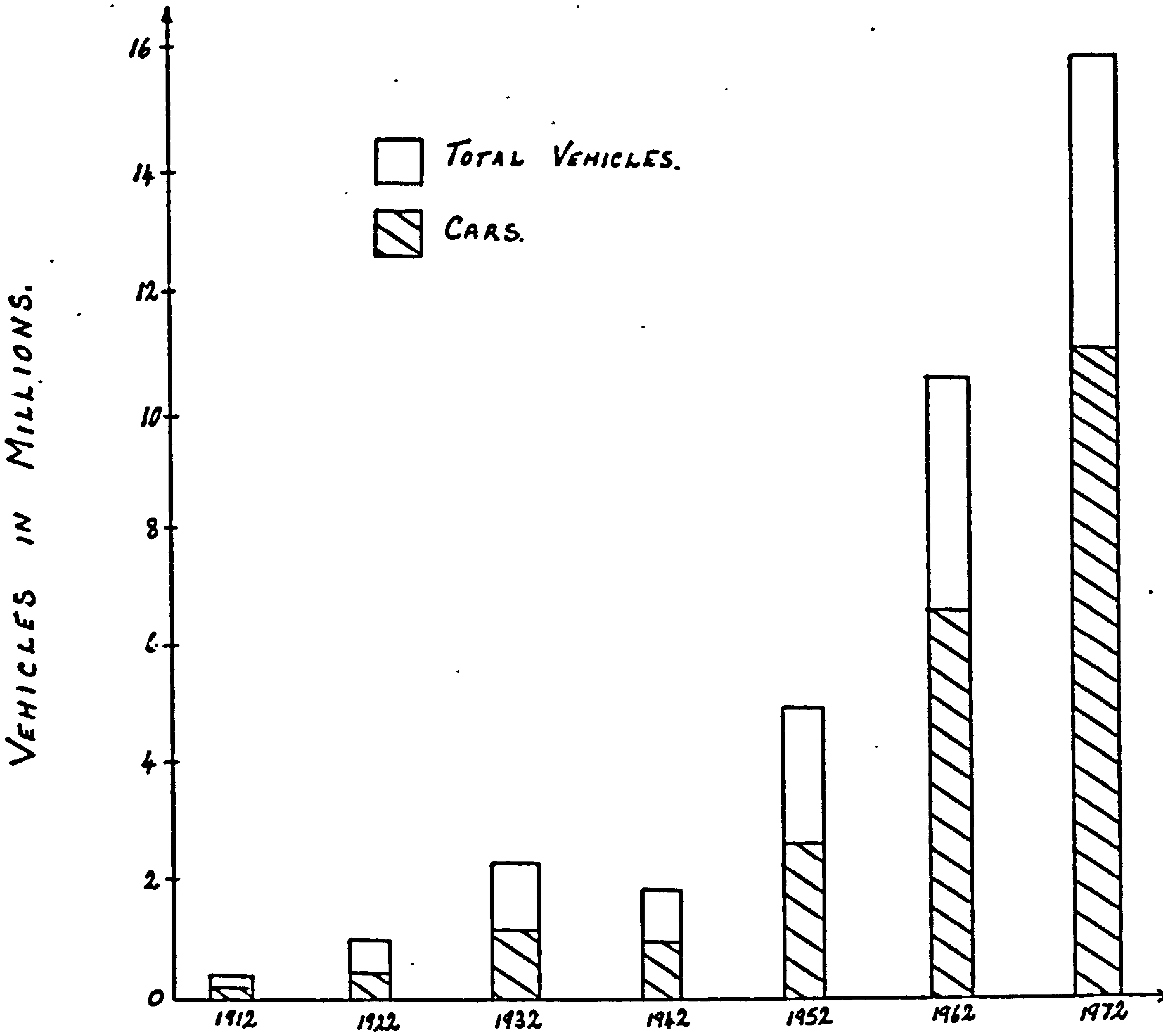


Fig. 1.1.1. THE GROWTH IN THE VOLUME OF VEHICLES IN G.B. 1912 - 1972.

A similar increase can also be noted in the growth of vehicle mileage. These figures are based on the regular traffic census returns of the Department of the Environment and the Transport and Road Research Laboratory (TRRL). If one takes the index of vehicle mileage for motor traffic in 1949 as being equal to 100, the comparable index for 1970 is 444. (Fig. 1.1.2).

1.2 EFFECT UPON URBAN AREAS

As already mentioned, the areas where these rapid growth rates have had the most striking effect have tended to be the larger centres of population and urban areas in general. The problem can be simply generalised by stating that most urban structures are the result of slow evolution since the time of the post-industrial revolution, to the present day, with the consequent result that most of these developed structures are unable to accommodate the increased traffic volumes, nor are they able to alter rapidly enough to create a satisfactory alternative environment. For this reason, most of the present urban environments can be viewed as inadequate to deal with the present urban transportation problem. This inadequacy can be seen in two ways. Firstly it is epitomised by the varying degrees of urban blight and social deprivation evident in all urban areas, especially where the conflict between the development of the transport network and the need for greater social awareness is most pressing. Secondly it is also epitomised by the high degree of failure within the urban transport system, as indicated by the high level of urban traffic conflicts or casualties. Conceptualised in this way, road

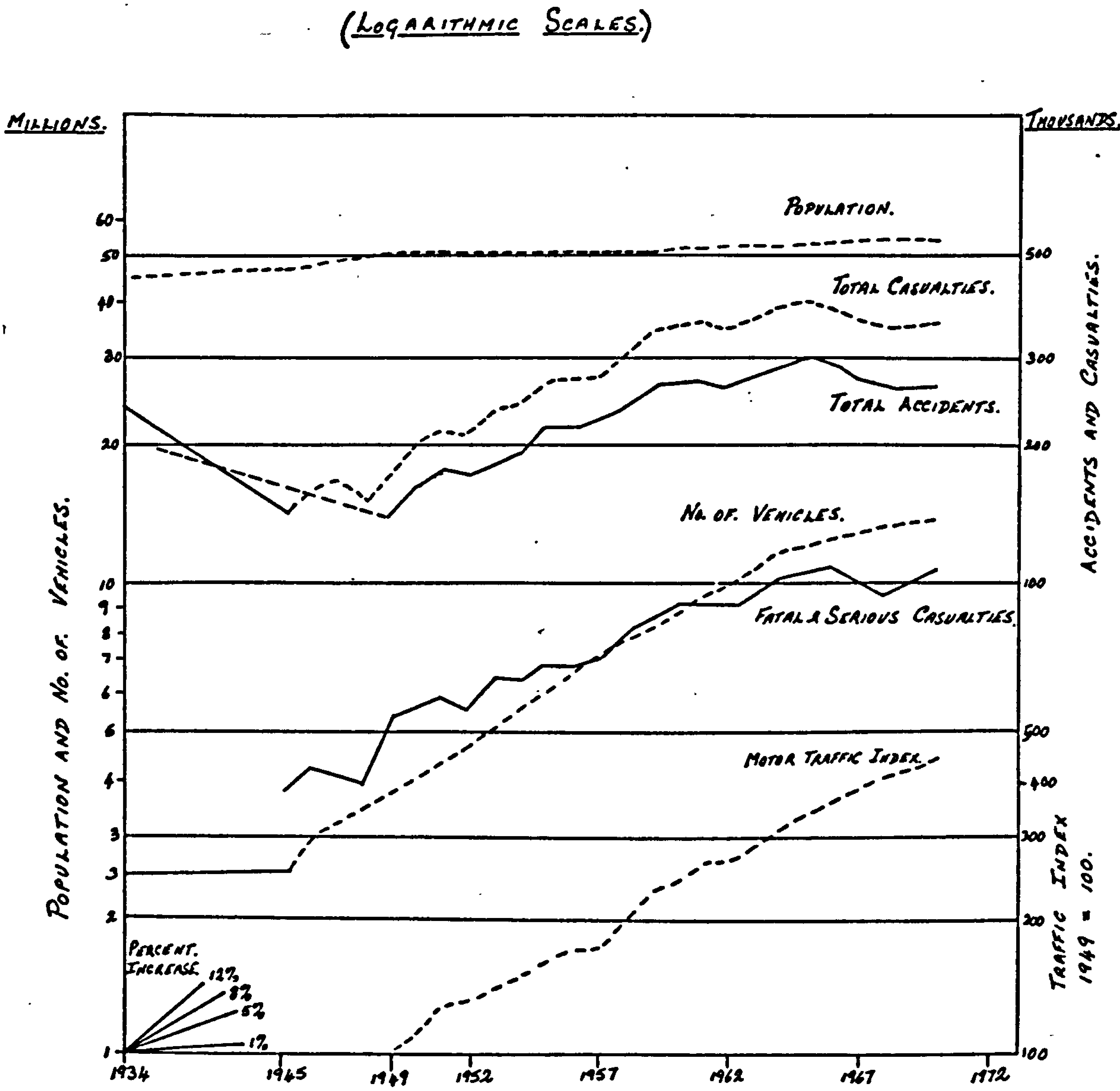


Fig. 1.1.2. POPULATION, ROAD MOTOR VEHICLES, TRAFFIC, ACCIDENTS AND CASUALTY RATES 1934 - 1970 (ROAD ACCIDENTS 1970)

accidents can be seen to be system failures, or the failure of the urban structure system to change sufficiently to accommodate the urban traffic system. Consequently, it can be argued that the level or degree of failure of any one individual area is a function of the degree of change within the urban structure system, either in physical, social or economic terms. By identifying these relationships it should then be possible to postulate the optimum lines of research, and internal modification of the various systems, for the future. However, before this point can be reached, it is necessary to document and classify the degree of failure of the various individual systems, and also the national system as a whole. To achieve this one has initially to look at the present road accident situation noting both the present trends and also any anomalies.

Chapter 2 THE PRESENT ROAD ACCIDENT SITUATION

2.1 INTRODUCTION

In order to gain some insight into the road accident situation in Britain at present, this chapter attempts to present the available road accident statistics (1970) both at the aggregrated level, and the disaggregated level. A further sub-division is effected by referring to urban road accident statistics, which indicate the seriousness of the problem in built-up urban areas.

2.2 AGGREGATED NATIONAL STATISTICS

If one considers the aggregated national statistics in the first instance, it can be seen that there has been a consistent overall increase in the number of road casualties, such that in 1960 there were approximately 348,000 people killed or injured on all the roads in Great Britain. (Table 2.2.1). By 1970 the comparable statistic had risen to approximately 363,000. Perhaps a more meaningful way of representing these figures is by comparing them to one or other basic variables. For example, if the casualties are related to the motor vehicle mileage index, it can be shown that whilst the number of casualties has doubled over the last twenty years, the index shows that over the same time period, vehicle mileage has trebled. However, if casualties are related to the increase in population, the picture is not so good. Road deaths have increased by 57% over a period when the total population has increased by only 10%, and therefore the probability of an individual losing his life on the roads has shortened from 150 to 1, to 100 to 1.

CASUALTIES FROM ROAD ACCIDENTS (10^3)					ACCIDENTS (10^3)
YEAR	KILLED	SERIOUS	SLIGHT	TOTAL	
1960	7.0	84	256	348	272
1961	6.9	85	258	350	270
1962	6.7	84	251	342	264
1963	6.9	88	261	356	272
1964	7.8	95	282	385	292
1965	8.0	98	292	399	299
1966	8.0	100	285	392	292
1967	7.3	94	269	370	277
1968	6.8	89	254	349	264
1969	7.4	91	255	353	262
1970	7.5	93	262	363	267

(ROAD ACCIDENTS 1970)

TABLE 2.2.1. THE GROWTH IN ROAD CASUALTIES 1960-1970 (G.B.)

2.3 DIS-AGGREGATED NATIONAL STATISTICS

When disaggregated, the national statistics also provide some pertinent information. During 1970 there were 85,370 pedestrian casualties, which was an increase of 18% since 1960. (Table 2.3.1). When related to total casualties, these pedestrian casualties account for 24% of all casualties in Great Britain. The most significant feature however, of these statistics can only be shown when the number of pedestrian casualties are broken down into a relevant age distribution. (Table 2.3.2). This distribution shows that two groups of the population, those aged 15 years or less, and those aged 65 years or more, account for approximately 62% of all pedestrian casualties, and the young pedestrians almost 50% of all pedestrian casualties on their own.

If these pedestrian statistics are put into a European context it can be seen that pedestrian casualties per 1000 population are the highest in the 5 - 14 years age group in all countries except for Sweden and Ireland. (Table 2.3.3). Studying this table even further, it can be seen that Great Britain has in fact the highest pedestrian rate in all of Western Europe. Although this can possibly be explained by the variation in such variables as level of urban development and population density, it still points to the necessity for a better understanding of the problem, and therefore hopefully, more solutions which may help reduce the problem.

The other major group of road users, drivers and passengers, has a similar problem to those described above for pedestrians. During 1970, 205,245 drivers and passengers were injured, about 56% of all casualties. However, as with pedestrians, almost a third of these casualties occurred within one age group, the eight year period

YEAR	PEDESTRIANS			
	KILLED	SERIOUS	SLIGHT	TOTAL
1960	2708	19831	49675	72214
1961	2717	19880	49702	72299
1962	2681	20062	49498	72241
1963	2740	20838	51551	75129
1964	2986	22478	55017	80481
1965	3105	23714	58248	85067
1966	3153	24786	57684	85623
1967	2964	24472	56843	84279
1968	2762	24440	56449	83651
1969	2955	24267	56486	83708
1970	2925	24875	57570	85370

TABLE 2.3.1. THE GROWTH OF PEDESTRIAN CASUALTIES 1960-1970 (G.B.)

AGE GROUP	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970
0 - 4	2058	2366	2236	2263	2557	2689	2782	2956	2936	2754	2642
5 - 9	4018	4155	4287	4338	4746	4892	5334	5533	5850	5614	5861
10 - 14	1712	1801	1844	2002	2087	2337	2487	2608	2725	2881	3106
15 - 19	972	1001	1185	1336	1680	1730	1892	1822	1755	1881	2034
20 - 29	1128	1229	1219	1446	1403	1613	1739	1727	1731	1853	2016
30 - 39	1035	1084	1028	1138	1183	1171	1258	1261	1125	1214	1218
40 - 49	1532	1417	1500	1536	1562	1662	1696	1560	1567	1639	1576
50 - 59	2475	2222	2343	2347	2458	2535	2685	2402	2307	2202	2219
60+	7386	7105	6873	6929	7577	7929	7881	7460	7089	7147	7123
UNKNOWN	223	217	228	243	211	261	185	107	117	37	5
TOTAL	22539	22597	22743	23578	25464	26819	27939	27436	27202	27222	27800

TABLE 2.3.2 THE NUMBER OF PEDESTRIANS KILLED AND SERIOUSLY INJURED BY AGE GROUPS 1960–1970 (G.B.)

between 17 and 25 years, indicating once again the vulnerability of young drivers, although the degree of exposure of this group must obviously be taken into account before any such definitive statement can be categorically accepted.

2.4 THE URBAN ROAD ACCIDENT PROBLEM

So far this account has only described the situation in all of Great Britain. However, since this research attempts to look at the distribution of road accidents in urban areas, it is also necessary to look at the situation related solely to these areas, as defined by roads having a speed limit of 40 mph or less.

The major problem in urban areas is, and has been, the conflict between pedestrians and road traffic. In fact what Swift wrote in the "Tatler" in 1710 could almost be applicable to present day conditions, some two hundred and fifty years later:-

"We are very glad to watch an opportunity to whisk across a passage, very thankful that we are not run over for interrupting the machine that carries in it a person neither more handsome, wise nor valiant than the meanest of us."

Similarly, Stone (1914) who was describing the situation at the turn of the century, also put most of the blame on the motor-vehicle:-

"It is practically impossible for the pedestrian to gauge the speed of a motor car approaching end on .... And with its elasticity,

NATIONS	Pedestrian Casualities per 1,000 population Age 0 - 4	5 - 14	15 - 24	All Ages	Pedestrian Fatalities (all ages)1 Population
G.B.	2.0	3.6	1.7(15-19)	1.55	55
Austria	1.5	2.2	1.2	1.44	100
Belgium	1.1	1.9	0.9	1.06	70
Denmark	1.2(0.6)	1.5(7-14)	0.6(15-20)	0.8	60[1]
France[2]	0.6	1.2	0.7	-	70[1]
W. Germany	1.7(0-5)	2.8(6-14)	1.2(15-20)	1.25	90
Ireland[3]	0.5	0.8	0.8	-	60
Italy	-	-	-	-.75	50
Luxemboug	1.1	2.7	0.7	0.8	-
Netherlands	-	-	-	0.65	50
Norway	-	-	-	0.65	50
Portugal[4]	0.9(0-5)	1.1(6-14)	0.8(15-20)	1.00	100
Sweden	0.5(0-6)	0.4(7-14)	0.4(15-19)	-	30
Switzerland	1.2	2.3	0.5	1.06	70
Yugoslavia	0.3	0.7	0.5(15-18)	0.50	60

\- NOT AVAILABLE (1) 1968 FIGURES
(2) 1966 "
(3) 1967 "
(4) Fatal and Serious

TABLE 2.3.3. INTERNATIONAL COMPARISON OF PEDESTRIAN CASUALITIES BY AGE GROUPS (1969)

even a judgement of pace would be of little use, since cars can vary their speed between say, 4 and 40 mph in the length of a few yards. This alone, therefore, throws the entire responsibility for accidents on the car and not the pedestrian." (Stone 1914).

Although this is perhaps a harsh judgement against the motor-vehicle, the actual pedestrian casualty figures for urban areas do emphasise the vulnerability of the pedestrian, against the more solidly constructed automobile. Thus in 1970, 262,553 road casualties occurred in urban areas accounting for 72.7% of the total number of casualties in Great Britain. Similarly, pedestrian casualties in urban areas accounted for 30.8% of all urban casualties, giving a figure of 80,751 pedestrian casualties in urban areas, which in turn is almost 95% of all pedestrian casualties in Great Britain.

Perhaps one of the major difficulties related to this pedestrian-traffic conflict has been the dichotomous approach usually used, and spotlighted in the Buchanan Report. (Buchanan et al 1963). On the one hand there are the needs of the driver sub-system, who demand increased network access, and higher speeds etc., whilst on the other hand, there is the pedestrian (and dwellers) sub-system who demand greater safety and freedom from traffic. Now although both of these two groups are not mutually exclusive, in most instances the reaction of an individual reflects his actual movement within either of the groups at any one point in time. Consequently, there has to be some balancing of suggestions and proposals to accommodate both sub-groups, and the advantage has usually gone to the group which could exert the greatest pressure. Until recently this has tended to be the motor-vehicle "lobby" with the result that recent developments have tended to aid the movement of vehicles around our urban areas, through increased capacities, or new constructions, all at the expense of our environment in terms of noise, safety and pollution. It is for this reason that little, if any, action had been taken as regards road safety proposals, until the environmentalist "lobby" had also started applying pressure.

This neglect of the pedestrian-traffic conflict can perhaps be best illustrated by the fact that most of the solutions traffic engineers are now suggesting for pedestrian safety in modern towns, such as shopping precincts, elevated pedestrian walkways, and subways etc., were all suggested many years ago, and in some cases centuries ago. Thus in 1500, Leonardo da Vinci proposed a scheme for ground level pedestrian-ways with the vehicular traffic running in tunnels. Similarly the Pantiles in Tunbridge Wells built in the 17th century are an example of an early pedestrian precinct.

Basically most of the approaches aimed at satisfying the pedestrian-traffic conflict all contained some degree of pedestrian-traffic segregation. Thus as Stone says, "If absolute safety for pedestrians is required, it is attainable only by the overbridge or subway, so that all interference is avoided...... That the overbridge has not become a feature of our own English cities is curious, when we consider the stringent regulations imposed on railway companies with regard to level crossings." (Stone 1914). A similar point was made by H.A. Tripp, an Assistant Commissioner of Police at Scotland

Yard, some twenty five years later:- "If we could segregate pedestrians completely from the wheeled traffic, we should of course abolish pedestrian casualties. Complete segregation cannot, however, be achieved.......so one has to concentrate on the major roads, and there the problem is definitely soluble." Tripp then continues by pointing out that, "where heavy through traffic invades old age towns and villages, one shall have to build roads exclusively for motor traffic and by-passing the centre of population. In places where pedestrians and heavy streams of traffic must still use the same thoroughfare, resorts must be had to either place-segregation (e.g. overbridges) or time-segregation (i.e. traffic signals)." (Tripp 1942).

Thus Tripp appears to understand the basic problem of road safety when he says that complete segregation is totally impossible. Although this would seem to be the Utopian ideal, only in very exceptional circumstances would it be at all possible and this would have to be within the framework of some new development. But as Tripp realises, the major problem concerns the older, more established towns, where modification is only something which can be completed in a piece-meal manner, thereby precluding the idea of total segregation, and restricting any segregation to very limited "isolated" areas.

2.5 THE NEW TOWN DEVELOPMENTS

Whilst the older towns provided obviously the greatest problems, it was hoped that the development of the New Town Ideal would create new environments which would be in themselves safer, whilst at the same time alleviating some of the pressures in the other older areas. Although some work had been started on New Towns, especially those Garden cities advocated by Ebenezer Howard, at Letchworth and Welwyn, little progress was really made until the post-1945 period. In fact Sir Patrick Abercrombie, one of the leading figures in the New Town movement after 1945, wrote in 1933 that - "It is interesting to speculate what might have been done if three or four million people out of the sixteen million who have been provided with detailed local residential planning in the country, had their housing made the subject of a really comprehensive plan. A hundred real new towns in different parts of the country would have been more constructive, if also more adventurous than a multitude of garden suburbs. In fact this post-war (post World War 1) housing as regards a general theory of civic planning, does not show much advance upon the Victorian muddle." (Abercrombie 1933).

The post 1945 developments, both first and second generation, New Towns seem to meet most of these criticisms, yet it is interesting to note that whilst the majority of these towns were built and planned explicitly with road safety in mind, the amount of research into the results of these experiments is very limited. Surely if money is to be spent it ought to be possible to determine whether this money is being used to a good purpose. With this thought in mind it is useful to compare the response of these entirely different urban systems within a road casualty context, with other less-planned urban areas.

2.6 THE APPLICATION OF COSTING TECHNIQUES

The present road accident problem can be put into context in many ways, as already indicated. One method which seems to be becoming more applied of late is that of cost-benefit analysis, and other related techniques. Wohl in a recent article concerning the analysis and evaluation of traffic safety measures, has formulated his argument somewhat in these terms. (Wohl 1969). He starts from two basic propositions -

"1) Safety measures that provide increased benefits (or reduced costs) to someone which, when totaled, outweigh the sum of increased costs to someone, should be adopted........

2) Safety measures that provide increased aggregate net benefits, but cause segments of the community to bear increased costs in excess of their increased benefits, should be "voted on" by society through its political process, and perhaps foregone."

He then continues:- "The traffic safety problem is to determine a) the aggregate levels of benefit and cost associated with the adoption of one safety measure or another, b) who will benefit and how much, and c) who will pay and how much. It is simply not good enough to say that "society as a whole will benefit" or that "safety will improve" if we adopt a certain safety measure."

Whilst one does not have to agree with the suggestions and arguments in Wohl's paper, it does indicate the present trend towards costing of various alternative proposals, and therefore the need for a vast increase in our knowledge about the results and implications of any improvement procedure. Without this knowledge such costing procedures would appear ambiguous at least, and meaningless at worst. Therefore it would appear to be a reasonable argument that until sufficient research has been completed upon road safety, the application of cost-benefit techniques would seem presumptuous and totally restrictive, especially in the short term. However, having made this observation it must be pointed out that the costing of road accidents to the community as a whole can help to indicate the seriousness and the degree of urgency needed to reduce the problem.

Every road accident is ultimately impossible to cost in terms of pain, grief and suffering to the individuals concerned, but it is possible to represent each accident as a quantifiable loss to the community in economic terms. Such costings have been studied both at the Ministry of the Environment and the Transport and Road Research Laboratory. The work at the TRRL was carried out by Dawson (1971), who estimated that the total number of accidents in Great Britain, cost the community approximately £318 million in 1970. These costs were mainly derived from costs leading to a diversion of current resources, and the gross loss of future output.

The work by the Ministry of the Environment, meanwhile, as published in the HMSO publication "Road Accidents 1970," arrived at the greater figure of £346 million (1970). (Table 2.6.1). They also calculated that the average economic cost of a fatal accident was £13,000, the average serious injury accident cost £700, and the average slight injury cost £230, and the average accident involving

	1970 £m	1969 £m
Medical treatment, ambulance and funeral costs	17	15
Police and Administration costs	28	26
Damage to vehicle and other property	198	183
Lost output	103	96
TOTAL:-	346	320

TABLE 2.6.1 COST OF ROAD ACCIDENTS (1969 AND 1970)

only damaged vehicles and property £100. They then continue by stating that, "on average road accidents result in an economic loss of approaching £1 million a day on top of the personal suffering caused by the daily toll of more then 20 killed, 250 seriously injured, and 700 slightly injured. These figures provide the facts behind a problem which must be the concern not only of the Government and safety organisations, but also of all road users"(HMSO 1972).

Chapter 3. ROAD ACCIDENT ANALYSIS - PAST AND PRESENT

3.1 INTRODUCTION

Since road accident analysis is a relatively young subject, it can be compared with many of the other embryonic social-science subjects such as sociology, economics, geography etc., all of which have their own related technical problems, and general overall, methodological limitations. However, when compared to these other disciplines, road accident analysis would seem to have been more restricted in its type and rate of development. This section tries to account for this apparent slowness in development whilst also suggesting approaches which may be more beneficial in the future.

3.2 PROBLEMS WITHIN ROAD ACCIDENT ANALYSIS

The reasons for the dearth of development within road accident analysis can perhaps best be understood by referring back to the subject matter itself. Whilst the other social sciences have tended to develop through the use of increased "scientific" rigour, and the greater use of statistical techniques necessary for hypothesis testing, road accident analysis has been restricted by three simple failings.

In the first instance, road accident analysis has been beset by the need for practical solutions to immediate problems, and therefore the development of any theoretical framework has been severly restricted. That is to say, road accident analysis has tended to develop as a method of program evaluation, where very often, evaluation is merely a formality after the program has been implemented. There has been little, if any, effort to understand the problems involved in any accident situation, and the most emphasis has been laid on reaching an empirically satisfactory solution. Yet such an understanding and hypothesis testing is essential in any approach to scientific explanation. Thus as Likert (1960) has pointed out, there are two functions of statistics in the process of evaluation:-

a) To provide information on the state of the system - "statistical measurements which reveal the current situation of the nation or economy, such as population data, price indices, etc."

b) To provide information on the nature of the system - "the basic conceptual model which serves as a guide to tell what dimensions of the nation should be interpreted in making decisions."

Whilst both of these points are important in pointing out the need for a more scientific approach in road accident analysis, the second point is perhaps the more pertinent one here since it explicitly indicates the need that there should be statistics available which can provide information on, or help build a "basic conceptual model." The implication of such a statement is simply that, without the basis of some theoretical understanding, the evaluation of any procedure, or program, is of very limited value. For this reason it is obvious that to progress, road accident analysis will have to move further in the direction of scientific explanation and theoretical conceptualisation. Before leaving this point, it would perhaps be worthwhile to point out that there are two basic routes used in scientific explanation, each of which has some advantages, depending upon the subject and the subject material being used. Therefore, since it is being suggested that road accident analysis should develop its level of scientific explanation, these two routes ought to be introduced and the differences noted. The two routes are generally known as inductive and deductive explanation.

(i) INDUCTIVE EXPLANATION - Inductive explanation can be classified as the classic model of scientific explanation as propounded by Francis Bacon. By this method, the scientist moves from a collection of perceptual experiences which when transformed into some language, forms a mass of poorly ordered statements into some kind of grouping or classifactory system. As Harvey points out:-

"In the early stages of scientific development, such ordering and classification of data may be the main activity of science, and the classifications so developed may have a weak explanatory function." (Harvey 1969).

Further study of these groupings, and the interactions between these groups may lead to the discovery of certain regularities, which through inductive inference may suggest the formulation of certain empirical (inductive) laws and ultimately postulated universal laws from the structuring of the empirical laws, into some theoretical framework. This mode of scientific explanation is shown diagramatically in (Fig. 3.2.1).

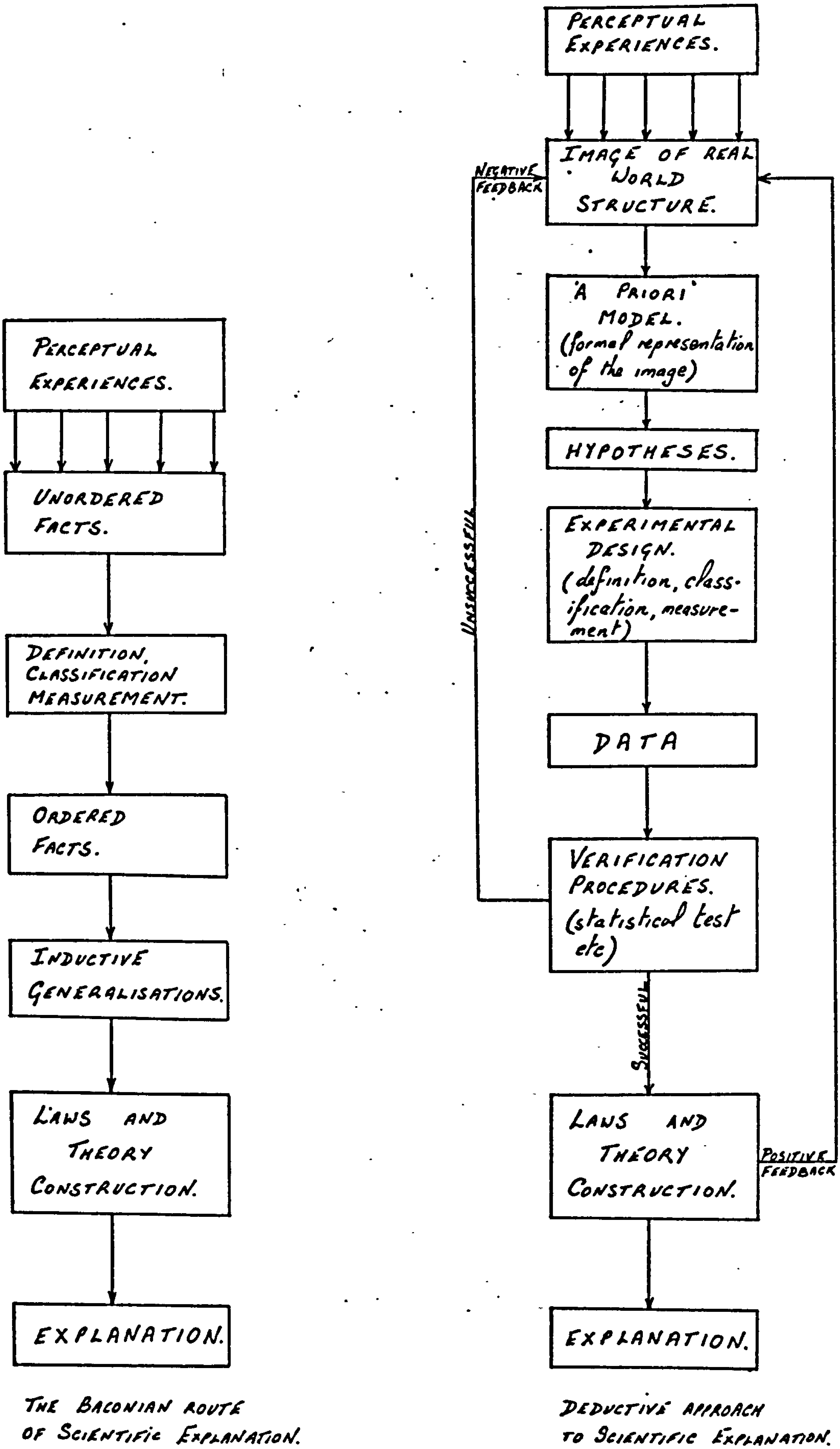


Fig. 3.2.1. TWO ROUTES TO SCIENTIFIC EXPLANATION

The problem with the inductive method of explanation is "the assumption that the processes of ordering and structuring data are somehow independent of the theory ultimately constructed." (Harvey op. cit). Thus as Churchman has argued, "facts, measurements and theories are methodologically the same." (Churchman 1961). Therefore, if one creates an "a priori" classification system to a set of data, it may be regarded identical to postulating an "a priori" theory.

(ii) DEDUCTIVE EXPLANATION - Deductive explanation meanwhile explicitly includes the "a priori" nature apparent in most scientific research, through the creation or formulation of some "a priori" model. The perceptual experiences are input into this theoretical model generating hypotheses, which can be then tested, according to some degree of confidence, by various statistical techniques. As (Fig. 3.2.1) shows, the process is cyclical in nature, only allowing the construction of laws and theories once the "a priori" model has been verified to a satisfactory degree. It will also be noted that classification only occurs once the "a priori" model and resultant hypotheses have been postulated. In this way the classification system is based on a hypothetical framework which can be tested and, therefore, ultimately accepted or rejected before further theoretical work is attempted. In this form the deductive approach to scientific explanation would appear to be the most satisfactory, especially in the field of social sciences where very complex systems, with a high degree of inter-relationships are encountered, and where the original data cannot be controlled in some experimental situation.

Whilst such an approach has been adopted freely in most of the other social science disciplines, in road accident analysis, this has not been the case, even where there has been an attempt at any coherent explanatory framework.

The problems indicated above pin-point the second major failing of road accident analysis. Like most social science subjects, advanced research has been restricted by the complex inter-relationships which must be studied and understood within any meaningful model. However, within road accident analysis, most models which have been developed have tended to be more simplistic in nature than realistic, basically because these models can be handled more easily by the researcher, and also because they circumvent the numerous data problems. Although some of these models appear to work reasonably well, it must be remembered that they are merely abstracted models which have little,if any, meaning when applied to a real world situation. That is not to say such models or relationships have no relevance whatsoever, but the extent of their limitations must be realised. One aspect of this debate in the literature has concerned the formula suggested by Smeed to predict the number of deaths from road accidents in various different nations (Smeed 1949). Smeed's work was "illustrative of a type of statistical modelling in considerable use today in various social sciences, where a limited number of variables are linked in a functional relationship without there being any available expla nation of causality. Thus:-

$$\frac{D}{N} = 0.003\left(\frac{N}{P}\right)^{-\frac{2}{3}}$$

where: D = deaths; N = number of vehicles; P = population,

remains as predictively valid (i.e. current data fits the curve), and as uninterpreted today as it did 20 years ago." (Katz 1971). This lack of interpretation was taken up by Peranio (1970) in a general discussion on the use of formulae in traffic engineering and road safety, and the Smeed formula in particular. He argued that all useful formulae must have certain attributes, and that without these attributes, and derived complex models, the possibility of useful formulae in road safety was very low. Thus, he states that in road safety topics, "Humans of varying characteristics play key roles. Variables include all aspects of the human highway users, vehicles, highways, and the physical and sociological environment. Whereas engineers work routinely.....with errors of about 1 per cent, those dealing with humans operating in open systems are fortunate if they can roughly predict what will happen....... In general, it is safe to say that when an obviously simple formula purports to describe a complex phenomenon in an open system, one can without trepidation question its validity or practical usability." (Peranio 1970).

Smeed replied to these criticisms by pointing out that what Peranio said was undeniable, but that it did not mean that even where one had to question the validity and practical usability of a formula, one necessarily had to reject that formula. (Smeed 1972). Similarly, Smeed also pointed out that any formula with any validity at all, "is almost bound to throw light on the phenomena it describes, even if it does not fully explain them." (Smeed op. cit.). Thus in the case of Smeed's formula, he argues that several possible reasons can be put forward to explain the tendencies revealed. These include the increased probability of being involved in a road accident with increases in the volume of motor traffic, and also the influence of some major counterbalancing factors; "as the population accident rate becomes higher the urge to do something about it becomes greater, and that something is in fact done. In addition, as the number of motor vehicles increases, which is in practice as time goes on, people are growing up and becoming more used to dealing with the situations which motor-traffic causes." (Smeed 1949). In a way, Katz puts this argument more succinctly when he states that, "Despite this critical lack of interpretation one important historical property of the system - the increase in safety attendant to system expansion - has been demonstrated." (Katz 1971).

It would seem that the crux of this argument revolves around the ultimate aim, and degree of progress desired by each of the writers. Peranio would appear to want to create some kind of complex model which enables both explanation and interpretation, even though he recognises the multitude of problems to be overcome before this becomes possible. Smeed on the other hand seems to be arguing that any model, even if it adds very little to our understanding of the situation, is better than nothing, especially in the short term, whilst the subject is still in the early stages of development.

Whilst both arguments carry some conviction, it would seem reasonable to argue that if road accident analysis is to develop and create some methodological framework, there would seem to be more hope for the future in the ambiguous proposals of Peranio. However, it must be emphasised that both proposals are complementary, rather than as usually portrayed, supplementary. For this reason, both approaches should be encouraged and the relevant progress noted.

The third major problem which road accident analysis has had to overcome, has been that until recently, it has had no academic core or base. Although this is less a philosophical problem and more a practical and organisational problem, it has caused many related problems within the subject matter. Until recently, when road accident analysis has tended to become a more coherent subject, the normal situation was to find road safety research being done by isolated individuals within entirely different, and therefore independent,academic disciplines. The result has been that interdisciplinary comparisons have been severely complicated, the collection of all relevant material restricted, and the construction of a satisfactory all-embracing methodological subject framework, prevented.

However, even though development within the subject has been severely restricted for the above reasons, some progress has been made and recent developments suggest that the major advances will be made along three basic lines of research.

3.3 POSSIBLE FUTURE DEVELOPMENTS IN ROAD ACCIDENT ANALYSIS

(i) MODEL BUILDING:- It was pointed out earlier in this section that models developed within road accident research had tended to be simplistic in nature and therefore restricted in use. However, one of the major trends of recent years has been the creation or formulation of a much more meaningful group of models, with a much higher degree of mathematisation. Similarly, associated with this increased use of mathematical models, there has been an increase in the application of statistical techniques and, therefore, objective hypothesis testing.

Since models have many different functions, dependent upon the aim of any piece of research, it is virtually impossible to define a model except in terms of its relevant function. However, Ackoff (1962) has tried to circumvent this problem by suggesting that a model can be conceptualised as being an "idealised representation of reality in order to demonstrate certain of its properties," and then continues by advocating a breakdown of models into three types:- iconic, analogue, and symbolic, in which each stage represents a higher degree of abstraction than the last. However, for the purpose of this discussion, it would appear more meaningful to use Lowry's classification of models based on function. (Lowry 1965).

Before looking at Lowry's classification in more detail, it is relevant at this point to ask the question, "Why do we bother creating models rather than study the facts themselves directly?" The answer lies in the inevitability, the economy and the stimulation of model building.

a) Model building is inevitable because there is no fixed dividing line between facts and beliefs; in Skilling's terms, "belief in a universe of real things is merely a belief...a belief with high probability certainly, but a belief none the less." (1964). Models are theories, laws, equations, or hunches which state our belief about the universe we think we see.

b) Model building is economical because it allows us to pass on generalised information in a highly compressed form. Like all rules there may be exceptions, but this does not in itself reduce the importance of the model as a first approximation in the learning process.

c) Model building is stimulating in that through its very over-generalisations, it makes clear those areas where improvement is necessary. Thus the building and testing of any model is of paramount importance in any subject, since each step leads onto further research and modifications. In other words, as has already been emphasised, model building and theoretical conceptualisation of any problem is a basic stage in any method of scientific explanation.

Returning, therefore, to Lowry's classification, one can see how the three classes; descriptive, predictive, and planning, all perform a different, but essential function in understanding any research problem.

(1) DESCRIPTIVE MODELS

According to Lowry, a descriptive model is one which replicates the "relevant features of an existing urban environment or of an already observed process of urban change."[1] (Lowry op.cit). The importance of such models are numerous, and for this reason should be the first models derived within any field of research. In the first instance, such models reveal much about the structure of any system and reduces the complexity of the real world into the more coherent and rigorous language of mathematical relationships. At the same time, descriptive models provide concrete evidence on the inter-relationships of the systems processes. Thirdly these models can also generate reliable values for "hard-to-measure" variables from input data consisting of "easy-to-measure" variables. Unfortunately, however, it is impossible for descriptive models to predict future relationships or situations, or to evaluate between different programs. These are the tasks of the other two classes of models proposed by Lowry.

(2) PREDICTIVE MODELS

When building a predictive model it becomes necessary to have an understanding of the relationship between form and process. Whilst it was sufficient in a descriptive model to note that two variables X and Y are co-variant, in a predictive model it becomes necessary to be able to specify a causal sequence. "If one is able to postulate the direction of causation knowledge of the future value of the "cause" enables one to predict the future value of the "effect."

---

1 Lowry was speaking in terms of models for urban planning and administration.

(Lowry op. cit.). Therefore, when building a predictive model, the researcher has to satisfy two conditions. In the first place, he has to establish a logical framework where the variables of interest are at the end of a causal sequence rather than the beginning. Secondly, he must be certain that any "causal" variables can be themselves predicted as far into the future as necessary. This second condition can be dropped if one is interested in a "conditional" prediction. That is where one is interested in the consequence that should be expected to follow a specified exogonous impact (change in variable X), if the environment were otherwise undisturbed. In road accident analysis, such an interest could be identified with the imposition of a speed limit on urban roads, and its consequent result upon the level of road accidents etc.

(3) PLANNING MODELS

A planning model is one which incorporates the method of conditional prediction and relates the outcome to the researcher's aims or goals. That is, the planning model is an evaluatory model for deciding between different alternatives proposed by the researcher during any review of the environment or system.

These three classes of models as described by Lowry, therefore, necessitate the model builder to progress through three complementary stages of analysis. He must begin with the identification of persistent relationships among relevant variables, move onto the study of causal sequences and ultimately develop a logical framework for the model. In so doing he should develop or borrow from relevant theories postulated at an earlier date. It is at this point that model building in road accident research is severely restricted and the lack of any theoretical understanding of the road accident system underlined. For although the terms "model" and "theory" are often used interchangeably to denote a logico-mathematical construct of interrelated variables, a clear distinction should be drawn. Unlike the reverse situation, the model builder is dependent upon the theorist, since he is explicitly applying theoretical constructs to empirical situations, and where the theorist can be satisfied with vague logically coherent statements, the model builder has to deal with exact functional relationships in order to generate empirically-relevant output from empirically-based input. For this reason it can be argued that the growth of relevant model building in road accident research will only increase following the expansion of our theoretical understanding of the road accident problem as a whole.

Therefore, in the context of road accident analysis, most of the models so far developed fall into the descriptive category, with perhaps the exception of some modified regression models which have tried to postulate some kind of functional relationship. This family of descriptive models include both the Smeed model already described, the various probability models, and also the various behavioural models such as the accident proneness model. Ultimately, of course, one would hope to be able to produce planning or evaluatory models, but at present it must be emphasised once again that a basic need is the further development and understanding of the various significant functional relationships.

It was pointed out at the beginning of this section, that associated with the increase in model building, there had also been an equivalent increase in the use of statistical techniques, especially for simple hypothesis testing and program evaluation. Unfortunately, some of the results of these tests have not been exactly what they appeared. In this context it is useful to discuss one technique which has been applied quite extensively in road accident analysis, since it has several implications which are relevant both to road accident research in general, and this research problem in particular. The technique referred to involves the application of Chi square analysis in the evaluation of road accident improvements, and is generally known as a "Before and After" study.

Before and After studies are usually associated with the removal of road accident "black spots" or high accident location treatment. The technique was described by Tanner (1958):-

"To estimate the effect of a change in road conditions at a particular site on the frequency of accidents there, the usual procedure is to obtain details of accidents at the site in convenient periods before and after, and to compare the ratio after to before with the corresponding ratio for a large control area. The latter may be the whole of the Police district in which the site lies, or some other area from which trends due to external factors can be reliably assessed. The significance of the difference between the two ratios can be tested in the usual way by means of X^2 with one degree of freedom."

Although this technique has been used extensively, it is necessary to point out two severe theoretical limitations. The first relates to the control area and the problem of allowing for unknown variables which may influence the road accident level, whilst the second relates specifically to the purpose of the technique, and the areal unit used. This second criticism needs to be expanded since it is relevant to the main body of this research.

An important study in relation to the above mentioned criticism, was carried out by Thomson (1968) when he looked into the value of traffic management in central London. The basic aim of traffic management is to make "the best use of the existing street system." Its techniques range from the micro to the macro level, involving such actions as roadway markings, one-way streets, new roundabouts and flyovers etc. "A program of traffic management in a large city consists of hundreds, perhaps thousands, of small individual actions over the course of a year, which it is believed will add up to progress in better and safer vehicle flow." (Katz 1971). Unfortunately, the results of Thomson's work led him to two conclusions:-

(1) "The true (road-speedflow) capacity of Central London has

been falling as a result of traffic management, as supported by various statistics of travel volumes, flows, and speeds," and,

(2) "Comparisons of accident rates do not reveal any improvements which might be attributed to traffic management." (Thomson op. cit).

Whilst these conclusions have been debated at some length in the relevant literature, (e.g. Ridley et al 1968; Stott 1968) the most important point to be raised in Thomson's analysis was that whilst measurement of the local effects of individual changes on capacity was positive, measurement of network effects were either non-existent or negative. The implication here is that a bad traffic situation is cured at one location simply by moving the problem to one or more other points on the road network system. As Katz points out, "If these conclusions are correct, and are transferred to safety work, then they have obvious and important consequences for high accident location treatment, indicating the need to deal in wider areas of system improvement. For measurement work the implications are that conclusions about progress must not be drawn from too small an area or from data on individual treatments alone." (Katz 1971).

Although such observations and comments are not new, they have usually been dealt with implicitly rather than explicitly as should have been the case. These facts, along with the reluctance of most researchers to assume that individual improvements within a system (urban area) will automatically lead to an improvement in the overall system performance, makes it imperative that some movement is made to the formulation of some model or system of models, which can take into account the complex inter-relationships evident in road accident analysis.

(ii) MICRO-LEVEL STUDIES

The second major line of development in road accident analysis has been the increase in micro-level studies. The pieces of research which fall into this group attempt to study some set of isolated or local incidents (accidents) at the point of occurrence and try and derive some trend from these individual findings. Although such studies involve a partial inductive mode of reasoning, this approach can be queried by asking whether it is statistically meaningful to generalise from specific incidents, which are neither randomly selected nor the follow up work to anomalies discovered in a more general overall study.

An example of this type of study is the "Causes and Effects of Road Accidents" carried out by the Department of Transportation and Environmental Planning at Birmingham University. (1969). Once again, although this study provided considerable useful background material essential for any road accident analysis, it would be hard pressed to justify any universal generalisations which would be applicable to any area in Great Britain, without considerable extensive work being carried out, along the same lines, in other areas. For this reason it would seem that the micro-level approach is subservient in usefulness to the third line of development, which can be described as the application of systems analysis.

(iii) APPLICATION OF SYSTEMS ANALYSIS

The third method of development and the one advocated in this research is the conceptualisation of the problem as a complex interlocking system, allowing the application and use of the techniques of systems analysis. Without moving too far into the realms of systems analysis, it is necessary to deal with this point in more detail in order to comprehend the advantages of the approach and the considerable amount of work still needed before the techniques can be applied with any advantage.

Systems analysis is not foolproof nor is it a complete research methodology. It is merely a helpful set of research ideas, procedures, and techniques which have been developed by investigators to assist in the formulation of policy. Considered in this way, systems analysis can be associated with the "planning models" advocated by Lowry.

The simplest way of defining a system is through the use of set theory. Following Klir and Valach (1967), the set of objects (identified by a set of attributes of objects) contained within some system S may thus be represented as a set of elements:-

$$A = \{a_1, a_2, a_3 \ldots\ldots a_n\}.$$

To this one may add an extra element a_o to represent the environment. One then introduces a set:-

$$B = \{a_o, a_1, a_2 \ldots\ldots a_n\}$$

which includes all the elements within the system plus an element that represents the environment. The interactions and relationships between these elements can then be examined. If r_{ij} represents the relationship between any element a_i and a_j (the case where $r_{ij} = 0$ would be the case where a_i has no effect upon a_j), then we can denote the set of all r_{ij} $(i,j = 0......n)$ by R. The definition of a system is then contained in the statement that every set

$$S = (A,R)$$

is a system.

Having defined a system, the first step in the process is to develop as clear a statement of the problem to be tackled as possible. Lack of a clear problem statement is a major source of imperfection in decision making. In particular, there is a danger of inadvertently adopting one solution to a problem as the problem itself. For example, one could do a careful imaginative study of how to accommodate an additional desk in an office so that an extra secretary could be hired, when actually the problem is how to increase the typing output of the office, which may be more easily, cheaply and permanently solved by electric typewriters and duplicating machinery.

In the context of road accident analysis, this statement of problem could take many forms, and it is essential in any piece of work to describe exactly the ultimate aim of that study. For example, one statement of problem could be, given our present level and type of urban development, how can we reduce the number of road accidents to a minimum level? Alternatively, it would be possible to postulate an opposite statement of problem such as, what form of urban development in the future would give us the optimum level of road accidents? Of course in some cases it may even be desirable to have a reduction in the type of accident as the statement of problem. For instance, one could propound the statement, "how can the cost of road accidents to the community be reduced by decreasing the severity of road accidents?"

It is obvious that all these statements of problem are realistic and, therefore, pertinent in different situations, and the need to define the relevant statement is, therefore, fairly obvious and as indicated, must be the initial operation in any systems analysis.

Given a clear statement of objectives, the next step is to develop an appropriate model of the phenomena which are being studied. During the modelling process, a deliberate attempt is made to broaden the study so that important variables and relationships are not omitted. At this time, the thinking of the researcher should not be constrained by problems of data collection, measurement difficulties, computational difficulties, small budget, or any other real world constraints. He must try to force himself to think about an ideal, totally unconstrained solution to the problem in which all important phenomena and relationships are considered. In addition, an attempt is made to structure the problem by specifying the nature and direction of the relationships between the phenomena which ideally should be included in the study.

Developing a systems model provides a framework in which one can perceive important variables and relationships which are being omitted. The output of an investigation should be as thoroughly mathematised as possible. However, the information available to do so is limited. To develop a system model based merely upon presently available information tends to perpetrate both piecemeal studies with narrow results and the status quo with regard to information availability.

There is a feedback between the statement of the problem and the system modelling. The statement of the problem should be constantly reviewed to ensure that the model is appropriate to the problem. Alternatively, as the model develops and existing knowledge about phenomena being studied is organised, the statement of problem should be evaluated for correction. Eventually, through an iterative process, the statement of problem and the system model become balanced and the researcher feels little is to be gained by trying to expand the basic system model and time, manpower and physical facility resources are allocated to the following activities:-

1) Mathematising and calibrating the basic model.
2) Establishing criteria for evaluating performance.
3) Testing and evaluating the given system.
4) Designing alternative solutions.
5) Testing and evaluating performance of the existing system and alternative.

Mathematising and calibrating the basic model involves devising appropriate measuring strategies, collecting information, and then reducing the data to the various relationships specified in the basic model. As this work proceeds, results are fed back into the statement of problem, the basic systems model and resource allocation and, if necessary, modifications can take place.

At the time that the basic model is being calibrated and revised, consideration should be given to establishing criteria by which the performance of existing and future systems and system components can be evaluated. Criteria for design and evaluation are also fed back into the statement of problem and the basic model. The criteria must conform to the goals of the study. Comparing criteria with the problem and the model gives the researcher an opportunity not only to evaluate the criteria, but also to expand or modify the problem and the basic model.

Performance of the given system and its components are evaluated after criteria for performance are established. The criteria for performance are established not only in terms of what the existing systems can do, but also in terms of the basic problem and what a system should or could do. Thus it may be discovered that the given system does not perform at all in terms of one or more of the standards which have been set.

The next step is to design alternative systems which improve poor performances and provide additional performance properties where none existed before. Imperfections and lack of performance in the existing system do suggest kinds of innovations to be made. If the initial statement of the problem is correct, likely innovations will not be arbitrarily excluded.

The costs of alternative systems are determined as a part of the process during which, alternative solutions are designed. The performance of each of the alternatives is compared with that of the others and with the existing system. During this phase, the mathematisation of the given system yields great benefits. Using simulation models provides the best method to date for evaluating alternatives without actually building them.

The evaluation strategy takes two forms. First, with a given budget, one may ask which alternative gives the highest performance. Second, with a given level of performance needed, one may ask which alternative costs the least. Both questions are appropriate depending on the nature of the problem. After alternatives are evaluated and one is selected, it only remains to implement the decision that has been made.

The application of such a methodological framework, or some similar framework, can be argued to be of paramount importance in the development of road accident analysis, for once the level and type of relationships between different elements has been recognised, the evaluation of a set of remedial procedures immediately becomes more objective than under any other approach. However, before this type of approach can be applied, one other problem must be settled. This involves what elements to include in the system, or how the system should be "bounded." Although it is possible to trace inter-relationships between everything and everything else, if any system is to be studied effectively, boundaries for the study have to be established. A system should be bounded in order to include all of the entities and activities which are meaningfully related, and which are relevant to the problem statement. Mathematically this problem poses no difficulties, since the boundaries are given by defining certain elements as being in the system and other (relevant) elements as belonging to the environment. However, in order to use the mathematical properties of system analysis, it is necessary to have some operational method of defining boundaries. In most instances, this has to be done subjectively and boundaries are imposed at points where the researcher believes the problem statement is being satisfied.

Although the use of systems analysis has many inherent practical difficulties, not least of which is the vast amount of research necessary before it can be made productive, it is suggested here that in the long term, it is by far the best and most meaningful approach at present available. For this reason, the basic need in road accident analysis now should be the formulation of mathematical relationships between the various elements of the road accident system and a greater understanding of the problem, (possibly through better modelling) in general.

Chapter 4 THE AIMS AND SCOPE OF THE STUDY

4.1 INTRODUCTION

Having described in the previous two sections the past and present situation regards the growth in road mileage, and therefore road accidents and casualities, it is necessary at this point to elucidate the aims of this research in more detail, and place them in context within the road accident problem as a whole.

4.2 RESEARCH AIMS

It has already been pointed out that there are two basic limitations within present road accident research; the lack of a suitable methodological framework and the sparseness of suitable background information, which would allow the formulation of satisfactory functional relationships between variables under study. Therefore, the essential need at present in road accident analysis is to try and overcome these present limitations, both theoretically and empirically. This situation is well described by Mackay et al (1969):-

"The failure to apply rational causation frameworks to accidents has, therefore, led to "common-sense" evaluations of often complex multi-factorial sequences being treated with more importance than they deserve. Thus, not only drivers and by-standers feel competent enough to present their own solutions for the cause and prevention of accidents, but often the police, local authorities and other public bodies have accepted such "common-sense" approaches....."

Noting these observations, the aims of this research can be summed up within two categories. On the one hand it has tried to review and criticise the major approaches in road safety analysis to date, whilst ultimately suggesting systems analysis as a suitable methodological framework. On the other hand meanwhile, and perhaps of more immediate impact, this research has tried to look at the relationship between road accidents and urban structure by studying the variation in the volume of road accidents between different urban areas. This was done by observing the calculated variations for each urban area between some derived "expected" number of accidents or "norm" and the actual observed number of road accidents for that same urban area. The logic behind this approach is that if towns have different responses to the same set of controls, then it should be possible to identify further influential factors which have not so far been considered, and group (classify) towns according to their varying response rates. This classification process besides being essential in any theory building, also enables individual towns to be taken and used as representative of their particular group. The importance of this is shown by the fact that further investigation of these towns should present representative problems, and perhaps solutions unique to their group. In other words, classification accepts that different urban areas will have different types of road accident problems for which "national" or even "regional" solutions might prove useless.

The structure of any urban area was defined by means of 18 independent variables which included variables describing the physical structure, social structure, and socio-economic structure of all urban areas. However, since these variables are described later[1] in this report no more explanation will be made at this point.

These structural variables were chosen because it is often claimed that the only real solution to the road accident problem in urban areas is that of total traffic-pedestrian segregation. Whilst this is almost possible in the new planned towns, especially the second generation new towns such as East Kilbride and Cumbernauld, it is virtually impossible in the older more established urban areas. For these towns, if anything more than piece-meal redevelopment is required, some compromise solution becomes necessary, and it would seem that this means some manipulation of one or more of these structural variables.

Similarly in road accident analysis, there are two basic lines of thought. One group argues that the quickest way to solve the road accident problem is through the improvement of the physical environment, and the greater application of engineering measures in general. The second group meanwhile,

---

1 See Chapter 6 "THE DATA"

claim that the road accident problem is the result of human psychological deficiencies. Thus one of the arguments put forward by some researchers (e.g. Nader 1965, Cohen and Preston 1968) is that any improvement to a stretch of road etc. through engineering measures, whilst making that stretch of road "physically" safer, will merely entice the motorist to lower his own responsibility, and therefore possibly maintain the same level of road accidents as was noted before the improvement. As Preston states, "It would seem that given safer conditions, some drivers will increase speed until the danger is as great as it was before." (Cohen and Preston op.cit.). If this is true, then this "feedback" mechanism between the road standard and the human psychological behaviour is of paramount importance and interest.

This dichotomous approach to the effect of road safety improvements is perhaps epitomised by the book of Cohen and Preston mentioned above. Because of this difference of opinion, the authors found themselves unable to complete one piece of work and instead settled for 'two distinct viewpoints side by side.' (op. cit.). Cohen's approach revolves around the need to study the psychological aspects of road safety, especially in relation to the study of the "motorist as a centaur" and the fusion of man and machine into some kind of "Bio-robot model." That is, Cohen believes that the practice of sub-dividing the system into individual sub-systems, especially those relating to the vehicle and the driver, has confused the issue rather than simplified it, due to the fact that the response of the newly created bio-robot is not equal to the sum of the two wholes! It is therefore necessary, Cohen argues, to study and understand the responses of this twentieth century creature before any improvement can be made in the road accident problem. Implicit in this argument, however, is the contention that until this "bio-robot" is totally understood, no improvements in the road accident system will have any worthwhile effect upon the ultimate list of road casualties.

Conversely, Preston puts forward the argument that although the solution to the road accident situation as a whole is through the increased understanding of various aspects of the system, short term improvements can be, or will be, beneficial if the more urgent problems are given concentrated attention. Accordingly she indicates nine areas where improvements could lead to immediate overall reductions in casualties[2]. Attempting to quantify these reductions she states, "If all these measures were carried out, then instead of nearly 8,000 people killed in a year there would be about 2,250, that is, somewhere in the order of 5,750 lives a year could be saved." (Cohen and Preston op. cit.).

Because of this divergence in opinion, exemplified above, it becomes necessary, therefore, before going any further into the psychological factors affecting driving behaviour, to determine which of these two lines of thought appear to be the more acceptable.

---

2 Preston's nine areas of improvement were:- (i) Use of Seat Belts, (ii) Drink and Driving legislation, (iii) Enforcement of speed limits, (iv) Training and other safety measures for motor cyclists, (v) Parking restrictions, (vi) Safety helmets for pedal cyclists, (vii) Surveillance of standards of goods vehicles, (viii) Provision of pedestrian facilities, (ix) Enforcement of road safety legislation.

In the case of environment and engineering measures, the answer to this question is of the utmost importance and relevance to future planning proposals. This is so because a lot of new town developments have been planned and developed with explicit aspirations to road safety through partial or total segregation. Thus if the physical structure of an urban area is important in causing road accidents, then these new towns should have a distinctly better road accident record than other comparable non-planned urban areas. If, however, such a situation is found to be non-existent then there must be some doubt as to whether the available capital has been expended in the optimum manner.

4.3 SPATIAL STUDIES - LIMITATIONS AND USE

As stated, the relationship between road accidents and urban structure was studied both spatially and temporally in this research by noting the variations in the volume of road accidents between different urban areas. Whilst this is not a new approach, most studies which have attempted to study areal variations in the volume of road accidents have tended to use national units. (Smeed 1949, Pfundt 1969, Bull 1969 etc.). However, most of these studies have met considerable difficulties due to the varying definitions used by different nations. An obvious example here is related to the definition of fatal road accidents, where the "time period" for a road accident varies from, "at the scene of the accident or immediately afterwards" in Portugal, to "within thirty days of the accident" in Great Britain. Unfortunately, even if this problem can be overcome, there are several other probable difficulties. Thus Pfundt (1969) has pointed out that "a truly scientific verdict on the relative risk of accidents in different countries can only be made by comparing the accident rates on roads of similar type in two or more countries." The problem with this kind of approach, however, is that any results obtained are only applicable to a disaggregated section of the various nations' road networks, and are meaningless as regards overall aggregate comparisons.

What Pfundt is really suggesting, therefore, is that several variables cannot be used experimentally, and should in such an instance be held constant, whilst the influence of other variables can be assessed independently. Although this can only be done in a controlled experimental situation, it can be closely approximated in various statistical techniques where the variables are held "statistically" constant, allowing variations only in the specified variables[3].

Some other researchers have tried to overcome this problem by reducing the scale of the observation units from international comparisons to intra-national comparisons. Unfortunately, these types of comparisons have largely been restricted to inter-state comparisons in the USA, and in many ways these are limited by the same restrictions as the national studies. The intention in

---

3 The technique of "step-wise regression" allows this function and this was one of the reasons for applying this technique in this research.

reducing the size of the unit is to maintain the level of certain national factors constant. Thus in this research, when comparing the response rates of different urban areas in Great Britain, such variables as vehicle inspection, road speed limits, etc. could be ignored, or at least, deemed to be held constant since such variables are imposed equally overall by decree of the national government. The result is that in any such intra-national study, fewer variables have to be incorporated in any model which attempts to describe or explain the road accident situation, than would be necessary in any international comparison.

Another spatial study which indicates further the problems which can arise, was an inter-state comparison study made by Colton et al (1968) when they set out to answer the question, "Do states with compulsory motor-vehicle inspection exhibit lower motor-vehicle accident mortality rates than non-inspection states?" Whilst this study could be generalised as a "cause present" and "cause absent" situation, it is perhaps marginally more sophisticated statistically than earlier studies, (e.g. Coverdale et al 1967), since it does attempt to account for variation in other independent variables (which could otherwise have affected the final results) through the use of a "weighted co-variance analysis." However, even with the application of this technique to remove the influence of these disturbing variables[4], the authors had to accept in their

4 These disturbing variables which Colton took account of in this study were:-
1) Log population density,
2) Per capita income,
3) Other accident mortality rates.

conclusions that "The evidence on the role of inspection in motor vehicle accident mortality in this report is retrospective and circumstantial. Further investigation is needed to classify this association." (Colton et al 1968).

The conclusion to be drawn from this study, therefore, is that in any areal comparison study, the necessity to control all variables which may "disturb" the actual relationship being studied, is paramount in importance. This is especially true where the areal units being used have been chosen quite subjectively. Once again this point is stressed very strongly by Colton (op. cit.):-

"In examining the role of inspection in motor vehicle accident mortality, it is important to recognise that the inspection states are a self selected group. Hence, the comparison of mortalities in inspection and non-inspection states should account for the possibility that the inspection and non-inspection states may be very different in several ways other than inspection, and that perhaps these other variables explain the observed mortality difference." (Colton op cit)

In most of these inter-state studies, the main aim has been, therefore, to discover the influence one variable has upon the volume of road accidents. That is, there has always been an implicit causal approach, coupled with the need for some evaluatory technique, to test the effect of any one counter-measure proposal. Thus most studies using motor vehicle inspection as their causal variable, have consistently tried to evaluate the effect upon the volume of road accidents of the introduction of compulsory inspection at the national level. However, because of the complexities of the road accident system, noted above, such evaluations have proved impossible leaving the results of these studies as minor indicative trends.

One way in which this evaluation process could be simplified is by introducing some intervening performance criteria, which has to be satisfied if the countermeasure introduced is to have any effect upon the road accident volume in the long term. For example, where the influence of motor vehicle inspection is being studied, the intervening criteria could be the improvement in the mechanical condition of the vehicle population. If this criteria is not satisfied, then it is obvious that the difference say between a non-inspection state and an inspection state is caused by some variable other than that of motor vehicle testing. (See McCutcheon et al 1969).

Since it has been shown that areal studies have several inherent problems, it could be asked what point is there in such studies. Essentially comparisons of geo-political units serve two functions. Not only do they reveal the nature and underlying structure of the accident generating system, but they also help spotlight differences between one area and another with the eventual purpose of initiating beneficial changes. In other words, spatial comparisons can be viewed as a search procedure for new relationships within the road accident system, whilst at the same time observing the actual performance of the various individual systems, and the relative effectiveness of countermeasure treatments.

The classic piece of work on the effectiveness of countermeasure treatment was that by Recht (1965). In this piece of work, whilst he accepted that most studies on road accident analysis had been aimed at explaining the causes of road accidents, he emphasised that his research was aimed at studying the influence of various activities which should directly or indirectly contribute to accident prevention. Thus, by rejecting the implication of causality, Recht was able to look at the consequence of various countermeasures which had been introduced in the American states under observation. Using multiple regression he introduced 218 program and non-program items into the model, where program factors were introduced safety countermeasure activities, and non-program factors were those relating to the environmental and physical conditions of the states. In arriving at his conclusions, Recht derived 110 regression equations with varying combinations of program and non-program items. The principal findings of the research was that the major proportion of the differences in accident experience is associated with the differences in the non-program factors, particularly vehicle density, age of population etc. However, when the fatality experience was adjusted for the non-program factors, there were many program activities that were significantly associated with the death rate.

The results of this piece of work by Recht are very important to this research, since most of Recht's "non-program" factors are included within the present research's structural variables, seemingly indicating the importance within the road accident system of these such variables.

Summarising, therefore, the study of spatial variations from one area to another can be an essential method in the analysis of the road accident problem. Such studies are, "well adapted to reveal the effects on accident generation of differences in environment, geography, and transport arrangements, that exist between areas. Once these structural factors are identified - that is a model exists - then they can be put under statistical control in order to compare the changes in accident and injury from year to year, in the same geo-political area." (Katz 1971).

As has been pointed out previously, most spatial studies have been concerned with the macro-level, or national level, but this does not mean to say that comparison between smaller units is unacceptable. In fact in many ways the spatial comparison of urban areas within one nation has several advantages over the larger units. For example, it can be argued that within one nation several variables are going to remain reasonably constant, thereby reducing the problems of experimental control. Indicative of these variables are; traffic laws, vehicle laws, social characteristics etc., all of which would theoretically have to be controlled in any international study.

4.4 SCOPE OF THE STUDY

This study of the spatial variation in road accidents between different urban areas in England, Scotland and Wales was effected by sampling a hundred urban areas with a population level greater than 20,000 (1969). An urban area was defined for practical purposes as that area contained within the relevant local government authority. In England and Wales, this included County-Boroughs, Non-County-Boroughs and Urban District Councils. In Scotland, the relevant authorities were the Large and Small Burghs.

In all, these urban areas produced a "total population" of 412 urban areas, from which the 100 sampled areas produced a sample size of approximately 25%. Because of the positive skewed distribution associated with the population level of these urban areas, a random-stratified sampling technique was applied, using the population levels of 100,000 and 50,000 as strata limits. The resultant list of a hundred urban areas is shown in Table (4.4.1).

In addition to this study of the spatial variation in road accidents, a temporal analysis was also used in order to study the influence of the relevant structural variables over time. Originally, it had been hoped that it would be possible to use a ten year period for observation 1960 - 1970. However, due to the lack and incompatibility of the data, this intention proved impossible. Instead the influence of the structural variables

1	Aylesbury	34	Dewsbury	67	Oxford
2	Barnsley	35	Dover	68	Plymouth
3	Barry	36	Dundee	69	Pontefract
4	Basildon	37	Durham	70	Poole
5	Bath	38	East Kilbride	71	Portsmouth
6	Birmingham	39	Edinburgh	72	Preston
7	Blackburn	40	Exeter	73	Reading
8	Blackpool	41	Farnworth	74	Rochdale
9	Bolton	42	Glenrothes	75	Rugby
10	Bournemouth	43	Glocester	76	Ryde
11	Bradford	44	Grimsby	77	Salford
12	Brentwood	45	Guildford	78	Sheffield
13	Bridgewater	46	Harlow	79	Southampton
14	Brighton	47	Hastings	80	Southend
15	Bristol	48	Hemel Hempstead	81	Stafford
16	Burnley	49	Hereford	82	Steverage
17	Buxton	50	Horsham	83	Stirling
18	Cambridge	51	Huddersfield	84	Stoke
19	Canterbury	52	Inverness	85	Stretford
20	Cardiff	53	Ipswich	86	Sunderland
21	Carlisle	54	Kings Lynn	87	Sutton Coldfield
22	Chelmsford	55	Kirkcaldy	88	Swansea
23	Cheltenham	56	Lancaster	89	Swindon
24	Cheshunt	57	Leeds	90	Taunton
25	Chester	58	Leicester	91	Thornton Cleveleys
26	Chesterfield	59	Lincoln	92	Welwyn G.C.
27	Colchester	60	Liverpool	93	West Bridgeford
28	Corby	61	Luton	94	Whitehaven
29	Coventry	62	Manchester	95	Winchester
30	Crewe	63	Newport (I. of W.)	96	Wolverhampton
31	Cumbernauld	64	Norwich	97	Worthing
32	Dartford	65	Nottingham	98	Wrexham
33	Derby	66	Oldham	99	Yeovil
				100	York

TABLE 4.4.1. THE 100 SAMPLED URBAN AREAS.

was studied only over the five year period 1966 - 1970, with disaggregation possible only for the two years 1969 and 1970.

4.5 SUMMARY OF RESEARCH AIMS

Before moving into the main text of this report, it would seem advantageous to draw together in a concise form all the aims and intentions so far mentioned. These aims can be classified under 6 headings, and are given below:-

(i) To review the present methodology used in Road Accident Analysis, suggesting the use of systems analysis as one of the more encouraging developments for the future.

(ii) Through the technique of multiple regression to calculate, an expected level of road accidents for each urban area, and therefore the corrected spatial variation in road accidents.

(iii) To discover which structural variables account for the greatest variation in the spread of road accident volumes.

(iv) To test to see if individual urban areas can be classified according to their response rate to the road accident problem, and therefore to discover any "group" problems, or "group" solutions.

(v) To observe the spatial spread of the residuals from the proposed "model" and discover any further variables which may appear important or relevant.

(vi) To study the importance of the various structural variables over time, in order to see how these variables react to different stimuli.

These then are the major aims of this research and in the remaining sections of this report they are elucidated in much more detail, along with the conclusions which can be drawn from the relevant results.

Chapter 5 THE ROAD ACCIDENT SYSTEM

5.1 INTRODUCTION

It has been argued in previous sections of this report, that the application of systems analysis provides the most hopeful trend for the development of the subject in the near future. If this is so, it becomes necessary to define and describe what is meant by the road accident system. That is, one has to know what elements are defined within the system, and what type and direction of relationships exist between these elements. Although it is impossible to do this in any detail at the present level of road accident research, this section does attempt to give a simple postulated system which can serve as a first approximation for the eventual true road accident system. Initially however, several definitions, relevant to this section and the main body of the research in general, are expounded.

5.2 ACCIDENT DEFINITIONS

The definition of what constitutes an accident is one of the major operational difficulties in any research dealing with the subject. A dictionary definition of an accident is, "...an event without apparent cause,....unexpected, unintentional act." However, in the instance of road accidents, the majority of such events are neither unexpected nor without cause. In fact, it is usually true to state that a large number of road accidents could be avoided by greater design control of the environment, and more self control by the human beings involved. For this reason, various authors have proposed that an accident should not be defined as a unitary concept, but that a "range" as opposed to a "class" definition should be adopted. (Suchman 1961, McFarland 1963). Thus Suchman suggested that the following criteria should be adopted in definition:-

(i) Degree of expectedness - the less anticipated an event is the more likely it is to be labelled an accident.

(ii) Degree of avoidability - the less the event could be avoided, the more likely it is to be described as an accident.

(iii) Degree of intention - the less the event involves deliberate action, the more the event is accidental.

Accordingly Suchman defined an accident as "an event characterised by a low degree of expectedness, avoidability and intent." This definition is an interesting one since it emphasises the precedent factors leading up to an event and eliminates any factors relating to damage or injury. It is, therefore, applicable to studying near misses as well as events involving damage or injury.

However, Barmack (1962) has suggested that it is not known what definitions provide the basis for obtaining the most useful information for understanding and preventing accidents, nor is it known what groups or classifications of accidents can best maximise the possibility of identifying common factors or combinations of factors.

For these reasons, most researchers have had to use a definition of an accident which has provided a practical range of possibilities for carrying out the particular piece of research under study. Thus Mackay et al (1969) defined a road accident as:-

"A traffic incident involving moving vehicles which results in damage to those or other vehicles, or road furniture, or results in injury to road users."

Similarly the definition used in this research bears many similarities to the above definition; however, since most of the accident data was provided by the Transport and Road Research Laboratory, the definition used is related to that suggested by the Ministry of the Environment in "Road Accidents" 1970. (HMSO 1972). Thus an accident is defined as:-

"An event in which at least one motor vehicle is involved in a collision on the public highway, and which results in one or more persons being injured."

Therefore, within the limits of this research, only personal injury accidents are considered, basically because most of the data on damage only accidents is incomplete.

Although such a definition leaves much to be desired, it is perhaps the most functional one available at present. Similarly as Suchman has also pointed out, "as our knowledge of causal factors increases we are more likely to describe an event in terms of those factors and less likely to label it as an accident." (Suchman 1961). A similar point is also argued by McGlade et al (1965) when they state that "for human causation research purposes, an 'accident' must be defined as a predictable event subject to human control and therefore avoidable." The relevant fact therefore is whatever definition may be used, as the subject develops so must the definition, and therefore what is suitable for one piece of research may be totally inadequate for another. For this reason, functional definitions as used above would appear to be the most practical in such a piece of research as studied here.

Other accident definitions, relating to the severity of a road accident are also the same as those used by the Department of the Environment:-

(i) FATAL ACCIDENT:- is an accident in which one or more persons die within 30 days of that accident.

(ii) SERIOUS ACCIDENT:- is an accident in which a person is injured such that he (or she) has to be detained in hospital as an "in-patient," or suffers any other serious injury (fractures, concussion, internal injuries, crushings, severe cuts and lacerations) whether or not the individual is detained in hospital.

(iii) SLIGHT ACCIDENT:- is an accident in which a person receives an injury of only minor significance such as a sprain, or a bruise.

5.3 DEFINING AN URBAN AREA

The problems inherent in the definition of any urban area have been well discussed by geographers for many years, and as such will not be argued at length in this report. (See Childe 1936, Weber 1958, Tisdale 1942, Dickinson 1947, Smailes 1947, Berry 1968. etc.). However, in the philosophical literature there has been a consistent growth in concern about "entitation." That is "any attempt to establish or explore relationships needs first to specify the objects which are to be related in some appropriate manner and the test of whether the definition is appropriate is that it should be meaningful in terms of the function and the performance of the objects." (Robson 1973). The stringency of this need is greatest when functional relationships are under study, and less so when the study is solely concerned with description. Unfortunately, the opportunities for such satisfactory definitions of urban areas are very limited and once again the problem is usually solved by movement to some operational definition. However, even with this approach there are numerous difficulties. For example, Dickinson's "spheres of influence" (1947) are restricted by the fact that any urban area has many such spheres dependent upon which function of the urban area is used as the relevant yardstick. Similarly Berry's attempt to define an "urban field" in place of the more familiar "urban area" using journey to work data in the USA (1968), merely complicated the situation even further from the point of view of definitions. Based on these calculations Berry showed that 96% of the total population of the USA lay within the urban regions which he produced, against only 67% within the census-defined Standard Metropolitan Statistical Areas.

It is obvious therefore that no definition of an urban area is going to be totally satisfactory although some will obviously be better than others. Although this problem is discussed further in relation to the problems of scale[1] in this report, suffice it to say that the definition used in this research was a simple operational one. An urban area was defined as that area lying within a relevant urban local authority area. In England and Wales these included County Boroughs, Non-County Boroughs and Urban District Councils, and in Scotland the relevant areas were large and small Burghs.

5.4 A SUGGESTED ROAD ACCIDENT SYSTEM MODEL

It has already been stated that it is possible to postulate a simple road accident system model which could act as a simple first approximation model. Such a model can predict the direction of relationship between entities within the system, and also define what entities should be "closed" within the system. It cannot however go into more detail about these relationships as regards type and strength. The system presented in (Fig. 5.4.1) is such a simple model, and is the one used throughout this piece of work.

---

1 See page 129.

Fig. 5.4.1. THE ROAD ACCIDENT SYSTEM

In many ways this system approximates the model suggested by Gordon (1949) who was one of the advocates of an epidemiological approach to accident analysis, with the application of techniques and methods available for the study and control of mass diseases. Thus Gordon found that when data on accidental injuries were analysed in a standard epidemiological manner, injuries appeared to follow related biological principles. The distribution of injuries showed characteristic variations in regard to age, sex and type of accident etc. Accidents occurred at different rates, and in relation to different agents and circumstances among different groups of the population.

By considering accidents as an ecological problem and by using the concept of causation "as a combination of forces from at least three sources - the host, the agent and the environment in which host and agent find themselves," Gordon (1949) was able to recognise patterns in the distribution of otherwise seemingly chance events.

The host, agent, environment model suggested by Gordon, approximates the proposed system most in its specification of its environment factors. Gordon splits his "environment" into three relevant sections:-

(i) Physical environment which relates to matters of weather, season and any physical features which can be natural or man made, concerning the world where man lives and drives or travels.

(ii) Biological environment - which included all living things connected with man excluding himself, and

(iii) Socio-economic environment which is the sphere of the environment where man is affected by the relevant interactions with other men.

It will be noted that in the case of road accident analysis, whilst such a sub-division is theoretically workable, the physical and socio-economic environments are by far the more important of these three sub-sections. Therefore in the system model described below, the biological environment is ignored.

It can be seen that the road accident model proposed here has been split into six different sub-systems, where a sub-system can be defined as a collection of closely related elements. However, of these six sub-systems, it will be noted only three directly influence the level of road accidents. The remaining three sub-systems operate upon this accident level through one or more of the other sub-systems.

In order to explain clearly this proposed model, it has been disaggregated into its relevant sub-systems, each of which will be described below. However, before that can be commenced, it must be pointed out that this model is a model which is only related to the "urban" road accident system. Obviously there will be an equivalent "rural" road accident system, which whilst consisting of the same basic sub-systems, must have different levels and directions of relationships between the different elements. Similarly, if an overall road accident system is to be built, these two models, rural and urban, will have to be combined and the complete pattern studied.

(1) PHYSICAL STRUCTURE SUB-SYSTEM

Since the basic aim of this research was to look at the relationship between road accidents and urban-structure, this sub-system, along with that relating to the socio-economic structure, are the two sub-systems studied in some detail. This physical structure sub-system can in actual fact be sub-divided even further into three smaller units:-

a) MACRO-STRUCTURE - DESIGN

b) MACRO-STRUCTURE - SIZE

c) MICRO-STRUCTURE

Although the third component - micro-structure - is obviously important in the overall road accident study, this research was more concerned with the macro factors. Accordingly, there is no detail from this research available concerning this facet of the system. However, research has been conducted in this area by other researchers and one study by the U.S. Bureau of Public Roads (BPR) is of particular significance.

This BPR study which elucidates many aspects of multi-factor modelling, was intended as an "Analysis and Modelling of Relationships Between Accidents and the Geometric and Traffic Characteristics of the Interstate Highway System" (Cirillo et. al 1969). In other words, the study tried to measure the effect of various geometric design features on accidents, and also evaluate different interchange designs.

The model used in this BPR study was a standard multiple regression model, and 75 independent variables were introduced in the first stage of the analysis, being reduced to 20 factors in the final analyses. Besides pure geometric design features (shoulder width, curvature etc.) the independent variables also covered such features as traffic variables, (average daily traffic etc.), road features (lighting etc.), and environmental factors (police hours etc.). Although such variables were therefore mainly traffic and road variables, 5 of the final 20, and 10 of the original variables were concerned with "other" items such as the percentage of commercial vehicles, the number of readable information signs, and if the location was rural or urban.

The major findings of this research, whilst being important in the main body of road accident research, has this to say about the importance of geometric design factors (micro-structure features),

"Geometric and traffic variables together account for only a fraction of the variance in accidents. The remaining portion is presumed to be accounted for by major factors not covered in this study, such as driver, vehicle and weather conditions." (Cirillo op. cit).

On these lines it is interesting to note that when the BPR study looks at the distribution of traffic and accidents over the 24 hours of the day, it points out that the time period from midnight to 4 a.m., although having the lowest average daily traffic of the day, has the highest rates of accidents per mile of travel. Obviously, road geometry has not changed during this period of time and also, since the traffic variables have been controlled statistically, it must be that driver and vehicle variables had not yet been accounted for, whilst logic, as well as knowledge of the work by Tamburis (1969), indicate that the interaction of specific geometric design features with driver characteristics may well be of considerable significance in elucidating this dilemma. Thus as Katz states:-

"It is an open question whether the BPR study has in fact, "expressed the significant relationships between the design and traffic characteristics and.....accident experience" as claimed - or whether the truly significant relationship between these variables isn't still to be found in the interaction with the "major" variables not included in the investigation." (Katz 1971).

In the light of these reservations, therefore, perhaps the most that can be said about the study is that considerable further information has been added to road accident research, data which when linked with additional similar information in other sections of the road accident system, will serve as the basis for the creation of more advanced and sophisticated models in this field of research.

Returning to the macro features of this sub-system, the relationships between these various variables will be discussed in the main body of this report, although certain superficial points can be made at this stage in the discussion.

The first point to be made is the high degree of correlation that is present between the two components as well as within each component. In the diagram (Fig. 5.4.1) the highest correlation coefficient is shown between any two variables, along with the direction of the relationship. Thus the degree of open space variable has its highest correlation coefficient with the density variable (r = -0.31) whilst the density variable has its highest correlation coefficient with the Road Density Index (which is an indicator of road junction density) where r = +0.84. As stated, this diagram only shows the highest correlation coefficient between variables, and although this gives some indication of the relationships present, it is much better if all the correlation coefficients are shown. Therefore, Table (5.4.2) presents the correlation matrix for all the 22 variables studied and used in this system model.

The second feature to note about the macro components and the physical sub-system in general, is the way it is postulated to influence the level of road accidents. Not only is it postulated that they directly influence the volume of road accidents, through the degree of interaction realised within an urban area, and the nature of the actual road network, but

	POP	AR	DEN	PTF	TTF	ETT	COW	MS1	SR1	TC1	C1	DOS	SH1	AV1	DV1	TD1	HT1	CHS1	PP1	PS1	R1	RD1
POP	1.00																					
AR	0.89	1.00																				
DEN	0.51	0.18	1.00																			
PTF	0.98	0.90	0.55	1.00																		
TTF	0.97	0.81	0.54	0.94	1.00																	
ETT	0.63	0.53	0.47	0.66	0.67	1.00																
COW	-0.25	-0.15	-0.34	-0.28	-0.25	-0.18	1.00															
MS1	0.18	0.09	0.24	0.18	0.21	0.11	-0.65	1.00														
SR1	0.03	0.03	-0.01	0.02	0.06	0.01	0.15	0.34	1.00													
TC1	-0.18	-0.18	-0.14	-0.19	-0.16	-0.15	0.01	0.06	0.02	1.00												
C1	-0.03	0.02	-0.12	-0.02	-0.04	0.08	0.12	-0.09	0.13	0.62	1.00											
DOS	-0.10	0.12	-0.31	-0.09	-0.12	-0.03	0.30	-0.21	-0.06	-0.18	-0.08	1.00										
SH1	-0.17	-0.21	0.07	-0.19	-0.18	-0.29	0.25	-0.09	0.20	-0.14	-0.08	0.14	1.00									
AV1	-0.11	-0.14	0.03	-0.11	-0.13	-0.19	-0.19	-0.08	-0.64	-0.19	-0.16	0.07	0.14	1.00								
DV1	0.10	0.03	0.11	0.10	0.14	0.21	-0.01	-0.11	0.49	0.08	0.15	-0.09	-0.03	-0.64	1.00							
TD1	0.99	0.90	0.50	0.97	0.97	0.62	-0.26	0.18	0.00	-0.19	-0.04	-0.10	-0.18	-0.10	0.09	1.00						
HT1	-0.08	-0.04	-0.14	-0.07	-0.09	-0.05	0.08	-0.39	-0.54	-0.04	-0.03	-0.00	-0.28	0.40	-0.39	-0.07	1.00					
CHS1	-0.163	-0.09	-0.26	-0.19	-0.20	-0.31	0.31	-0.22	0.09	-0.06	0.09	0.13	0.43	0.08	-0.04	-0.17	-0.21	1.00				
PP1	0.12	0.06	0.27	0.16	0.17	0.38	-0.05	-0.01	-0.04	-0.04	-0.13	-0.07	-0.21	0.00	0.04	0.13	0.08	-0.26	1.00			
PS1	0.08	0.05	0.08	0.09	0.09	-0.01	-0.02	0.05	0.01	0.18	0.02	-0.17	0.01	-0.05	0.03	0.09	-0.13	0.18	0.16	1.00		
R1	-0.02	-0.02	0.02	-0.03	-0.03	-0.23	-0.38	0.25	-0.14	0.14	-0.06	-0.01	0.07	0.20	-0.23	-0.02	0.02	0.07	-0.04	-0.13	1.00	
RD1	0.29	0.01	0.84	0.33	0.30	0.26	-0.38	0.18	-0.20	-0.06	-0.11	-0.33	-0.06	0.18	-0.05	0.29	0.09	-0.32	0.26	0.01	0.22	1.0

TABLE 5.4.2 CORRELATION MATRIX FOR THE 22 INDEPENDENT VARIABLES (1970)

it is also postulated that they are causal variables as they interact with other sub-systems. The most obvious interaction of this kind is through the human sub-system where it is presumed that the psychological attitude of the driver is influenced by such factors as traffic density, type of road, congestion etc. which in turn, will influence the driver's attitude to "risk taking" and ultimately, therefore, the volume of road accidents.

The third point to make meanwhile, is the direct "feedback" link between the actual level of road accidents and the "control" of the physical structure sub-system of any urban area. Since one can conceive of the road accident system as being a "controlled system"[2] this "feedback" mechanism is of vital importance, since in one way it indicates a certain "tolerance" level within society, whilst at the same time, it acts as a motivator in reconstruction within the physical environment. Thus it can be argued from this point, that this feedback mechanism is essential in determining the level of road accidents (or costs) that any society is liable to deem acceptable, when related to the actual cost of physical improvement and/or redevelopment. The understanding

---

2 There are usually deemed to be four types of systems:-

(i) HOMAEOSTATIC SYSTEM:- This is a system that maintains "a constant operating environment in the face of random external fluctuations." (Rosen 1967). Such systems resist any change in environmental conditions and gradually return to an equilibrium or steady state behaviour after any such change. A spring is a good analogy for this kind of system.

/contd...

of such a feedback mechanism can only come, however, at a much later date in the modelling of the road accident system, when the relationships between all the elements are more clearly understood. When that situation has been reached, the feedback mechanism becomes a type of evaluatory process.

(2) SOCIO-ECONOMIC STRUCTURE SUB-SYSTEM

It can be seen that this sub-system is basically made up of social status and housing variables. Once again since these variables were studied in this research, the highest correlation links are shown both by value and direction.

Although it is postulated that this sub-system does not directly influence the volume of road accidents, it was deemed significant in affecting the ultimate level of accidents through other sub-systems. Thus the level of car ownership will affect the traffic volumes in the physical structure sub-system, and the sex ratio will affect the human sub-system, because of the different accident patterns and rates, noted for male and female drivers. (See McKay 1969, Robertson et al 1966 etc.). A similar argument can be put forward for the two age distribution variables.

---

2 contd. (ii) ADAPTIVE SYSTEM:- is a system for which there exists for each possible input a set of one or more preferred states, or outputs. The system is also such that if at any time the system is not in a preferred state, then the system will so act as to alter the state until one of the preferred ones is achieved. The study of such systems obviously involves the concept of "goal seeking." The concept of a feedback mechanism is obviously vital in these types of system changing the various states by directly affecting the condition of the environment or by altering the perameters of the system itself.

The other major sub-system linked with the socio-economic sub-system is the vehicle system. The line of reasoning here is that the socio-economic level of an urban area, will directly influence the type, age and therefore level of maintenance, of the vehicles used in that area. The lower an area is socially, the higher the proportion of "second-hand" and old cars one would expect to find in the total vehicle population. On these lines, if any confidence can be placed on the tentative results of the motor vehicle inspection studies mentioned previously, then one would also expect a relatively higher degree of road accidents in these areas.

(3) VEHICLE SUB-SYSTEM

Of all these sub-systems within the road accident system, the vehicle sub-system has perhaps commanded in the past, the major amount of research in this field. Unfortunately, this sub-system is complicated by the fact that it can influence road accidents both in the "pre-collision circumstances" and also in the "post-collision circumstances," and as such they ought to be dealt with individually.

---

2 contd. (iii) DYNAMIC SYSTEMS:- are systems through which the feedback mechanism keeps the system moving through a sequence of unrepeated states usually termed the trajectory or line of behaviour of the system. Implicit in such systems is the learning process where the feedback may cause new preferred states to be identified. Examples of such systems include economic growth models etc.

Post collision investigations can be generalised under the heading of protective devices; their use and degree of success in the post-crash condition. Such investigations have therefore studied the use of safety belts and the energy absorbing steering column etc. Although these devices could be tested experimentally in the laboratory, the true degree of their success can only be statistically demonstrated in the field, since in the case of the safety belt there was considerable scepticism as to what would be the belt's real contribution to the overall traffic safety progress.

A good example of the resultant studies, and one which helps to indicate the influence of other variables and sub-systems, upon the vehicle sub-system is the study completed by Bohlin (1967) who looked at 28,000 accidents involving Volvo vehicles in Sweden 1965 - 1966.

The first and perhaps most significant finding was that only 25% of the vehicles with seat belts, actually used them with the result that any effect of seat belts on overall injury

---

2 contd.. (iv) CONTROLLED SYSTEMS:- are those where the operator has some degree of control over inputs. Systems control theory is therefore very important when dealing with these systems. Controlled systems are most frequently found in the fields of planning, and government policy where some inputs are controlled in order to achieve some desired level of output. For further reading on these systems see (Rosen 1967, Ashby 1966, 1963).

must be very limited. Returning to the main findings of the study meanwhile, using a "cause present" and "cause absent" group technique, Bohlin was able to study these two groups and see if there was any significance in the seriousness of injuries sustained. However, before uncritically accepting the findings of this study, which conclusively showed the effectiveness of seat belts as a crash injury reducing agent, it must be pointed out that the two groups used exhibited significant differences in composition. For example, as regards age, in the total group of 37,511 injured front seat occupants, the author states that "the substantial part were the very young drivers aged 18 - 24." On the other hand, since belt usage varied from 17% in the 20 - 25 years age group, to 31% in the 45 - 55 years age group, it is obvious that the distribution of ages in the 9569 injury cases who wore seat belts is entirely different from that in the remaining 27,952 cause absent, non-belted group. If this is coupled with the assumption that people who wear safety belts might prove to be more "careful" persons overall (Morgan 1967) it can be seen that the actual influence of these vehicle variables cannot truly be assessed until the interactive influence of other variables (age, speed, sex, carefulness, etc.), has satisfactorily been accounted for and studied.

Similarly, the studies which have tried to investigate the influence of vehicle factors on the pre-crash conditions, have also been restricted by several difficulties. The major hypothesis here has been to try and determine whether changes in the handling, and performance characteristics of different vehicles, can in any way be a factor in the causation of road accidents. Thus a typical study was that by Cornell Aeronautical Laboratory which attempted, "to find out to what extent power steering cars might differ from standard steering cars in injury producing accidents." (Gensler 1966). Their analysis sub-divides the "with" and "without" groups according to weight of car, age and sex of driver, time and type of accident, impact speed, area of impact, direction of force, roll-over violence and two car - single car configuration. These observations led to two major conclusions.

(i) Power steering cars have more object collisions as contrasted to vehicle collisions, and,

(ii) Power steering cars have more rear end involvement as contrasted to front involvement etc.

However, the complexity of such studies, and therefore this part of the vehicle sub-system as a whole, can be best illustrated by quoting from the Cornell study:-

"None of the findings can be related to the steering mechanism. There is no obvious way by which power steering could be thought to cause object collisions rather than vehicle collisions ....etc. ..... Yet these tendencies are statistically significant..... The interpretation of these findings is made problematic if not impossible by the fact that the data contain no information on the exposure of the two groups of cars.....It may be that power steering cars are on the road ......under different traffic conditions.....Preponderance of object collisions could be expected if power steering cars did more rural highway driving.... These explanations of the observations must remain quite speculative." (Gensler op cit.).

The major conclusions to be drawn from both types of study described above, (those looking at pre-crash conditions, and those concerned with post-crash conditions) is that although one can determine several relationships, the relationships between vehicle factors and other variables cannot be explicitly stated without further data and further study. However, it can be concluded that the vehicle sub-system is definitely affected by the socio-economic sub-system and the prevention sub-system (through police surveillance and control) even though the type, level and direction of these influences cannot be specifically stated at this point in time.

(4) NATURAL PHENOMENA SUB-SYSTEM

Although the influence of this sub-system in Britain on the level of road accidents (in urban areas) can be postulated as being minimal, it should not be discounted altogether.

There are two ways suggested in which this sub-system can influence the volume of accidents. In the first instance it can work through the micro-features of the physical structure sub-system, whereby it makes "normally" safe physical features "less" safe under bad natural conditions. Thus the degree of satisfaction concerning surfacing work can vary according to the level of such things as rainfall, ice, relief etc.

Similarly the natural phenomena sub-system can be influenced due to its interaction with the human system. It should be realised there are two climatic conditions, one inside the vehicle as well as one outside the vehicle. The interior conditions within a vehicle can therefore be important in determining the physiological condition of the driver etc., and therefore the way in which he responds to any incident.

In this research, because of data difficulties, the only variable included from this sub-system was that of rainfall. Although its correlation with the other variables is shown in Table (5.4.2) it would appear that these are rather meaningless, since the above mentioned intermediary variables are not included and therefore the postulated important interactions unavailable.

(5) PREVENTION SUB-SYSTEM

Once again this sub-system can be very influencial in determining the level of road accidents in any area through one or other of the various sub-systems. It would appear that this sub-system can be split into two separate sections which can influence the accident level in different ways.

One section can be described as the police action section. This includes the action of the police in relation to car maintenance, traffic control, and also motoring offences. Thus this section can influence the level of road accidents by means of its interaction with the vehicle, physical and human sub-systems respectively. For example, it can be argued that where a police force is well known for its "unsympathetic" attitude to motoring offences, this should affect the attitude of individual drivers, through the human psychological sub-system. It is interesting to note here that the maximum correlation coefficient for the Police Prosecution Index used in this research is $r = 0.38$ and links this variable with the level of through traffic, in the physical structure sub-system. This could indicate perhaps one of two hypotheses. Either it could be argued that police action tends to be more stringent in traffic control where there is a high degree of through traffic and therefore, possibly traffic congestion, or alternatively it could be argued that where a police force follows a stringent line on motoring offences, the number of police prosecutions is related directly to the volume of through traffic and therefore the volume of strangers through an area who are not conversant with the local police attitude. Whichever argument is accepted however, it is obvious that the police action is definitely interacting with one or more of the other sub-systems.

The second section within this sub-system can be looked upon as a training section. That is, the action of such people as road safety officers, and driver trainer bodies. This section therefore obviously interacts with the human sub-system by trying to alter the psychological attitude of both drivers and pedestrians of all ages, by increasing the level of knowledge and understanding of these people.

Before leaving this sub-system, it should be pointed out once again, the direct feedback link between the actual level of road accidents and the prevention sub-system. The logic behind this is that the prevention organisations are the implementors of the "tolerance" level of society to the volume and cost of road accidents. Thus it can be seen that there is a direct relationship between the volume of road accidents and police action as regards motoring incidents.

(6) HUMAN SUB-SYSTEM

Basically this sub-system is related to the psychological and physiological response of the human individuals to any situation. The importance of this sub-system can be seen by the fact that it is influenced by every other sub-system within the road accident system, and can therefore perhaps be regarded as the central core of the whole system. Unfortunately the internal operation of this sub-system is also the least understood and any studies which have been attempted in this direction have been of very little success. It can be argued, however, that before this sub-system can be satisfactorily "opened" the relevant interactions of the other sub-systems will have to be known, and the understanding of the human system will be the last clue in the problem.

Studies which have tried to look at this sub-system in isolation, have tended to revolve around the various early theoretical models of road accident causation, such as accident proneness and highway hypnosis. Consequently, as these theories have become less acceptable, so have the results of these studies, although once again important background information may have been collected. However, it is still true to state that very little is still known about this sub-system and further advanced work by perhaps psychologists is one of the major necessities in road accident analysis at present.

This chapter has therefore attempted to postulate a simple road accident system, and the various internal sub-systems, which can serve as a first approximation to the final road accident system. It must be emphasised this model only suggests the various directions of the interrelationships and therefore no mathematical functions can be put forward at this time. These functional relationships can only be obtained by further detailed work into the nature of each of the sub-systems. Accordingly, this research has used the suggested framework and has attempted to look at two of these sub-systems; the physical structure sub-system and the socio-economic sub-system, the analysis and results of which, are given in the following sections of this report.

Chapter 6 DATA

6.1 INTRODUCTION

Since the aims of this research involved the study of the spatial and temporal variation in road accident statistics in England and Wales, various sources had to be used for the extraction of the relevant data points. Similarly it was necessary to collect two sub-sets of data, one relating to the dependent variables (road accident statistics) and another relating to the independent or explanatory variables. The source and definition of these variables will be expounded in more detail in the following sections of this report. Numerically twenty two independent variables (Table 6.1.1) and forty five dependent variables (Table 6.1.2) were assertained for each of the one hundred areas sampled. In order to elucidate the explanation of these two sets of variables, each will be referred to individually within each relevant sub-set.

6.2 DEPENDENT VARIABLES

The major source for this sub-set of variables was obtained from the TRRL's computerised output of the Police Stats 19 form. Use of this source of data enabled uniformity of data spatially, whilst unfortunately restricting the temporal analysis due to the variation in the availability of the same data population for different years. Because of this difficulty, it was only possible to attain heavily disaggregated data for the years 1969

1. POP:- POPULATION LEVEL
2. AR:- AREA OF URBAN DISTRICT
3. DEN:- DENSITY OF POPULATION
4. PTF:- PEAK TRAFFIC FLOW
5. ETT:- ESTIMATE OF THROUGH TRAFFIC
6. TTF:- TOTAL TRAFFIC FLOW
7. COW:- CAR OWNERSHIP
8. MSI:- MEAN SOCIAL INDEX
9. SRI:- SEX RATIO INDEX
10. TCI:- TOWN CLASSIFICATION INDEX
11. CI:- COMPACTNESS INDEX
12. DOS:- DEGREE OF OPEN SPACE
13. SHI:- STANDARD OF HOUSING INDEX
14. AVI:- AGE VULNERABILITY INDEX
15. DVI:- DRIVER VULNERABILITY INDEX
16. TDI:- TOTAL DWELLING INDEX
17. HTI:- HOUSING TENURE INDEX
18. CHSI:- CHANGE OF HOUSING STOCK INDEX
19. PPI:- POLICE PROSECUTION INDEX
20. PSI:- POLICE SEVERITY INDEX
21. RI:- RAINFALL INDEX
22. RDI:- ROAD DENSITY INDEX

TABLE 6.1.1. LIST OF 22 INDEPENDENT VARIABLES (1970)

1.	TC70:	Total Casualties
2.	TA70:	Total Accidents
3.	TFS70:	Total Fatal and Serious Casualties
4.	TPED70:	Total Pedestrian Casualties
5.	TMC70:	Total Motor cycle Accidents
6.	TPCY70:	Total Pedal cycle Accidents
7.	TDRC70:	Total Driver and Rider Casualties
8.	TCPED70:	Total Child Pedestrian Casualties
9.	TYDR70:	Total Young Driver and Rider Casualties
10.	TJUN70:	Total Junction Accidents
11.	TPDCR70:	Total Accidents at Pedestrian Crossings
12.	TTRN70:	Total Accidents due to a
13.	TRT70:	Total Accidents at Roundabouts
14.	TTJN70:	" " at T Junctions
15.	TYJN70:	" " at Y "
16.	TXJN70:	" " at X "
17.	TUCRD70:	Total Accidents on Unclassified Roads
18.	TCRD70:	" " " Classified Roads
19.	TCPP70:	Total Casualties per 10^3 population
20.	TAPP70:	" Accidents " " "
21.	TFSPP70:	Total Fatal and Serious casualties per 10^3 population
22.	TCVM70:	Total Casualties per 10^6 Vehicle Miles
23.	TAVM70:	" Accidents " " " "
24.	TFSVM70:	Total Fatal and Serious casualties per 10^6 Vehicle Miles
25.	SR70:	Fatal and Serious casualties/Total Casualties.
26.	CPEDCR70:	Child pedestrian casualties/Total pedestrian casualties
27.	YDRCR70:	Young driver and rider casualties/Total driver and rider casualties.
28.	TVNAR70:	Total turning accidents/Total accidents
29.	PDCRAR70:	Total accidents at pedestrian crossings/Total accidents
30.	TAMVC70:	Total accidents on unclassified roads per mile of unclassified road.
31.	TAMC70:	Total accidents on classified roads per mile of classified road
32.	RTAR70:	Total roundabout accidents/Total junction accidents
33.	TJNAR70:	Total T Junction " " " "
34.	YJNAR70:	" Y " " " " "
35.	XJNAR70:	" x " " " " "
36.	TPEDPP70:	Total pedestrian casualties per 10^3 population
37.	PEDCR70:	Total pedestrian casualties/Total casualties
38.	PCYWT70:	Total pedal cycle accidents/pedal cycle work trips (10^3)
39.	PEDWT70:	Pedestrian casualties/Total walk work trips (10^c)
40.	PCYAR70:	Total pedal cycle accidents/Total accidents
41.	MCAR70:	Total motor cycle accidents/Total accidents
42	TDRMV70:	Total driver and rider casualties per 10^6 Vehicle miles
43.	DRCR70:	Total driver and rider casualties/Total Casualties
44.	TCTA70:	Total Casualties/Total Accidents
45.	TJNAR70:	Total Junction accidents/Total Accidents.

TABLE 6.1.2. LIST OF 45 DEPENDENT VARIABLES 1970.

and 1970. However, the temporal analysis was extended for three of the basic dependent variables; total casualties, total casualties per thousand of population and total casualties per vehicle mile (10^6), for the years 1966 - 1970. Such variables related to all accidents involving fatal, serious and slight casualties for each of the sampled local authority areas. Data prior to 1966 was only available for fatal and serious accidents thereby restricting the use of these years in any extended, or meaningful temporal analysis. Therefore it is important to note that the 45 disaggregated dependent variables obtained from the Stats 19 form, are essentially only related to the two year period 1969 and 1970.

The Police Stats 19 form has been operative for many years, but the data used in this research came from the 1968 revised form. This form which has to be completed by the police for every personal injury accident, contains extended information ranging from the type and class of casualties to attendant road and manoeuvre circumstances. (Fig. 6.2.1). The computerisation of these forms in 1967 by Harris (1971) at the TRRL at Crowthorne, enabled the extraction of the necessary data to take place for the relevant local authority areas. This allowed a greater disaggregating of the data than would otherwise have been possible. The definition of the various disaggregated dependent variables is given below:-

Y_{01} <u>TOTAL CASUALTIES:</u> (TC 70) the total number of fatal, serious and slight casualties reported by the police for each relevant local authority.

MINISTRY OF TRANSPORT
SCOTTISH DEVELOPMENT DEPARTMENT
WELSH OFFICE

Form Stats. 19 (Rev. 1968)

REPORT OF ROAD ACCIDENT RESULTING IN PERSONAL INJURY

POLICE REFERENCE No.

CASUALTIES

CASUALTY CLASS (1st 2nd 3rd 4th 5th 6th) 1
Driver or rider 12
Passenger 11
Pedestrian 0
Or:—
Pedal cycle 1
Moped solo 2
Motor scooter solo 3
Motor cycle solo 4
Combination (m.p., m/s, or m/c) 5
Car or taxi 6
Public service vehicle 7
Goods vehicle 8
Other vehicle 9

SEVERITY OF INJURY 2
Killed † 1
Seriously injured 2
Slightly injured 3

SEX OF CASUALTY 3-4
Male 12
Female 11

AGE in years *

CAR PASSENGER *
In front seat 12
In rear seat 11

PEDESTRIAN ACTION 5
Crossing road at pedestrian crossing 12
Crossing road within 50 yds. of ped. crossing 11
Crossing road elsewhere 0
In road, not crossing 1
Masked by stationary vehicle 2
On footpath or verge 3
On refuge or centre strip 4

PSV PASSENGER
Boarding or alighting 5

CAR OR VAN OCCUPANT *
Safety belt in use 1
Safety belt fitted but not in use 2
Safety belt not fitted 3

VEHICLE OCCUPIED 6
First vehicle 1
Second vehicle 2
Third vehicle 3
Fourth vehicle 4
Fifth or subsequent vehicle 5

VEHICLES INVOLVED

REGISTRATION NUMBERS
1st .... 2nd ....
3rd .... 4th ....
(1st 2nd 3rd 4th: 7 8 9 10 (7) (7) (7))

TYPE 11-12

MAKE 13

MODEL OR SIZE

SKIDDING 14
Skidded 12
Jackknifed 11

TOWING & ARTICULATION
Articulated vehicle 1
Caravan 2
Other tow 3

VEHICLE DEFECTS 15, 16, 17 (12—0, 1—6)
Load 12
Tyres—report for process 11
Brakes 0
Lights, front 1
Lights, rear 2
Brake lights 3
Trafficators 4
Vision—glass wipers, washers 5
V.E. report requested 6

AGE (entered by Ministry) (1st 2nd 3rd 4th) 20

MANOEUVRES 21
Entering or leaving layby 12
Entering or leaving hard shoulder 11
Reversing 0
Parked 1
Stopping 2
Starting 3
Turning round 4
Turning left or waiting to 5
Turning right or waiting to 6
Going ahead but held up 7
Going ahead overtaking moving or held up vehicle 8
Going ahead, other 9

OVERTURNING 22
Overturned 11

PART DAMAGED 23, 24
Front 1
Back 2
Left 3
Right 4
All four sides 5
None 6

SEATS OCCUPIED (CARS ONLY) * 25
Front Passenger 11
Rear: One passenger 1
Rear: Two passengers 2
Rear: 3 or more passengers 3

DRIVERS AND RIDERS

Vehicle: 1st 2nd 3rd 4th 18

SEX
Male 12
Female 11

BREATH TEST
Positive 1
Negative 2
Not required 3
Failed to provide 4

LEARNER-DRIVER 19
"L" driver 11

DRIVER'S ACTIONS
Disobeyed junction control 1
Disobeyed double centre or offset line 2
Failed to give precedence at pedestrian crossing 3

DRIVER'S OR PASSENGER'S ACTION
Opened door negligently 4

AGE GROUP (estimated) Vehicle: 1st 2nd 3rd 4th 26
Under 17 years 12
17 years 11
18 years 0
19 years 1
20 years 2
21 years 3
22—24 4
25—28 5
29—34 6
35—54 7
55—64 8
65 or over 9

*Report these items until 31.12.70, at which date a review becomes operative.

ATTENDANT CIRCUMSTANCES

POLICE FORCE 27 28

LOCAL AUTHORITY 29 30 31

DAY OF WEEK

DATE 32 33 34 (Day Mth)

TIME (24-hour clock) 35 36 37 (Hour 5Min.)

ROAD CLASS 38
M 1
A, motorway standard 2
A 3
B 4
C or unclassified 5

ROAD NUMBER 39 40 41 42

SPEED LIMIT 43
30 mph or less 3
40 mph 4
50 mph 5
60 mph 6
70 mph 7

ROAD TYPE 44
Clearway in force 11
One-way street 1
Dual carriageway 2
Other road 3

LANE MARKINGS 45
No lane markings 1
Two lanes 2
Three lanes with offset double lines 3
Normal three lanes 4
Four lanes 5
Lane markings unknown 6

LIGHT 46
Daylight 1
Dark, with street lights over 20 feet high, lit 2
Dark, with street lights under 20 feet hight, lit 3
Dark, street lights unlit 4
Dark, no street lighting 5

WEATHER 47
Raining 1
Snowing 2
Fog 3
Other 4

ROAD SURFACE CONDITION 48
Dry 1
Wet 2
Snow or ice 3

JUNCTION DETAIL 49
Roundabout 1
'T' or staggered junction 2
'Y' junction 3
Crossroads 4
Multiple junction 5
Other junction 6
Using private drive or entrance 7
Not at or within 20 yards of junction 8

JUNCTION CONTROL 50
Authorized person 1
Automatic traffic signal 2
Stop sign 3
'Give Way' sign or marking 4
Uncontrolled 5

PEDESTRIAN CROSSING 51
(If on or within 50 yds.)
Manually controlled 1
Light-controlled at junction 2
Light-controlled not at junction 3
Uncontrolled 4

MOVEMENTS BEFORE ACCIDENT 52
Stationary vehicle involved 11
One moving vehicle 1
Two moving vehicles, same direction 2
Two moving vehicles, opposite directions 3
Two moving vehicles, different roads 4
More than two moving, all same direction 5
More than two moving, same road, not all same direction 6
More than two moving, not all same road 7

CARD CLASS (entered by Ministry) 53

SPECIAL CONDITIONS AT SITE 54 55 56
Dog in carriageway 12
Other animal in carriageway 11
Object in carriageway 0
Parked vehicle(s) (contributory) 1
Automatic level crossing involved 2
Lamp post hit 3
Telegraph pole hit 4

CLASS OF ACCIDENT 57
Fatal 1
Serious 2
Slight 3

LOCATION 58 59 60 61 62 63

† If a person under 15 is killed, attach or write below a brief account of the accident.

FIG. 6.2.1. THE POLICE STATS. 19. FORM.

Y_{02} TOTAL ACCIDENTS: (TA 70):-the total number of accidents involving any personal injury reported by the police for each relevant local authority.

Y_{03} TOTAL FATAL AND SERIOUS CASUALTIES: (TFS 70) the total number of fatal and serious casualties reported by the police for each relevant local authority.

Y_{04} TOTAL PEDESTRIAN CASUALTIES: (TPED 70) the total number of pedestrian casualties, fatal, serious and slight reported by the police for each relevant local authority.

Y_{05} TOTAL MOTOR CYCLE ACCIDENTS: (TMC 70) the total number of accidents involving a motor cycle, motor scooter or moped, reported by the police for each relevant local authority.

Y_{06} TOTAL PEDAL CYCLE ACCIDENTS: (TPCY 70) the total number of accidents involving a pedal cyclist reported by the police for each relevant local authority.

Y_{07} TOTAL DRIVER AND RIDER CASUALTIES: (TDRC 70) the total number of drivers and riders injured and reported by the police for each relevant local authority.

Y_{08} TOTAL CHILD PEDESTRIAN CASUALTIES: (TCPED 70) the total number of pedestrian casualties below the age of 15 years reported by the police for each relevant local authority.

Y_{09} TOTAL YOUNG DRIVER AND RIDER CASUALTIES (TYDR 70) the total number of driver and rider casualties in the age group 16 - 25 years reported by the police for each relevant local authority.

Y_{10} <u>TOTAL JUNCTION ACCIDENTS:</u> (TJUN 70) the total number of accidents occuring at or within 20 yards of a junction involving personal injury and reported by the police for each relevant local authority.

Y_{11} <u>TOTAL ACCIDENTS AT PEDESTRIAN CROSSINGS:</u> (TPDCR 70) the total number of accidents occurring at or within 50 yards of a crossing involving personal injury and reported by the police for each relevant local authority.

Y_{12} <u>TOTAL ACCIDENTS INVOLVING A TURNING MOVEMENT:</u> (TTRN 70) the total number of injury accidents involving a vehicle turning round, turning left or waiting to, or turning right or waiting to do so, reported by the police for each relevant local authority.

Y_{13} <u>TOTAL ROUNDABOUT ACCIDENTS:</u> (TRT 70) the total number of injury accidents occurring at, or within 20 yards of a roundabout reported by the police for each relevant local authority.

Y_{14} <u>TOTAL ACCIDENTS AT 'T' JUNCTIONS:</u> (TTJN 70) the total number of injury accidents occurring at or within 20 yards of a 'T' junction or staggered junction, reported by the police for each relevant local authority.

Y_{15} <u>TOTAL ACCIDENTS AT 'Y' JUNCTIONS:</u> (TYJN 70) the total number of injury accidents occurring at or within 20 yards of a 'Y' junction, reported by the police for each relevant local authority.

Y_{16} TOTAL ACCIDENTS AT CROSS-ROADS: (TXJN 70) the total number of injury accidents occurring at or within 20 yards of a cross-roads junction, reported by the police for each relevant local authority.

Y_{17} TOTAL ACCIDENTS ON UNCLASSIFIED ROADS: (TUCRD 70) the total number of personal injury accidents occurring on urban roads graded as class C roads or unclassified roads, reported by the police for each relevant local authority.

Y_{18} TOTAL ACCIDENTS ON CLASSIFIED ROADS: (TCRD 70) the total number of personal injury accidents occurring on urban roads graded as class M, A/M, A and B, reported by the police for each relevant local authority.

Y_{19} TOTAL CASUALTIES PER THOUSAND OF POPULATION: (TCPP 70) the total number of road casualties reported by the police for each relevant local authority, divided by the population (000s) for that local authority area.

Y_{20} TOTAL ACCIDENTS PER THOUSAND OF POPULATION: (TAPP 70) the total number of personal injury accidents, reported by the police for each relevant local authority, divided by the population (000s) for that local authority area.

Y_{21} TOTAL FATAL AND SERIOUS CASUALTIES PER THOUSAND OF POPULATION: (TFSPP 70) the total number of fatal and serious casualties, reported by the police for each relevant local authority, divided by the population (000s) for that local authority area.

Y_{22} CASUALTY RATE PER MILLION VEHICLE MILES: (TCVM 70) the total number of casualties reported by the police for each relevant local authority, divided by the number of million vehicle miles calculated for that local authority area.[1]

Y_{23} ACCIDENT RATE PER MILLION VEHICLE MILES: (TAVM 70) the total number of personal injury accidents reported by the police for each relevant local authority, divided by the number of million vehicle miles calculated for that local authority area.

Y_{24} FATAL AND SERIOUS CASUALTY RATE PER MILLION VEHICLE MILES: (TFSVM 70) the total number of fatal and serious casualties reported by the police for each relevant local authority, divided by the number of million vehicle miles calculated for that local authority area.

Y_{25} SEVERITY RATE: (SR 70) the total number of fatal and serious casualties reported by the police for each relevant local authority, divided by the total number of casualties reported by the police for each relevant local authority.

---

1 Vehicle mileage is defined as:-

$$Vm_i = TTF_i \cdot RM_i$$

where TTF_i = Total traffic flow for $area_i$ (see notes on independent variables)

RM_i = Total road mileage for $area_i$.

Y_{26} <u>CHILD PEDESTRIAN CASUALTY RATE:</u> (CPEDCR 70) the total number of pedestrian casualties below the age of 15 years reported by the police for each relevant local authority, divided by the total number of pedestrian casualties reported by the police for each relevant local authority.

Y_{27} <u>YOUNG DRIVER AND RIDER CASUALTY RATE:</u> (YDRCR 70) the total number of drivers and rider casualties in the age group 16 - 25 years reported by the police for each relevant local authority, divided by the total number of driver and rider casualties reported by the police for each relevant local authority.

Y_{28} <u>TURNING ACCIDENTRATE:</u> (TUNAR 70) the total number of personal injury accidents resulting from a turning manoeuvre[2], reported by the police for each relevant local authority, divided by the total number of personal injury accidents reported by the police for each relevant local authority.

Y_{29} <u>PEDESTRIAN CROSSING ACCIDENT RATE:</u> (PDCRAR 70) the total number of personal injury accidents occurring at or within 20 yards of a pedestrian crossing reported by the police for each relevant local authority, divided by the total number of personal injury accidents reported by the police for each relevent local authority.

---

2 Actions included in a turning manoeuvre are:-
1) Turning round in the roadway.
2) Turning left or waiting to turn.
3) Turning right or waiting to turn.

Y_{30} ACCIDENTS PER MILE OF UNCLASSIFIED ROAD: (TAMUC 70) the total number of personal injury accidents occurring on unclassified roads reported by the police for each relevant local authority, divided by the mileage of unclassified roads[3] in that local authority area.

Y_{31} ACCIDENTS PER MILE OF CLASSIFIED ROADS: (TAMC 70) the total number of personal injury accidents occurring on classified roads reported by the police for each relevant local authority, divided by the mileage of classified roads[3] in that local authority area.

Y_{32} ROUNDABOUT ACCIDENT RATE: (RTAR 70) the total number of personal injury accidents occurring at or within 20 yards of a roundabout reported by the police for each relevant local authority divided by the total number of personal injury junction accidents reported by the police for each relevant local authority.

Y_{33} 'T' JUNCTION ACCIDENT RATE: (TJNAR 70) the total number of personal injury accidents occurring at or within 20 yards of a 'T' or staggered junction reported by the police for each relevant local authority, divided by the total number of personal injury accidents occurring at or within 20 yards of a junction, reported by the police for each relevant local authority.

---

3 Supplied by the Ministry of the Environment and the Scottish and Welsh Offices.

Y_{34} 'Y' JUNCTION ACCIDENT RATE: (YJNAR 70) the total number of personal injury accidents occurring at or within 20 yards of a 'Y' junction, reported by the police for each relevant local authority, divided by the total number of personal injury accidents occurring at or within 20 yards of a junction reported by the police for each relevant local authority.

Y_{35} CROSS-ROAD JUNCTION ACCIDENT RATE: (XJNAR 70) the total number of personal injury accidents occurring at or within 20 yards of a cross-road junction, reported by the police for each relevant local authority, divided by the total number of personal injury accidents occurring at or within 20 yards of a junction, reported by the police for each relevant local authority.

Y_{36} TOTAL PEDESTRIAN CASUALTIES PER THOUSAND OF POPULATION: (TPEDPP 70) the total number of pedestrian casualties reported by the police for each relevant local authority, divided by the population (000s) of that local authority area.

Y_{37} PEDESTRIAN CASUALTY RATE: (PEDCR 70) the total number of pedestrian casualties reported by the police for each relevant local authority, divided by the total number of casualties, reported by the police for each relevant local authority.

Y_{38} TOTAL PEDAL CYCLE ACCIDENTS PER THOUSAND PEDAL CYCLE WORK TRIPS: (PCYWT 70) the total number of pedal cycle injury accidents reported by the police for each relevant local authority, divided by the estimated pedal cycle work trips[4] for that local authority area.

Y_{39} TOTAL PEDESTRIAN CASUALTIES PER HUNDRED WORK WALK TRIPS: (PEDWT 70) the total number of pedestrian casualties reported by the police for each relevant local authority, divided by the estimated work walk trips[4] for that local authority area.

Y_{40} PEDAL CYCLE ACCIDENT RATE: (PCYAR 70) the total number of personal injury accidents involving pedal cycles, reported by the police for each relevant local authority, divided by the total number of personal injury accidents reported by the police for each relevant local authority.

Y_{41} MOTOR CYCLE ACCIDENT RATE: (MCAR 70) the total number of personal injury accidents involving a motor cycle reported by the police for each relevant local authority, divided by the total number of personal injury accidents reported by the police for each relevant local authority.

Y_{42} TOTAL DRIVER AND RIDER CASUALTIES PER MILLION VEHICLE MILES: (TDRVM 70) the total number of driver and rider casualties reported by the police for each relevant local authority, divided by the number of million vehicle miles calculated for that local authority area.

---

4 Both the pedal cycle work trips (Y_{38}) and the total walk work trips (Y_{39}) were obtained from the "Journey to Work" tables presented in the 1966 census returns. These estimates were included to try and obtain some degree of exposure of these two disaggregated accident variables.

Y_{43} DRIVER AND RIDER CASUALTY RATE: (DRCR 70) the total number of driver and rider casualties reported by the police for each relevant local authority, divided by the total number of casualties reported by the police for each relevent local authority.

Y_{44} CASUALTY RATE: (TCTA 70) the total number of casualties reported by the police for each relevant local authority, divided by the total number of personal injury accidents reported by the police for each relevant local authority.

Y_{45} JUNCTION ACCIDENT RATE: (TJNAR 70) the total number of personal injury accidents occurring at or within 20 yards of a junction reported by the police for each relevant local authority, divided by the total number of personal injury accidents reported by the police for each relevant local authority.

As will be obvious from the list of dependent variables used, their nature and type are very diverse, some are absolute figures, whilst others are compound variables. Within the road accident literature, there has been some, but not enough discussion, as to what type of variable is the best kind of safety indicator. (Katz 1971 Rapoport 1967). Following Katz it can be argued that there are two groups of indicators, one which measures how intense or prevalent accidents are within a given population, and another which measures the amount of accident incidence in relation to a related measure of exposure to that activity. Thus the first group which Katz calls the "static" or "population" family includes such indicators as:-

(i) accidents per 1,000,000 persons

(ii) accidents per 1,000 vehicles

(iii) accidents per mile of road

(iv) accidents per 1000 male drivers etc.

The second group meanwhile, which can be called the "exposure" or "activity" family of indicators, includes:-

(i) accidents per mile of bus travel

(ii) accidents per mile of annual driving

(iii) accidents per number of pedestrians crossing the road

(iv) accidents per mile of travel of road length etc.

The distinction between these two groups is both relevant to this research, and also any future research, since of the two, the exposure indicators are by far the more meaningful ones yet to date such information is either scarce or non-existent, and when applied to sub-national geographical areas, open to considerable error values due to the dearth of accepted information collection and derivation. Therefore of the 45 dependent variables used in this research, only 6 can be said to be of the exposure group whilst the remainder belong to the static family of indicators. The reason for preferring the exposure indicators is due to the limitation inherent within the static indicators, which reflect both the amount of activity and the degree of danger associated with that activity, within one statistic. Thus where one is trying to measure the effectiveness of some countermeasure action, such indicators would appear meaningless. However, where one is trying to study spatial variations about some declared "norm" of acceptability, as in this research, whilst the static indicators may not be the optimum ones, they do have considerable relevence and are therefore meaningful in relation to consequent conclusions.

6.3 INDEPENDENT VARIABLES

Unlike the dependent variables, there was no one source of data for the independent variables, although in several instances the 1961, 1966 and 1971 censuses provided the major components. Of the 22 independent variables 8 were constant during the period of the study, whilst the remainder varied from year to year. A description and the derivation of each of these variables is given in the following section.

X_1 POPULATION:- (POP 70) The incidence of road accident conflict is obviously related to the size of an urban area. Therefore data of the relevent population levels were obtained from the 1961, 1966 and 1971 census returns for England, Wales and Scotland. The data for the intervening years was taken from the Registrar General's annual estimates of population.

X_2 AREA OF AN URBAN AREA:- (AR 70) Once again this information was obtained from the respective census returns. Additional information was also obtained from the data given within the Municipal Year Book, and the annual reports of the New Town corporations, where relevant. The area of an authority was measured in acres.

X_3 POPULATION DENSITY:- (DEN 70) This variable was a combination of the previous two variables and can be defined as

$$D_{it} = \frac{P_{i.t}}{A_{i.t}}$$

where $D_{i.t}$ = the density of population area i at time t (persons per acre)

$P_{i.t}$ = the population of area i at time t

$A_{i.t}$ = the area of an area i at time t (acres)

X_4 PEAK TRAFFIC FLOW:- (PTF 70) During the period 1964 - 1966 the Road Research Laboratory made a study of road traffic flows crossing cordons around the control areas of eight towns of varying population size. One of their conclusions was; "The peak $\frac{3}{4}$ hour flow a.m. or p.m. in passenger car units was found to vary with population p as-

$$f.p.c.u = 6p^{0.6} \text{ p.c.u/h}$$ (Munt et al 1968)"

Accordingly this formula was used in this research as an estimator for the peak traffic flow, for each of the relevant urban areas, and for each year under study.

X_5 TOTAL TRAFFIC FLOW:- (TTF 70) Because of the possibility of error in the previous variable it was deemed expedient to introduce a second traffic variable which could be calculated from an entirely different source. Although some data on traffic flow was available from the numerous Traffic and Transportation studies already published, due to differing methods of collection and estimation, comparisons between different areas was deemed unacceptable. Similarly whilst at one stage it was originally intended to relate total traffic flow to a calculated total capacity for that urban area following Tanner, (Tanner 1966) this was also eventually rejected. This line of reasoning was due to the fact that any such capacity level would necessitate a high degree of calculated estimation, and that this would create a high level of inaccuracy within any such derived index. Therefore the eventual index that was derived was both simple and nationally comparable, but even so it must include a substantial error term due to the method of estimation. However, it was assumed that this error term would vary only slightly from one area to another, thereby enabling some confidence in the areal variations.

The data source used was therefore the 1966 Sample Census, "Workplace and Transport Tables," which gives the number of "person-work-trips" by mode of transport, and place of destination. Three categories of destination are given:-

(i) Persons resident and working in the area,

(ii) Persons working in the area but resident elsewhere,

(iii) Persons resident in the area but working elsewhere.

The first two groups indicate person movement into the city area, whilst the third gives movement out of the city area: Thus during the work period, all three groups are going to be involved in some kind of "work-transport-journey" within the total urban area.

The mode of transport is also givenin six categories:-

(i) Bus
(ii) Car
(iii) Goods
(iv) Motor Cycle
(v) Pedal Cycle
(vi) On foot.

The first five categories will give some indication of the number of vehicles upon the road network during the "work-travel" period, whilst the last category, if included, gives some indication of the number of pedestrians around the road network during the same time period. To relate the data presented in these returns to some kind of traffic volume index, certain extra assumptions had to be made. In the first instance the figure giving the number of people going to work by car does not give the number of those who are car passengers. If the number of cars being used is to be calculated, an assumption concerning car occupancy rates has to be made. Similarly the number of buses is dependent upon an expected bus occupancy rate. Tanner (Tanner 1966) whilst estimating the total capacity of urban areas used the two occupancy rates of 1.45 people per car and 42.5 people per bus. These rates were based on 1966 empirical observations in London and other urban areas. It was noted however that the two rates were altering, the former increasing and the latter decreasing. Therefore for the purpose of this research the two occupancy rates used were 1.5, and 40.0 respectively. Using these rates it was then possible to calculate the total number of vehicles used in the journey to work, and also the traffic flow in terms of passenger car units. The p.c.u. values used were those recommended by the former Ministry of Transport for urban areas:-

(i) Bus = 3.0 p.c.u.

(ii) Car = 1.0 p.c.u.

(iii) Motor Cycle = 0.75 p.c.u.

(iv) Pedal Cycle = 0.33 p.c.u.

The calculated p.c.u. figure is therefore an estimated figure of the number of vehicles used in the journey to work. By doubling this figure a two-way work journey figure can be assumed. From empirical observations recorded in the various Land Use and Transportation studies available, it can be seen that the volume of work journeys ranges between 28 - 35% approximately, of all journeys, during the 24 hour day. (6.3.1). If the previously calculated figure is multiplied by a factor of 3, one then has some estimate of the Annual Average Daily Flow. This calculated Annual Average Daily Flow (AADF) was the figure used for the Total Traffic Flow variable for 1966. Taking this figure as the base figure, variations from year to year were estimated by using

SURVEY	YEAR	WORK TRIPS AS % ALL TRIPS	CAR WORK TRIPS AS % ALL CAR TRIPS
LEICESTER	1963	40.51	35.03
W. MIDLANDS	1964	54.40	51.67
M.A.L.T.S.	1966	39.70	29.70
S.E.L.N.E.C.	1966	43.60	27.60
EXETER	1965	30.24	31.57
BELFAST	1968	18.00	21.00
L.T.S.	1964	24.00	34.00
NORWICH	1967	32.00	43.00
E. CEN. SCOT.	1966	38.00	32.00
CAMBRIDGE	1969	26.00	32.00
WINDSCR	1968	29.00	N.A.

TABLE 6.3.1. VOLUME OF WORK JOURNEYS AS A PERCENTAGE OF TOTAL TRIPS AS GIVEN BY VARIOUS RECENT TRANSPORTATION SURVEYS.

the annual indices of vehicle travel for all urban roads, (Dunn 1970) attained from the various traffic census returns of the Ministry of Transport. These annual correction factors are given in (Table 6.3.2). The total traffic index can therefore be defined as below:-

$$TTF_{it} = 6\left\{\left(\sum_{j=1}^{3} B_{ij}/0.075\right) + \left(\sum_{j=1}^{3} C_{ij}/1.5\right) + \left(\sum_{j=1}^{3} M_{ij}/0.75\right) + \left(\sum_{j=1}^{3} P_{ij}/0.33\right)\right\} C_t$$

where:-

TTF_{it} = Total Traffic Flow in area i for year t (p.c.u.)

B_{ij} = The number of people going to work by bus in category of destination j, in area i.

C_{ij} = The number of people going to work by car or goods in category of destination j, in area i.

M_{ij} = The number of people going to work by motor cycle in category of destination j, in area i.

P_{ij} = The number of people going to work by pedal cycle in category of destination j, in area i.

C_t = Correction factor for year t

X_6 ESTIMATE OF THE LEVEL OF THROUGH TRAFFIC:- (ETT 70) Since it was only possible to obtain "rough" statistics of the volume of through traffic for some of the larger urban areas from individual

YEAR	TRUNK ROADS (ALL)[1]		ALL ROADS (URBAN)[2]	
	TRAVEL INDEX	CORRECTION FACTOR (1966)[3]	TRAVEL INDEX	CORRECTION FACTOR (1966)[3]
1960	100	0.62	100	0.66
1961	109	0.68	108	0.72
1962	117	0.73	114	0.76
1963	125	0.78	122	0.81
1964	140	0.87	135	0.89
1965	153	0.95	143	0.95
1966	161	0	151	0
1967	169	1.05	155	1.03
1968	172	1.07	157	1.04
1969	172	1.07	161	1.07
1970	178	1.11	166	

[1] Used for Through Traffic growth.

[2] Used for internal Traffic growth.

[3] 1966 is the base year for the correction factor.

TABLE 6.3.2. TRAFFIC (GROWTH) VOLUME CORRECTION FACTORS 1960-1970 USING 1966 AS THE CORRECTION BASE YEAR.

traffic surveys some other estimate had to be derived to validate this variable. It was assumed that if traffic was to be classed as through traffic, then it would be proportional to the amount of traffic entering an urban area on the main arterial roads of the national transport network. Thus using data published by the Ministry of Transport (HMSO 1967), the volume of traffic along each of the major roads, around each of the relevant urban areas, was calculated, and the resultant figure was deemed to be the volume of through traffic 1966. The correction factor used for other years was once again taken from the annual indices of motor vehicle travel on all trunk roads. (Dunn 1970). Corroborating evidence was also obtained from the Atlas of Great Britain and N. Ireland (1963) and further census returns of the Ministry of Transport.

X_7 <u>DEGREE OF CAR OWNERSHIP:-</u> (COW 70) The original source of data for this variable was once again the 1966 Census County Reports, which provide information on "Household by number of cars," for each relevant urban area. Although it would have been desirable to extrapolate these figures, and update them to 1970 because of the lack of data for non-county borough areas, this proved impossible. It could have been possible to extrapolate for county borough areas using the estimates given by Herrmann (Herrmann 1968). However, in the first instance, these estimates were based on different indices (vehicles per head of population), and secondly even the author puts little reliability on the extrapolated figures due to the lack of local knowledge. Thus the author states; "Therefore it must be stated that the forecasts given here are not given in the belief that the figures are a close approximation to what is likely to happen. Their purpose is rather to indicate the local implications that may follow from the national forecasts that are currently accepted," and "It is suggested therefore that the forecasts given here for any particular area should only be given any acceptance if no forecasts have been made for that area on the basis of better local knowledge." (Herrmann 1968). Because of these limitations it was decided that this variable would be held constant during the period of study, solely using the 1966 observations.

X_8 MEAN SOCIAL INDEX:- (MSI 70) This index was calculated in the same manner as that used in the recent Birmingham Road Accident Survey. (McKay et. al. 1969). The index is based on the social class totals given in the County Reports of the 1966 census returns. These returns list five groups of social class, based on Socio-economic groupings, and gives the number of people in each group. The variable can be defined therefore as below:-

$$MSI_i = \frac{\sum_{j=1}^{5} (K_{ij} \cdot N_{ij})}{\sum_{j=1}^{5} N_{ij}}$$

where,

MSI_i = the mean social index for area i

K_{ij} = the number of the j th social class in area i.

N_{ij} = the number of people in the j th social class in area i

Obviously the index will have a range between the values 1 and 5, but as was found in the Birmingham survey, the index tends to oscillate between 2.20 and 4.50, which proved an even greater range than in this research because of the larger areal units used. The index is such that as the value increases, the lower the social grouping of the area. Once again because of the lack of data, and also because of the very small variance in this index, this variable was held constant during the period of study.

X_9 <u>SEX RATIO INDEX:</u> (SRI 70) This variable which measures the relative distribution of males and females within a population is important in the context of road accident analysis for two reasons. In the first instance the two groups have totally different response rates to the road accident situation (Mckay 1969), whilst secondly, the two groups exist within different exposure "sets" during the working day. Data for this variable was obtained from the census returns since 1961, with interpretation between these standard years. The variable can simply be defined as:-

$$SRI_i = \frac{M_i}{F_i} + 1$$

where,

SRI_i = the sex ratio index for area i.

M_i = the number of males in area i.

F_i = the number of females in area i.

Thus as the value of the sex ratio index increases the proportion of males relative to females also increases.

X_{10} TOWN CLASSIFICATION INDEX:- (TCI 70) This index was designed such that it gave some indication of the structural shape of the urban area, both in respect to the actual shape and the developed road network system. Although these two sections are treated as one within the final index to facilitate ease of explanation, the two sections will be dealt with independently in this account.

Geographical literature has often expounded the problems of finding a satisfactory mathematical representation of the shape of any one area, (Bunge 1966, Haggett 1965 etc.) yet even with these problems, several such indices have been developed and used. In this town classification index however, it was thought that no great advantage would be gained by using one of the more intricate techniques, and a simple measure of shape by Haggett (Haggett 1965) was eventually employed. Haggett's shape index can be represented as:-

$$Sh_i = (1.27A_i) / l_i^2$$

where,

Sh_i = the shape index for area i.

A_i = the area in sq. kms. of area i.

l_i = the length of the long axis (km) for area i.

This index gives a range of values between zero and one. Values near zero indicate extreme elongation, whilst those near one tend to circularity. Intermediate values, calculated by Haggett are; triangles = 0.42, squares = 0.64 and hexagons = 0.83. (Table 6.3.3).

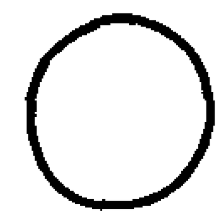
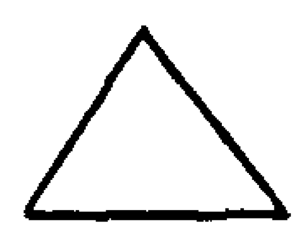
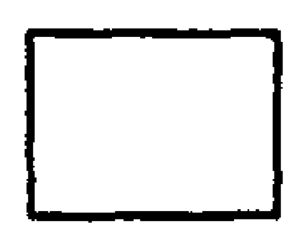
SHAPE	HAGGETTS INDEX	DAG. REP.	BLAIR/BLISS INDEX
CIRCLE	1.000		1.000
POLYGON (10 SIDED)	0.935		0.999
HEXAGON (6 SIDED)	0.827		0.996
EQUILATERAL TRIANGLE (3 SIDED)	0.413		0.909
SQUARE (4 SIDED)	0.637		0.977
ELONGATION (1 x 7.5 UNITS)	0.167	 7.5	0.500

TABLE 6.3.3. COMPARISON OF SOME STANDARED SHAPES USING THE HAGGETT SHAPE INDEX AND THE BLAIR-BLISS INDEX.

The calculation of the network spacing was based on work by Dacey, (Dacey 1967), who used a random walk technique in testing the network spacing of different river drainage patterns. The random walk is constructed by locating an originating point from a pair of random co-ordinates. From this point a vector is constructed by selecting a real number from the range zero to two. The vector terminates where it intersects the most distant line of the network system. From this point a second randomly-orientated vector originates; it terminates once again at the point at which it intersects the most distant line under investigation. The walk continues until the number of intersections is equal to the sample size desired, which in this research was deemed equal to 60 intersections. Earlier work by Clark (Clark 1956), has shown that in a random spacing of points along a long traverse line $\frac{2}{3}$rds of the points have a reflexive (first) nearest neighbour, and the proportion of points which have a reflexive nearest neighbour of order "n" = $(\frac{2}{3})^n$. When the percentage of points having n^{th} order nearest neighbours is significantly greater than $(\frac{2}{3})^n$ the lines are positively grouped as in the Manhattan Grid. Conversely when the proportion of points having n^{th} order nearest neighbours is significantly less than $(\frac{2}{3})^n$ the lines are negatively grouped as in the Radial Grid. This technique of studying network spacing produces a range of zero to one, if the proportion of reflexive points to the total number of points is used.

Combining these two previously mentioned indices, the Town Classification Index can now be formulated based both on shape and network density. This index can be defined as below:-

$$TCI_i = \left(\frac{1.27A_i}{\ell_i^2}\right) \cdot \left(1 - \frac{r_i}{N}\right)$$

where TCI_i = the town classification index for area i.

A_i = the area in sq. kms. of area i.

ℓ_i = the length of the longest axis (kms) for area i.

r_i = the number of reflexive first-order nearest neighbour points in area i.

N = total number of sampled intersections on the random walk traverse (60)

Thus the index will have a range of zero to one with a circular town with a radial network having a value approaching one, and an elongated town with a Manhattan grid having a value approacing zero. Besides allowing the calculation of the town classification index, the above mentioned approach and results gave some interesting insights into the structure of the sampled urban areas. This point is developed more in Appendix (1). Data for the calculation of the town classification index was obtained from the Ordinance Survey 6" maps of Great Britain.

X_{11} COMPACTNESS INDEX: (CI 70) Unlike the Town Classification index just described, the compactness index purely related to the shape and circularity of the various urban areas, and no account was taken of the internal road network system. The index used in this research was the "Blair-Bliss" compactness index (Blair et al. 1967) which is defined by the formula:-

$$CI_i = \frac{A_i}{\sqrt{2\pi \int_a R^2 \, dx \, dy}}$$

where,

CI_i = the compactness index for area i.

A_i = area of urban area i (sq. kms.)

a R = radial axes from gravity centre to small area, 2 (km.)

The index values vary from between zero and one, with values approaching one indicating a high degree of compactness and circularity. Because the Blair-Bliss index is more related to areal spread around the centroid of an urban area, the range of values experienced with this index, as compared with those obtained from Haggett's shape index, is substantially reduced. For example, the values calculated in this research for each of the administrative units had a range of only 0.339, varying from 0.645 to 0.984.

X_{12} DEGREE OF OPEN SPACE:- (DOS 70) Because of the lack of data this was another variable which was held constant during the period of study. The original data was obtained from the 1967 Municipal Year Book which published returns from a questionnaire sent to all local authorities in 1964 - 65. The returns gave information as to the total acreage of public open space, and also the proportion of total acreage, per 1000 population. It was this last figure which was used as the degree of open space variable in this research.

X_{13} STANDARD OF HOUSING INDEX:- (SHI 70) This variable attempts to provide information as to the relative standard of the housing stock in different urban areas. The source of the data was the census returns since 1961, which provide tables for all local authorities, of the number of dwellings which do not provide the basic amenities of hot and cold water, inside W.C. and a bath etc. It was therefore possible to ascertain two sets of figures from these census returns; the number of dwellings with these basic necessities, and those without. The index derived therefore can be defined as:-

$$SH_i = B_i/C_i$$

where

SHI_i = Standard of Housing Index for area i

B_i = The number of dwellings with basic necessities in area i.

C_i = The number of dwellings without basic necessities in area i.

Obviously therefore the higher the value of this index, the better the standard of housing in that area.

X_{14} AGE VULNERABILITY INDEX:- (AVI 70) The two most vulnerable sections of the population as far as pedestrian accidents are concerned, involve the two age groups, 0 - 14 years and over 65 years. It is therefore important to know the potential vulnerability of an area, as regards these two groups. Data on the actual numbers within these two age groups was obtained from the Census returns, county reports. These numbers were then expressed in relation to the total population of that area. Thus,

$$AVI_i = A_i/P_i$$

where,

AVI_i = the age vulnerability index of area i.

A_i = the number of people under 14 years of age and over 65 years of age in area i.

P_i = the total population of area i.

The higher the index the more vulnerable the area is, especially in relation to pedestrian accidents.

X_{15} DRIVER VULNERABILITY INDEX:- (DVI 70) This variable was included to supplement the previous age vulnerability index. It is generally assumed that the most vulnerable age group for drivers occurs within 5 to 10 years of learning to drive. In most instances this involves people in the age group 16 years to 25 years. However because of the limitations of the available data in the relevant Census returns, the age group used for this variable ranged between 15 years and 25 years, for each relevant area. As before the numbers in this age group were related to the total population of the area such that,

$$DVI_i = D_i/P_i$$

where,

DVI_i = the driver vulnerability of area i.

D_i = the number of people between the ages of 15 years and 25 years in area i.

P_i = the total population of area i.

The higher the index, the more vulnerable is any area, especially in relation to the number of driver accidents.

X_{16} <u>TOTAL DWELLING INDEX:-</u> (TDI 70) Data on the total number of dwellings within any urban area was again obtained from the relevant Census returns, and also the Housing returns for England and Wales and also those for Scotland. Such a variable although obviously related to the population size of an area, was included as another indication of the magnitude of the urban area, and the resultant impact upon the road accident situation.

X_{17} <u>HOUSING TENURE INDEX:-</u> (HTI 70) It was postulated that the type of urban development between public and private might cause varying responses within the road accident system.

An indication of these two types of development can be obtained by comparison of the number of people living in local authority owned dwellings and the number of people living in private housing. Data was obtained for this index from the relevant Census returns, with additional annual information from the HMSO publications, "Local Housing Statistics for England and Wales," and "Housing returns for Scotland." The index can be defined as,

$$HTI_i = Pr_i/Pb_i$$

where,

HTI_i = the housing tenure index for area i.

Pr_i = the number of dwellings of private origin in area i.

Pb_i = the number of council owned dwellings in area i.

The lower the value of the index, the more the area has been developed by the relevant local, or public authority, and therefore the more likely the area is to have been subject to controlled comprehensive planning. The lowest index values in this research obviously occurred therefore within areas containing New Town Developments.

X_{18} <u>CHANGE OF HOUSING STOCK:-</u> (CHSI 70) Where the structure of an urban area is changing rapidly one would tend to expect a proportional response rate in observed road accidents. One way of measuring this structural change is by observing the change in the housing stock. Once again data was obtained for this index from the Census returns and also the HMSO publications of Housing Returns, and the resultant index can be defined as below,

$$CHSI_{i \cdot t} = \left(\frac{I_{i \cdot t}}{T_{i\ t-1}} \times 100 \right) + 10$$

where,

$CHSI_{i \cdot t}$ = the change of housing stock index for year t in area i.

$I_{i \cdot t}$ = the absolute increase in the housing stock during the year t, in area i.

$T_{i\ t-1}$ = the total number of dwellings in area i for the year t-1.

Thus areas with a high level of change in the housing stock would be characterised by an index value considerably above 10.

X_{19} <u>POLICE SEVERITY INDEX:-</u> (PSI 70) This index along with the following variable was introduced to study the possible impact of police action within the various urban areas. The main problem encountered however, centred around the fact that data was only available according to Police Force Areas. Similarly, the various re-organisations of the police forces since 1966 have made it very difficult to obtain comparable area statistics. The first of these problems was overcome by assuming that the influence and/or attitude of a police force would be approximately constant over all the constituent parts of that area. This is reinforced by the fact that in areas where the attitude may have been different, in areas of high population density, these areas tend to have their own police force. At the same time, to solve the problem of lack of temporal data both of these police variables were held constant over time.

The police severity index was designed to show the change in the attitude of a police force in any area from 1968 - 1970. The index can be defined as giving the proportional change in the rate of prosecutions per thousand of population 1968 - 1970. These prosecutions concern "offences relating to motor vehicles" as published by the (HMSO). The index was equal for all urban areas falling within the same police force area, and can therefore be defined such:-

$$PSI_i = (P_{ci} / P_{si \cdot 70}) \times 100$$

where,

PSI_i = the police severity index for area i.

P_{ci} = the increase in number of prosecutions per thousand population 1968 - 1970 in area i.

$P_{si \cdot 70}$ = the number of prosecutions per thousand population 1970 in area i.

X_{20} POLICE PROSECUTION INDEX: (PPI 70) Whilst the previous index attempted to show the dynamic attitude of the police forces, the prosecution index simply shows the absolute attitude of any police force within its area. The data source and application to individual urban areas, is exactly the same as in the previous variable, and can be defined simply as the number of prosecutions (relating to motoring offences) per thousand of population within the police force area in 1970:-

$$PPI_i = P_{si}/P_i$$

where,

PPI_i = the police prosecution index for area i.

P_{si} = the number of police prosecutions within area i (1970)

P_i = the population within the area i in thousands (1970).

X_{21} RAINFALL INDEX:- (RI 70) the mean annual rainfall for the years 1916 - 1950 obtained from the "Atlas of Britain and N. Ireland" (D.P. Bickmore et al 1963) was the data used for the rainfall index in this research, and obviously this variable was constant for the period studied.

X_{22} ROAD DENSITY INDEX:- (RDI 70) Defined as the mileage of roads per acre, this variable is useful on two counts. In the first instance, it helps to show the density of the road network within an urban area, and in the second instance it also gives a measure of the road junction density. This follows the work of Borchert (Borchert 1961) on the road pattern in the twin-cities of Minneapolis - St. Pauls, which showed that there was a very high degree of correlation between the density of road junctions and the total road length (+0.99). The data for this variable was obtained from Highway Statistics (HMSO) for county borough areas, and the Ministry of Transport for non-county boroughs.

This then completes the list of independent variables used in the body of this research. A summary of thm is given in (Table 6.1.1) along with some of the more relevant descriptive statistics for 1970.

Chapter 7 METHODOLOGY: I

7.1 INTRODUCTION

In order to compare the various road accident statistics spatially and temporarily, it is necessary to hold the effect of various variables constant. Without this precaution any derived index would be spatially incomparable. Since the theme of this research is to study the variations from some acceptable "norm" the index to be used is one which reflects an "expected" number of road accidents for an urban area, to those which actually occurred in that same urban area. The essential problem then was to decide upon the method by which these "expected" number of accidents, per urban area could be derived. Three such groups of methods were studied starting from a simple scaling procedure to the more complex statistical modelling of probability theory and multiple regression. Although it is often true that the simplest models are often the most useful in practical application, it must be pointed out that as far as the researcher is concerned he is more intent upon modelling as clearly as possible the real world in order to understand and explain that very world. Very often the simpler approach is pursued to the detriment of meaningful results. As regards spatial comparisons such an opinion was expressed by some of the participants at the recent O.E.C.D. Symposium held at Crowthorne, (O.E.C.D. 1969) following the session on International Comparisons. If this is true at the macro level it must also be true at the micro level. Complexity is not an end result in itself, but it is more likely to produce more realistic explanations in the long run. With this thought in mind the progress and direction of this research was not deterred by the increasingly complex techniques used, and often the simpler approach was rejected because of the low level of reality.

Before progressing on to the methodology used, it is useful to explain the other approaches considered and point out the reasons for their rejection.

7.2 "SCALING DOWN" METHOD

The simplest approach studied was that of scaling down of national data and applying these rates to individual urban areas. An example of this approach was that used by Cumbernauld Development Corporation to compare the response of Cumbernauld with other new towns, and older established towns. (Cum. Dev. Corp. 1967). Using national data on the number of accidents, it was possible to derive a "National Average rate of Accidents per thousand of Population" for each year under observation. Once this national average had been computed it was an easy step to apply it to individual urban areas, knowing their population levels. In this way an expected number of accidents could be calculated for all areas and compared with the actual number of accidents (Table 7.2.1). Although this method produced instant practical results, it has several severe limitations, which together could invalidate any conclusions derived from such a method. Firstly it would seem very naive to assume "a priori" that only one variable, (population) is responsible for all road accident variations. Similarly one could argue as to which one variable to accept. In most instances

	1966 POPULATION	TOTAL POSSIBLE ACCIDENTS DURING THE FIVE YEARS 1962/1966	TOTAL ACTUAL ACCIDENTS DURING THE FIVE YEARS 1962/1966	TOTAL ACTUA EXPRESSED A % OF TOTAL POSSIBLE ACCIDENTS
CUMBERNAULD[1]	16.640	330.73	77	23
NEWTON AYCLIFFE[1]	17,202	504.82	128	25
CUMBERNAULD[2]	19,640	249.45	91[3]	36
GLENROTHES[1]	21,830	572.73	248	43
HATFIELD	24,800	338,60	235	69
CWMBRIAN[2]	38,000	1,126.93	709	63
WELWYN GD. CITY[2]	42,600	1,314.90	1,071	81
CORBY[2]	45,065	1,397.75	948	68
STEVENAGE[1]	48,000	1,418.28	918	65
EAST KILBRIDE[1]	48,952	1,138.66	614	54
BASILDON[2]	66,700	1,619.35	1,122[6]	69
HARLOW[7]	70,800	2,160.08	1,506	77
ALLOA[8]	15,000	473.64	310	65
GRANGEMOUTH[8]	21,000	665.91	544	82

NOTES:
1 Accidents occurring on roads developed by the Corporation.
2 Accidents occurring on all roads within the designated area.
3 Covers year 1965/66 only.
5 Covers years 1964/66 only.
6 Covers years 1962/65 only.
7 Accidents occurring on all roads within Harlow Urban District
8 Existing towns not developed as New Towns.

TABLE 7.2.1. THE 'SCALING TOWN' METHOD USED FOR CUMBERNAULD NT., AND SOME OTHER COMPARABLE TOWNS.

dependent upon which variable is chosen, one could prove any hypothesis. The most obvious approach is one which applies various variables, testing each individually, whilst holding the other variables statistically constant.

The second and perhaps most damning criticism of this type of approach, involves the so called "problems of scale." Such problems have received increased attention from geographers since the work of McCarty (McCarty et al 1956) and Duncan (Duncan et al 1961). Both of these pieces of research indicated similar problems, which can be simply expounded by stating that in each case the type and direction of results was dependent upon, and varied with the size and level of the areal unit used. Thus Duncan who was studying the simple question of what is the population density of an area of downtown Chicago, came up with two sets of figures, one of which was twice as large as the other, derived from two different definitions of vicinity. Extending the problem to comparisons of different cities, Duncan also shows that the comparisons are entirely dependent upon the areal definition of these cities. This variation is shown in (Fig.7.2.2) for two cities Chicago and Detroit. Still more striking contrasts are also shown with comparisons over time where the population of the USA appears to be becoming more concentrated on the evidence of the counties, but were dispersed on the evidence of the states. (Table 7.2.3). A similar development of this scale problem has been noted by Harvey, "Similar problems of comparability and

AREAL DEFINITION	CHICAGO(C)* (POP/SQ. ML.)	DETROIT (D)* (POP/SQ. ML.)	RATIO (D/C)
CITY	17,450	13,249	0.76
URBANISED AREA	7,713	6,734	0.86
ST,METROPOLITAN AREA	1,519	1,535	1.01

SOURCE: DUNCAN, CUZZORT & DUNCAN 1961

* 1950

TABLE 7.2.3. URBAN POPULATION DENSITIES UNDER ALTERNATIVE CENSUS BOUNDARIES.

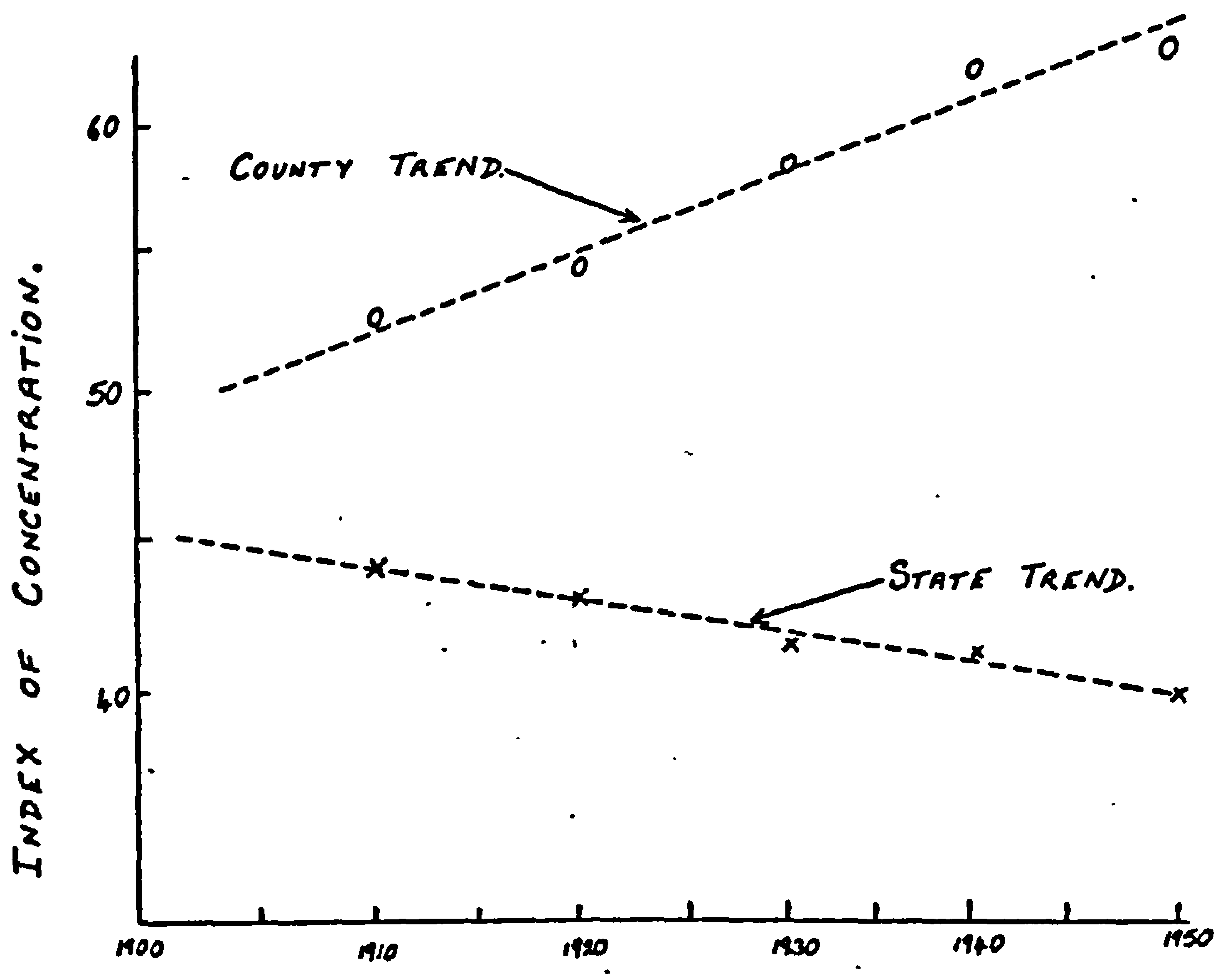


FIG. 7.2.1. CONTRASTS BASED ON DIFFERENT SIZES OF COLLECTING UNITS.

inference exist when different levels in the hierarchy are simultaneously studied..... We shall begin by considering the special case in which distinct steps in the hierarchy of areal units can be recognised, i.e. areal units at one level can be included in areal units at the next level. In such a "nested" hierarchical situation it should be observed that comparisons can only be made between similar individuals, (i.e. individuals at the same level in the hierarchy) and that inferences made about relationships at one level cannot be extended, without making strong assumptions, to any other level. This is not to say that conditions at one level are irrelevant to conditions at another. It does indicate that the nature of the analysis is contingent upon whether the individuals being compared or analysed are at the same or different levels." (Harvey 1969 p.352). Harvey then continues by pointing out that where continuous data is being used areal units can be "imposed" spatially in order to create a hierarchical situation within which comparison is possible. However, in most situations such assumptions can prove to be groundless. For example, whilst administrative units may be regarded as singular areal individuals from the point of view of administrative structure, it is unrealistic to continue comparability to other variables, especially if these are continuous variables. "In such a situation there is nothing natural in the way the data are aggregated and it may well be that the administrative units are not "similar" or comparable from the point of view of the data aggregated within them." (Harvey 1969, p.353). The main problem therefore is that since most areal data is only available for territorial units, and will probably remain so for the next few decades, (except in the case of computer grid mapping techniques) how can one attempt to overcome these areal scale problems? Theoretically there are perhaps three different types of solution, but once again they all include their own implicit limitations. Robinson suggested that one solution was to "weight" the various parameters by the size of the areal unit. (Robinson 1956). Thus he suggested modified equations for the standard deviation, and regression co-efficient as below,

$$\sigma_A = \sqrt{\left\{(\Sigma AX^2)/\Sigma A\right\} - (\Sigma AX/\Sigma A)^2} \quad \text{and}$$

$$b_A = \left\{\Sigma A_i \Sigma A_i X_i Y_i - (\Sigma A_i X_i)(\Sigma A_i Y_i)\right\} / \left\{\Sigma A_i \Sigma A_i X_i^2 - (\Sigma A_i X_i)^2\right\}$$

where A_i was the area of the ith observation.

However, such modifications will only in very special circumstances remove the size effect from areal comparison studies, as was pointed out by Thomas and Anderson (1965).

Another attempt by Robinson to overcome this same problem suggested the use of regular hexagons as the unit of study in preference to irregularly sized counties. (Robinson et al 1961). For each hexagon he interpreted a standardised value. However, as King has pointed out such a solution is neither practical nor in many cases, very useful. (King 1969). "The interpolation is not only time-consuming and arduous, but there is no guarantee that it will not introduce as much error, perhaps even more, as would result from the use of the original county data." (King 1969 p.155).

The work by Thomas and Anderson (1965) tries to solve the problem through the notion of statistical inference, making a distinction between a population as the total set of outcomes or events from which a sample may be drawn and a universe which is "a more abstract group" containing "all events as they happened and as they might have happened if everything else had remained the same but the random shocks." Thus it follows that "the data for each of .....several study areas may be treated as a random sample from some hypothetical universe of possible values although each is in fact, a population." (Thomas et al 1965). However once again this approach involves severely limiting assumptions and as Curry has pointed out, the distinction between a population and universe is very tenuous. (Curry 1966). He further concludes that "the real problems in the study of areal associations are not statistical but rather the dearth of theory on the process producing the association."

The conclusions to be drawn from the above discussion are therefore totally inconclusive, and the best that can be said at present is that given our acceptance of these areal data limitations, one can only proceed as before, but with increased caution. The one glimmer of hope is provided by Harvey who following his critique on the use of non-euclidean data collection units, comes to this tentative conclusion. "It may seem reactionary to defend the crazy patchwork of administrative units of different sizes which have traditionally been used to record data, but such units have tended to be adaptable in the face of temporal change and tend to be much smaller in areas of great social and economic activity than they are in areas of little economic activity. In some cases it may be, therefore, more appropriate to use data collected in old-style administrative units than data collected in the new-style uniform Euclidean cell." (Harvey 1969 p.355).

The above discussion on the problems of scale has been dealt with in some detail because of its importance not only as a criticism of the "national rate" method used by Cumbernauld, but also within the bounds of this present research. As regards the first point, the "national rate" method is a perfect example of the cross-hierarchical comparison that Harvey was discussing, and as such would seem to be dealing with much too many implicit assumptions, and has therefor little reliability. The present research meanwhile has attempted to explicitly avoid this problem through various means. Initially, one hieratchical level is taken, that of sub-national local authority divisions, even though it is accepted that these divisions are arbitrarily imposed, and must therefore contain the resultant restrictions mentioned previously. Similarly the difference between the units in terms of size, shape etc. have been acknowledged by introducing these terms as independent variables into the multiple regression equations. Finally, since it has been theorised that road accidents in urban areas occur within an urban and road accident system it would seem reasonable to accept these units as equally comparable units. However, even having argued these points, it must be mentioned that the limitations of the various data sources is well recognised and although some modification would have been desirable, this has proved impossible, due to the vast quantities of data utilised. Thus one must remember that for this reason any conclusions and results must be tempered with caution. A good example of this problem is the differing results obtained by this research and that mentioned above, by the Cumbernauld Development Corporation. This diversity of results can almost totally be explained by the fact that two different areal units were used. One dealt with that area under the aegis of the Development Corporation, whilst the present research has dealt with the whole Cumbernauld District Council. Once again therefore, caution must be emphasised in interpretation of the results of this research if comparison is necessary with other studies using different areas. The only solution to this problem and the problem of scale in general is further micro-level analysis, say at the individual urban area level, but before this can be contemplated the further necessary theoretical, and classifactory research must be completed at the macro level.

7.3 APPLICATION OF PROBABILITY MODELS

The second major methodological approach contemplated, in order to calculate an "expected" level of road accidents for each urban area, involved the use of probability models and techniques. The use of probability models in road accident analysis is quite extensive, and although many of the different purpose models overlap considerably, up to the present there is no suitable "general purpose" model. Thus any study using these models must develop its own specialised model.

The use of probability models and various poisson models in particular, can be justified if a road accident is considered to be the consequence of a combination of a large number of circumstances, where each circumstance (or event) has a small probability of leading to an accident. Following Erlander (1971) this can be demostrated statistically, quite simply if one considers a Poisson model for the number of accidents experienced by an individual. If we let the outcome of an event i be X_i and define X_i such that

$$X_i = \begin{cases} 1 & \text{if event } i \text{ results in an accident} \\ 0 & \text{otherwise} \end{cases}$$

then,

$$S_n = \sum_{i=1}^{n} X_i$$

denotes the total number of accidents experienced in connection with n events. If we assume that the outcome of events are mutually independent, (i.e. that the random variables X_i are mutually independent) and that the probability of an accident at event i equals P_i then the outcome of the successive events represents a sequence of "Bernoulli Trials" with variable probabilities. It is known that as $n \rightarrow \infty$ the probabilities P_i depend on the number of events n in such a way that the largest P_i tends to zero and the sum,

$$\sum_{i=1}^{n} P_i = \lambda$$

remains constant, then S_n has in the limit the Poisson distribution

$$P(S_n = k) \quad \frac{\lambda^k e^{-\lambda}}{k!} \qquad \text{as } n \rightarrow \infty$$

Hence for large n the distribution S_n can be approximated by the Poisson distribution.

The intuitive appeal of such a model for road accident occurrence is due to three principal reasons. Firstly the number of events that may lead to accidents is very great. Secondly the outcome of different events seems to be fairly independent, although problems arise here over fatal and serious accidents which of course interrupt the flow of events for the particular individual concerned. Thirdly the probability of an accident in any specific event is very small. Thus the major assumptions of the Poisson model seem to be satisfied by the observed characteristics of road accidents.

Whilst it is also possible to use other probability distributions to describe road accidents, especially some kind of weighted poisson distribution, where it is assumed that the parameter λ

is itself a random variable with distribution function $G(\lambda)$, (this allows account to be taken of the differing risk levels of different individuals) the added complications are not justified by much improvement in the empirical results.

The Poisson model described above will predict the number of accidents experienced by one individual but it is also possible to apply the model to a collective of individuals, such as those living in any one urban area. The sum of independent Poisson variables is poisson distributed and, similarly the superposition of independent poisson processes is itself a poisson process. All that is necessary is to sum the parameter values. Assume that there is a collective consisting of N individuals who experience $X_1 \ldots\ldots X_N$ accidents during the time period (O.T). Assume further that for any randomly chosen individual the occurrence of accidents can be described by a poisson process, then the total number of accidents for the whole population of N individuals will also be a poisson process. Hence

$$P\left(\sum_{i=1}^{N} X_i = k\right) = \int_0^{\infty} \frac{(\lambda T)^k}{k!} e^{-\lambda T}$$

When studying the whole population of N individuals the ordinary poisson process with fixed but unknown parameter value will usually be more preferable than the use of the weighted distribution $G(\lambda)$.

The major advantage of the use of probabilistic models lies to a great extent in their capacity to treat quantitatively, events, which have a common structure but differ in many details. Unfortunately they also have the disadvantage of ignoring and wasting any valuable background data which may be available other than the total number of accidents, and in some cases the time. For this reason, other types of models which attempt to study the influence of various variables on the number of accidents, are preferable, especially those using the technique of multiple regression analysis. However, this is not to denigrate the use of all probability models. Various researchers have attempted to combine such models with the technique of multiple regression. Although such models are perhaps the most satisfactory at the present time they are usually only possible when dealing with very limited areal units such as a certain stretch of road etc., and are therefore limited in use when dealing with urban systems and the resultant spatial comparisons. Even so, the potential of these compromise models is very great, and this can perhaps be demonstrated most rewardingly by looking at a recent model developed by Jorgensen of the Danish Council of Road Safety Research, in order to evaluate the traffic safety implications of different choices of construction stages of a new motorway scheme in Jutland. (Jorgensen 1969).

If one assumes that the density of road accidents (d) is some function of the traffic flow, such that,

$$d = f(N)$$

where N is some estimate of traffic flow. If the data is disaggregated into different road classes within a unit area, then a relationship can be assumed as below,

$$d = a \, . N^{p}$$

where a and p are constants characterising that road class and N is some observed measure of annual average daily traffic.

If one now assumes that the annual number of accidents on a certain road section is a Poisson distributed variable then it is possible to extend the simple model postulated above. Let U_{ij} be the number of accidents in a certain year on section number j of road class i. It is assumed that U_{ij} is a poisson variable such that,

$$E(U_{ij}) = a_i . N_{ij}^{P_i} . L_{ij} = a_i . Y_{ij} \qquad (1)$$

where $Y_{ij} = N_{ij}^{P_i} . L_{ij}$ and

L_{ij} = the length of road section (ij)

Thus the distribution is,

$$P_r(U_{ij} = u) = \frac{(a_i . Y_{ij})^u}{u} \; e^{-ai . Y_{ij}} \qquad (2)$$

N_{ij} and L_{ij} are known from observed data and the only unknowns are a and p. Although it is impossible to estimate these parameters using the standard least squares method, values can be derived using an iterative process by setting P_i at chosen values and estimating for a_i from the equation,

$$a_i = \frac{\sum_j U_{ij}}{\sum_j Y_{ij}} = \frac{U_i}{Y_i} \qquad (3)$$

The p_i value chosen is the one at which the sum of squares of the deviations $S(P_i)$ between the estimated curve given by (1) and the observed values U_{ij} is minimised.

The application of these two mentioned parameters is very important for, since they characterise the different road classes, and can also be derived for different areal units, they can be compared to the β coefficients of multiple regression, and used as indicators of the spatial variation of road accidents and other related variables.

This model by Jorgensen is not therefore a simple standard probability model since it is necessary to know beforehand the original relationship between the number of road accidents and the relevant variables. In this instance it was assumed for simplicity sake that $d = a \cdot N^p$, but this must be verified using some other statistical technique, usually regression or correlation, before the model can be reasonably applied. Consequently, such models are perhaps more useful once the amount of information and theories about the basic properties of road accidents have been greatly increased. For this reason such a model was not pursued in this research, but the technique advanced was that of multiple regression.

Chapter 8. METHODOLOGY II: MULTIPLE REGRESSION ANALYSIS

8.1 INTRODUCTION

Although the methodology used in this research can be generalised as a multiple regression analysis, in practice the method involved the use of several other statistical techniques and so as an aid to comprehension, a flow diagram, showing the relevant techniques and approach is shown in (Fig. 8.1.1). The main purpose of this chapter is to elucidate and explain this flow diagram, whilst explaining the methodology used in this research, and also pinpointing the limitations inherent in any such approach.

8.2 THE MULTIPLE REGRESSION MODEL

Of all the techniques used in road accident analysis, that of multiple regression has proved the most popular and also the most worthwhile. This is so because of the complex nature of the road accident situation and the multitude of variables which have to be considered in any realistic model. Similarly, the use of the multiple regression model also enables the inclusion of background information whenever it becomes available.

There are several examples in the literature of the use of multiple regression models in road accident analysis, but since these are mainly dealt with in other sections of this report only three simple uses will be mentioned here.

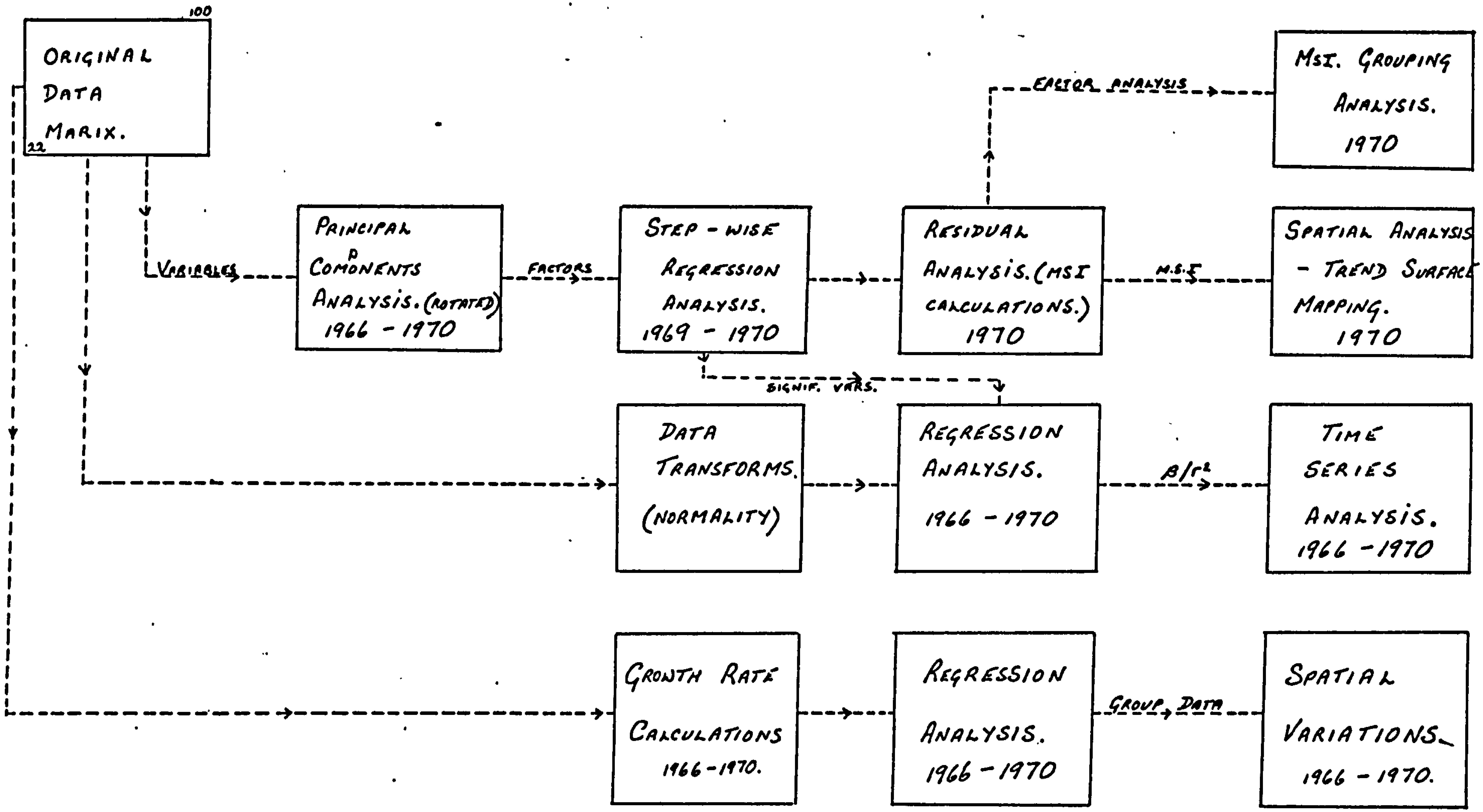


Fig. 8.1.1 FLOW DIAGRAM OF METHODOLOGICAL APPROACH

The model developed by Smeed (1949), whilst studying the variation in road accident volumes in 20 different countries, was basically a modified regression model. Using 1938 accident data he derived an equation which related the number of deaths in road accidents in any one year to the number of licenced motor vehicles and the population level. A line corresponding to the formula,

$$D = 0.0003(NP^2)^{\frac{1}{3}}$$

where,

N = the number of licenced motor vehicles

P = the population level

D = the number of road accident deaths,

was found to be the best fit line for the collected data.

Along with the work by Rempler and Blomquist (1956) in their study of road accidents in Sweden, the Smeed study used compound variables such as the volume of traffic, but it would seem more meaningful to disaggregate such data and perhaps divide traffic variables among the different traffic elements. A good example of this kind of disaggregation can be seen in the work of Rempler and Westergren (1966). In the present research, such disaggregation was also attempted, but the possibilities proved very limited due to the size and nature of the areal units used.

As stated previously, the main aim of the methodological approach in this research was to predict an "expected" number of road accidents per urban area. In statistical terms and within the multiple regression model used, this is known as point estimation. To understand how the multiple regression model is able to provide this point estimation, it is best to describe the regression model as simply as possible. The multiple regression model attempts to fit a linear surface to a set of variable data points, such that the sum of the squared deviations from that surface is minimised. The model can be represented by

$$Y = a + \sum_{i=1}^{k} b_i X_i + e$$

where, Y is the dependent variable

X_i is the k independent variables

a and b_i are the regression coefficients, representing the parameters of the model for a specific population.

e is a stochastic disturbance term.

When the number of independent variables is above five or six the calculation of the various parameters becomes very complex, and some standardising procedure is necessary, and very often the formulation in terms of matrices is the one preferred since it is the method most easily applied in computer based solutions. Before leaving this point of the discussion it is important to note that the computed b values of the regression are known as "partial regression coefficients." These coefficients each give the rate of change in the dependent variable for a unit change in the independent variable while the remaining independent variables are held statistically constant. However, because the partial regression coefficients are not independent of the particular metrics used for the variables, it is meaningless to attempt to compare directly the magnitudes of the b values, as in any way indicative of the relative importance of the different variables in the regression equation. The solution to this problem is to standardise the partial regression coefficients in such a way to enable direct comparisons. The normal method used is to calculate the beta values (β_{js}) which are dimensionless terms. They can be derived as follows:-

$$\beta_j = b_j \, (S_j/S_o)$$

where,

β_j = the standardised beta value

S_j = the standard deviation of the independent variable

S_o = the standard deviation of the dependent variable.

The use of this model in this research is important on two accounts. In the first instance the best-fit regression plane allows the calculation of a value for Y_{oi} for any set of X_i independent variables. In other words, for every urban area a value relating to the number of road accidents can be calculated from the relevant independent (structural) variable values. Thus the regression plane can be conceptualised as the "norm" from which to study spatial variations. Secondly the β coefficients allows one to study and compare the influence of each of the structural variables upon the level of road accidents. Additionally, over time, the β coefficients will show the variation in the influence of any one independent variable over time. Summarising therefore the multiple regression model enables the calculation of an "expected" number of road accidents for each urban area; the comparison of the influence of different structural variables upon the level of road accidents and finally the β coefficients allow the study of the influence of these variables over time.

Unfortunately, the multiple regression model is based upon various implicit assumptions which must be satisfied if the results attained are going to have any realistic interpretation. There has been considerable discussion within the geographical literature as to whether these assumptions have received enough, if any attention. (King 1969, Harvey 1969, Tobler 1966 etc.). A recent paper by Poole et. al. (1971) has pointed out that there are seven basic assumptions which must be satisfied if anything more than point-estimation is desired. As regards the present research three of these problems proved difficult to accept, and are discussed below.

The most limiting assumption was that of multi-colinearity. This assumption states that "the independent variables, X_i are linearly independent of each other." (Poole et. al 1971) That is the correlation between the independent variables is minimal. Obviously in this research such an assumption was untenable since several of the variables introduced had high correlation values with other variables. The correlation matrix for these independent variables is shown in Table 5.4.2.

A second limiting assumption is that of autocorrelation. The relevant assumption states that the error terms in the regression model are independent and normally distributed. When the least squares technique is used to estimate b in the regression equation Σe_i^2 is minimised. Therefore if Σe_i^2 is 0 or near to zero, then there is no autocorrelation problem, and all the relevant variables have been included. Similarly if Σe_i^2 is relatively large the problem can be twofold. There may be a sampling error or a specification error, whereby the wrong relationship has been postulated in the model. However, in both instances if the sample size is increased Σe_i^2 does not tend towards zero. Therefore it is possible to test for the occurrence of autocorrelation simply by increasing the sample size and noting whether the value of Σe_i^2 is reduced or increased. Whilst most of the discussions on autocorrelation have revolved around time-series data, it is obviously just as important in the analysis of spatial data, where one refers to cross-sections through time. Spatial autocorrelation which occurs when dealing with contiguous areas is probably more complex and therefore more difficult to accommodate than time-series autocorrelation, because the time-series data involves only one dimension whilst spatial-autocorrelation can involve many more dimensions. One of the methods used for testing for spatial autocorrelation is the Cliff-Ord Statistics (Cliff 1970). However, in this research, because the areas dealt with were in most instances non-contiguous, the problem of spatial autocorrelation was not explicitly tested for, but the autocorrelation in the regression residuals was studied by means of the trend surface analysis.

The third limiting assumption is that of normality of data. Although this assumption is of little importance as regards point estimation, it is important if any inferential statements are to be made of the regression as a whole, or the various parameters. Normality can be tested by using the formula proposed by Lambe (1967) which is given as:

$$\left| \frac{\eta}{\sigma} - 0.7979 \right| < \frac{0.4}{\sqrt{N}}$$

where,

η = the mean deviation

σ = the standard deviation

N = the number of observations.

The results of this test on all the independent variables is shown in Table (8.2.1). A study by Anderson et. al. (1968) tested this assumption of normality amongst the number of accidents in Sweden, by studying the derived trends and autocorrelations, using histograms of the standardised residuals, and correlograms. In each case, the deviation was small and tended to zero, and although the sample size was too small to make any definite conclusions, it seems plausible to accept that the assumptions of normality and

VARIABLE	LAMBE TEST	KURTOSIS	SKEWNESS
POPULATION	0.158	12.765	3.155
AREA	0.084	5.055	2.108
DENSITY	0.029*	0.700	0.458
P.T.F.	0.073	4.561	1.936
TTF	0.202	16.074	3.592
ETT	0.031*	2.428	1.320
COW	0.012*	0.768	-0.324
MSI	0.009*	0.428	-0.531
SRI	0.101	3.791	-1.249
TCI	0.016*	-0.254	0.060
CI	0.049	1.961	-1.450
DOS	0.207	20.644	3.967
SHI	0.362	30.536	5.284
AVI	0,187	14.203	2.482
DVI	0.167	7.373	1.787
TDI	0.157	11.696	3.033
HTI	0.186	12.635	3.015
CHSI	0.200	13.221	2.907
PPI	0.081	2.461	1.142
PSI	0.055	3.452	1.574
RI	0.039*	1.352	1.213
RDI	0.042*	1.115	0.573

* INDICATES A NORMAL DISTRIBUTION ($\leq$0.04)

KURTOSIS - is given by the formula:-

$$K = \frac{\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{s}\right)^4}{N} - 3$$

SKEWNESS - is given by the formula:-

$$S = \frac{\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{s}\right)^3}{N}$$

TABLE 8.2.1. NORMALITY TESTS FOR THE 22 INDEPENDENT VARIABLES 1970.

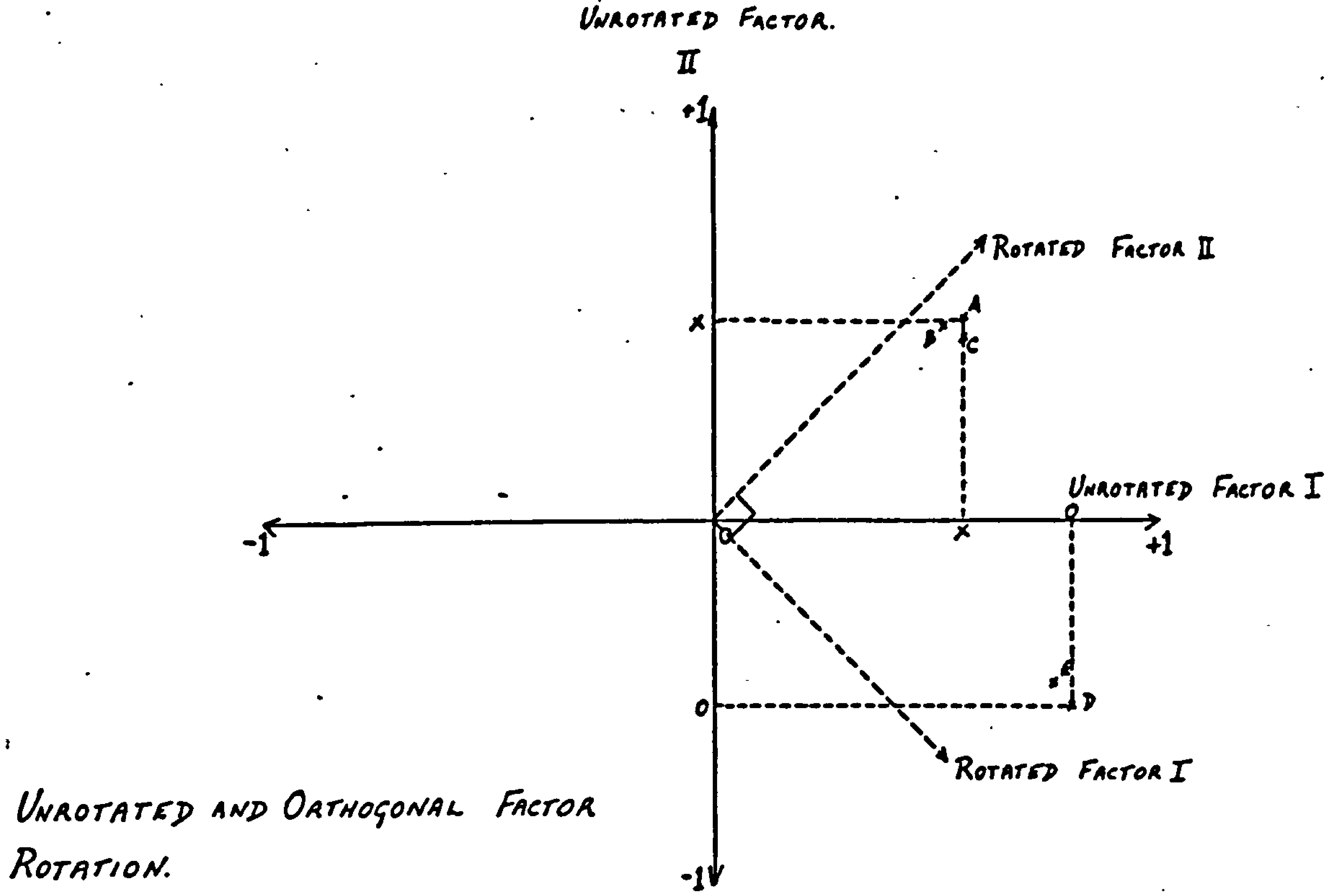
independency of the residuals, can generally be accepted.

8.3 PRINCIPAL COMPONENTS ANALYSIS

Because of the extent of these various limitations, it was decided that the original data ought to be transformed. This was done by applying a principal components analysis upon the original data matrix of 18 variables. Principal components analysis can be conceptualised as a technique which is capable of collapsing a set of intercorrelated variables into a smaller number of basic dimensions, or "composite variables." What is more important however, is that these basic dimensions are such that they are totally uncorrelated one to another (orthogonal). The solution, moreover, is such that the first dimension extracted accounts for the highest proportion of the total variance present within the original data matrix.

The interpretation of these resultant basic dimensions is one of the major problems of any research project, and this is done by reference to the "loading coefficients." In practice these loadings are simply the correlation coefficients between the original variables and the derived components. However, very often any one component will have a high loading with several of the original variables making any interpretation difficult, especially when dealing with the higher order components. For this reason it is common practice to rotate the components, in order to reduce the number of high loadings on any one component.

Mathematically, rotation can be viewed simply as the result of the indeterminacy of the principal component analysis, since an infinite number of locations could have been derived for the various axes, and the problem is which location is to be preferred. (Fig. 8.3.1). The most common rotation used is the Varimax rotation. This method attempts to simplify the columns of the factor loadings matrix, by a series of orthogonal transformations of pairs of factors. (See Harman 1960). Thus the varimax solution can be simply described as an iterative procedure aimed at finding an optimum solution for the location of rotated factors. Two limitations of this solution are that the pattern of loadings obtained may change considerably if additional factors are included in the solution, and also the solution is not very useful in cases where there is a dominant general factor. (Lawley et al 1971). For this reason it was decided to use an oblique rotational method which whilst useful in aiding interpretation of the various factors, suffers from the loss of orthogonality. The oblique solution methods were first studied by Thurstone (1947) and Thomson (1951), who both used a graphical solution. However, these solutions "involve the researcher in many arbitrary decisions about the choice of factors." (Lawley et al 1971), and so further research was directed into the discovery of analytical criteria that could be used to produce unique and less subjective solutions. The earliest work in this field was carried out by Carroll (1953), Kaiser and Dickman (1959) and Hendrickson and White (1964), who evolved the "Promax Solution." It was this solution which was used



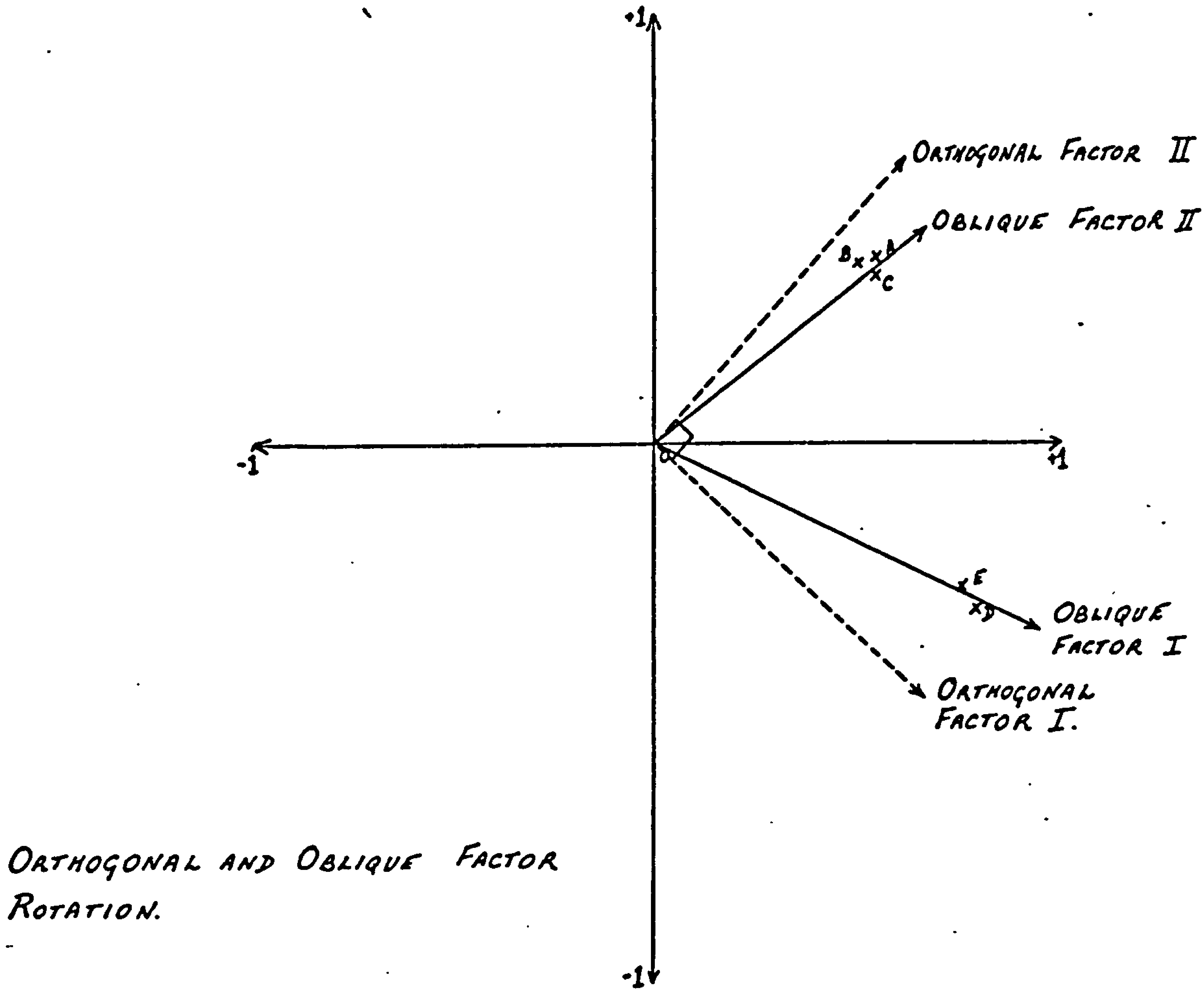


Fig. 8.3.1. POSSIBLE LOCATIONS FOR DERIVED FACTORS WITH ORTHOGONAL AND OBLIQUE ROTATION METHODS.

in this research necessitating only the one subjective decision, that of what value to use for "m" in the simplification of the "pattern matrix." This value for "m" determines the degree of obliqueness desired in the solution, and consequently therefore, the level of correlation between the various factors. It is generally accepted that values upto $m = 4$ will simplify the pattern of loadings by increasing the absolute magnitude of relatively large loadings whilst decreasing those with relatively small loadings. However, where $m > 4$ the usefulness of the simplified pattern loading matrix may be overcome by the high level of correlation produced between the various factors. For this reason a value of $m = 4$ was used in the present rotation.

The use of a principal components solution in a "two-stage regression" is quite extensive and has been reviewed at some length by Massey (1965), but only a few have actually used a rotated solution (Wong 1963, Hodge 1965) mainly because of the "communality" problem which is associated with any factor analysis solution. However, it can be argued that if the communalities are held to be equal to one, then the method used here remains a principal components solution, with rotation being used merely as an aid to identification. In the context of a two stage regression approach, the results from such an analysis would appear clearly acceptable, especially since there is the added bonus that derived scores input into the regression are standardised with a mean of zero and a standard deviation of one.

Thus the use of this analysis was useful in satisfying most of the assumptions of the multiple regression model, whilst allowing some observation of the urban structure system through the identification of the various derived components. This line of discussion will be continued at a later point in this report, (Chapter 9). Before leaving this discussion on principal components analysis and its use in a two-stage regression analysis, one further point must be elucidated, concerning the number of components to be introduced into the regression. This problem arises because of the very small amount of variance explanation, provided by the higher order components. Very often such components are deemed insignificant and thereby ignored. Alternatively some test is performed upon them to test for their actual significance. These include those tests suggested by Bartlett (1950) or Jeffers (1967) who suggests that in general, it is satisfactory to disregard all components with a variance of less than 1. However, it is suggested that when the components are to be used as regressors in a regression analysis, it is not necessarily true to maintain that just because a component explains an insignificant amount of the total variance in the original data matrix, it is also going to be insignificant in "explaining" the variation in some other variable. If this point of view is to be accepted then the obvious conclusion must be that all the components ought to be available for the regression analysis, and their significance determined accordingly. This was the approach of this research and the significance of each component was determined by using a "Step-wise regression model" rather than the traditional multiple regression model.

8.4 THE STEP-WISE REGRESSION MODEL

When the simple multiple regression model is utilised, all the available variables are introduced into the model at the same time. This is satisfactory when the relationship between all the variables is known and when one is not interested in further research hypotheses, but as is generally the case, most research problems necessitate the selection of an optimum collection of variables, from all the available variables, adding only variables which significantly affect the explanation level of the regression, whilst the other variables are held statistically constant. Essentially therefore, the problem is to derive and select the "best" regression equation. Unfortunately, since there is no statistical procedure which can guarantee an "optimum" solution, or an "optimum" group of variables, some compromise algorithmic technique has to be used. Although there are several such algorithmic procedures, the one used in this research was the "step-wise" procedure.

The step-wise regression method recursively constructs a prediction equation one independent variable at a time. The first step is to choose the single variable which is the best predictor. The second independent variable to be added to the regression equation is that which provides the best prediction (level of explanation) in conjunction with the first variable. One then proceeds in this recursive fashion, adding variables step-by-step until the desired number of independent variables has been achieved, or until no other variable will make a significant contribution to the prediction equation. It is important to note at this point that such a procedure, besides dictating the selection and therefore addition of the independent variables, also dictates the removal of variables from the regression equation. This is essential because a variable which may have been the best single variable to enter at an early stage may, at a later stage, be superfluous because of the relationship between it and other variables now in the regression. To check on this, the partial F criterion for each variable in the regression at any stage of calculation is calculated and compared with a preselected percentage point of the appropriate F distribution. This provides a judgement on the contribution made by each variable as though it had been the most recent variable entered, irrespective of its actual point of entry into the model. Any variable for which the calculated F value is less than the preselected acceptance value, and which therefore provides a non-significant contribution, is removed from the model. This process continues until no more variables can be added or removed from the regression equation.

Step-wise regression is based upon a common method of solving the system of linear equations in multiple regression, that is Gauss elimination with row and column interchanges. This computational method is also useful since it provides the two pieces of information necessary in the selection process. The first is the normalised regression-coefficient value b that the prospective independent variable would have if it were brought into the equation on the next step. It is the value of this coefficient which is tested by the F statistic during the selection process, and as already indicated, if F is too small, there is little reason to add that independent variable to the regression equation.

The second piece of information used in the selection process is the pivot element which would be involved in bringing that variable into the equation. This value is known as the tolerance level. If the tolerance is small then that variable is nearly a linear combination of variables already in the equation. If it is really a linear combination of independent variables already in the equation, then the tolerance value will be zero. A large tolerance indicates that a new "dimension" is being added to the prediction equation. The tolerance has a value range of zero to one. The amount of additional variance explained by adding the new variable is the product of the normalised regression coefficient b squared and the tolerance. Thus, even if the prospective b is large, a small tolerance value will negate the value of that variable being added to the equation. Consequently, the step-wise procedure never enters a variable into the equation if the tolerance is below a specified minimum. This is, therefore, another way in which the computational accuracy of the program is upheld.

This step-wise procedure described above, was the model used in this research to derive the desired regression planes, and the resultant level of road accidents in any one urban area ascertained. The program used was the standard BMD step-wise regression program which was converted from the original Fortran II program, to 360 Fortran IV(H-Level) at Salford University.

The confidence interval used to obtain the inclusion and deletion values for the F statistic was the $\alpha = 0.1$ level which gave F values of 4.00 for inclusion and 3.92 for deletion. Similarly the default value of 0.001 for the tolerance level was used, although in most instances this was never called into consideration.

8.5 THE STUDY OF RESIDUALS

Having obtained the necessary "expected level" of road accidents, the next problem was then how to relate these values to the actual number of road accidents in any urban area. In a way this problem can be generalised as that of how to deal with the "residuals" from regression analysis. That is, the researcher is interested in the difference between observation and theory. Because of the complexity of spatial variations there is obviously considerable unexplained variation and the residual values can be conceptualised as reflecting in part, the effect of other possibly unknown variables, and as such can be used to gain a greater insight into the behaviour of particular phenomenon. For example, John Stuart Mill in a "System of Logic" (1874) presents the "methods of residues" as the fourth Canon, and consistently praised it as a highly respectable research technique. Thus, "Subduct from any phenomenon such part as is known by previous inductions to be the effect of certain antecedents, and the residue of the phenomenon is the effect of remaining antecedents." Similarly, whilst speaking generally of the method of residues, "Of all the methods of investigating laws of nature, this is the most fertile and unexpected result; often informing us of sequences in which neither the cause nor the effect were sufficiently conspicuous to attract of themselves the attention of the observers. The agent C may be an obscure circumstance, not likely to have been perceived unless sought for, nor likely to have been sought for until attention had been awakened by the insufficiency of the obvious causes to account for the whole of the effect." (Mill 1874). Thus in a way, Mill's Method of Residues is simply a search procedure used in the analysis, for the discovery of further important variables or even new theoretical concepts. As such this approach has developed into the more modern "deviant-case analysis" which is important because as Horst points out, "the most interesting and useful parts of a prediction study should be the investigation of the cases which have been incorrectly predicted in the new sample." (Horst 1941).

Whilst there has been a rapid increase in the use of "residual analysis" one of the major problems encountered has been the form of the residual index. Thomas identifies four of the more common forms used, favouring the standardised residual,

$$(X_{oi} - Y_i)/S_{oi}$$

where, X_{oi} = the computed value for area i

Y_i = the observed value of Y for area i

S_{oi} = the standard error of estimate

but there are many other forms which are viable and can only really be determined by the individual research project. Thus Taaffe et al (1963) in their study of the transport development in Nigeria and Ghana used the simplified, non-standardised residuals,

$$(X_{oi} - Y_i)$$

and yet the mapping and study of these residuals led the authors to suggest five additional factors, which they had previously not considered and which ultimately added further findings concerning the transport systems within those two nations.

The form of the residual used in this research relates to the third form as mentioned by Thomas (1960) and is given as the magnitude of the observed value for the ith unit as a proportion of the computed value for that area. When multiplied by 100 this residual was used as the value for the "Mean Safety Index" for each area at time t,

$$MSI_{i.t} = (Y_{i\,t}/X_{oi\ t}) \times 100$$

where,

$MSI_{i.t}$ = the mean safety index for area i at time t.

$Y_{i.t}$ = the observed value for area i at time t

$X_{oi.t}$ = the computed value for area i at time t

As Thomas points out however, this, "is not a residual according to the strictest of definitions" due to the fact that the difference between the two values is ignored in preference to the measure of their relative values. However, since this index is closely related to the more common form of residual $(X_{oi} - Y_i)/Y_i$, the mapping of such an index is obviously useful in the study of unexplained variance.

The derived mean safety index is therefore designed such that it gives a variation about 100. Where the index exceeds this base value, the observed value is greater than the computed value, and the "response" of that urban area is worse than would be expected given the present set of independent variables. Similarly when the index is less than 100 the observed value is less than the computed and therefore the "response" of that urban area is better than theorised.

8.6 TREND SURFACE ANALYSIS

Whilst the Mean Safety Index is meaningful in its own right, as already suggested, much more can be achieved by mapping this index and by creating some "continuous" surface which can possibly predict the relationship in other "unsampled" urban areas, and also for suggesting new variables which ought to have been taken into account. Although such mapping procedures as simple isopleth maps are useful, the problem of discontinuous data can prove a serious problem; for this reason a further stage was added to this research approach by means of the application of Trend Surface Analysis techniques.

Trend surface analysis is simply an extension of the multiple regression model into a third dimension. It is a method of fitting progressively higher order polynomial equations to a set of X,Y and Z observations in which the X and Y values are co-ordinate data and the Z values are observations of some variable. This type of analysis has been used extensively in the various "earth sciences" where the problem has been mainly extrapolation of existing data or the reconstruction of surfaces from fragmentary data (as in the "ghost stratigraphies" of Whitten (1959). However, trend surface analysis has also been used in order to extract regional trends and examining the deviations from these trends. Thus Krumbein and Graysbill (1965) define trend surface analysis "as a procedure by which each map observation is divided into two or more parts: some associated with the "larger-scale" systematic changes that extend from one map edge to the other and the others associated with "small-scale" apparently non-systematic fluctuations that are imposed on the large-scale patterns." (p.321). In other words therefore one of the main aims of trend surface analysis is to differentiate between the broad regional distribution of variability, from the non-systematic, local and chance variations. In terms of road accident analysis, this shows the differentiation between those factors which necessitate national attention and those which must be dealt with at the local level.

With the derivation of computer-based solutions to the problem of fitting the various polynomial surfaces, Trend-surface analysis has been applied in many other fields other than those simply related to geology. Obviously the usefulness of Trend Surface mapping is greatest where continuous data is available, and regularly spaced sampling points are used. However, the latter problem can be overcome, although the solution of the polynomial equations becomes much more complex, and the former problem has been discussed by Chorley et al (1965). They argue that surfaces need not be continuous and that "population, like light may be profitably regarded either as a series of discontinuous quanta or as a continuum." The fitting of trend surfaces to social and economic data may provide then a "useful yardstick for describing geographical patterns." (Chorley et al 1965). Subsequent work has followed these suggestions with trend surfaces being applied to such diverse subjects as spatial perception studies, (Gould 1966), the location pattern of urban services among a set of towns (Fairbairn et al 1967) and the growth of urban areas in England and Wales, (Robson 1973).

The formulation of the trend surface equations is given below for a first order surface (linear) and a second order surface (Quadratic).

$$Z = a + BX + cY \quad \text{(1st order)},$$

and, $$Z = a + bX + cY + dX^2 + eXY + fY^2 \quad \text{(2nd order)}$$

where, Z = the value of the research variable.

X & Y = two grid co-ordinates

a,b,c,d,e,f = parameters.

The shape of the various surfaces upto the quartic surface are shown in (Fig. 8.6.1.). Thus the trend surface model is derived by creating a hierarchy of mathematical surfaces which are fitted to the original data, and the explained portion of the observed variability, conventionally expressed as the per

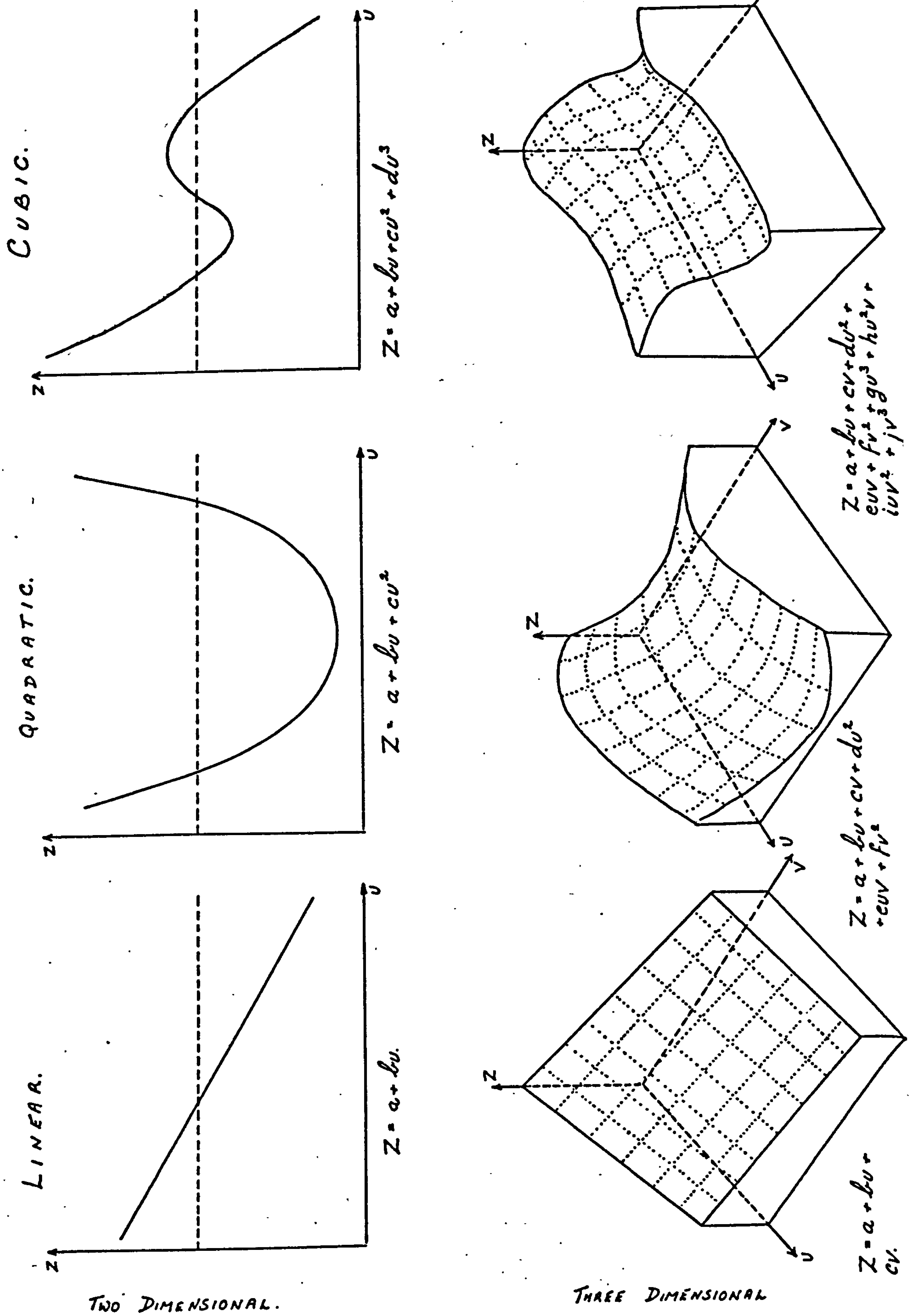


Fig. 8.6.1. DIAGRAMMATIC REPRESENTATION OF THREE TREND SURFACES.

centage reduction in the total sum of squares, may then increase with surfaces of successively higher degree. In most instances, however, the lower order surfaces have been preferred since they tend to indicate the regional trend, whilst surfaces which exceed the 6th order tend to become very complicated, and are near approximations to the original complex distribution. Therefore in this research, surfaces were produced for twenty dependent variables (1970) all of which extended at least as far as the 6th order surface, but none exceeded the 7th order (Appendix 2).

8.7 TEMPORAL VARIATION ANALYSIS

So far this analysis of the methodology used in this research has been limited to the study of the spatial variation in road accidents. Turning to the problem of temporal variation, one has to return to the multiple regression model. As already indicated, the multiple regression model is useful in two ways when studying the change in road accident characteristics over time. In the first instance, the amount of the total variation in a dependent variable, that is explained by a set or group of independent variables can be studied over time simply be reference to the R^2 value. If the impact of these independent variables is constant over a given time period, then one would also expect the level of explanation by the model, also to remain constant. Variation in either statistic therefore suggests a changing relationship. Secondly the influence or relationship between the dependent variable and the individual independent variables can also be studied over time by recourse to the standardised β coefficients. Unfortunately since the transformation of the original data matrix by means of the principal components analysis, was completed individually for each year, the various derived components also varied year by year. Therefore it was necessary to return to the original data matrix. To escape the problems noted previously in the use of untransformed data and multi-colinearity etc. a new regression equation was developed for each dependent variable used in the time analysis, which included only thoseindependent variables which were shown to be significant by the step-wise regression during the spatial analysis. Where a component was heavily loaded on more than one variable, the variable with the highest loading was the one entered into this section of the analysis. Once these modifications had been completed the time analysis was completed as described above. However, since the road accident data was very difficult to obtain prior to 1966 and because the disaggregation of the data was very limited, it was only possible to conduct this time analysis for three dependent variables. Total casualties, total casualties per thousand population, and total casualties per vehicle mile (10^6). 1966 - 1970.

Because of the limitations posed by the lack of data, it was deemed useful to look at the temporal variation in another way, basically through the growth rate in the number of road accident casualties from 1966 to 1970. The hypothesis put forward was, how much was the level of road accidents in time t + 1 the response of the level of road accidents in time t. Alternatively, do towns of different size and therefore different road accident levels have accident growth levels which are proportionally equivalent. This hypothesis can be tested by the regression equation,

$$X_{t+1} = a + bX_t + e$$

where,

X_t = the number of casualties at time t.

a,b, = constants

e = normally distributed error term.

If the absolute growth of road accident casualties is proportionate and independent of the original level of road accident casualties, then one would expect b = 1.0, with 'a' representing the overall growth in road accident casualties for all the urban areas studied. Obviously where the value for b is less than 1.0 the areas with a small road accident casualty level must be growing at a faster rate than those with an original high level. Of course one must also take into account the effect of the error term which should be at a minimal level if any significance is to be put on the consequent results.

A second hypothesis which was postulated was whether the actual growth rate in road accident casualties was a function of the original road accident casualty level. If the growth rate in road accident casualties is defined as,

$$G.A_{t.t+1} = (TC_{t+1} - TC_t) \cdot \left(\frac{TC_{t+1} + TC_t}{2}\right)^{-1}$$

where,

$GA_{t.t+1}$ = the growth rate in casualties from time t to t + 1

TC_t = the total number of casualties at time t then this second hypothesis which can be equated as,

$$GA_{t.t+1} = f(TC_t)$$

can be tested by the simple regression equation,

$$GA_{t.t+1} = a + bTC_t + e$$

These two hypotheses and their consequence to road accident analysis will be developed further in the following discussion on the results obtained during this research. At present it is sufficient to mention the problem and the methodology used without delving any further into the respective consequences.

8.8 SUMMARY OF METHODOLOGY

The last two sections of this report have attempted to explain the methodology used in this research. Because of the complexity of the approach it would seem best to simply summarise the methodology before moving onto the results obtained.

As regards the spatial analysis, the methodological approach can be conceptualised as a five stage analysis. (Fig. 8.1.1). These stages are listed below:- (i) Collection and manipulation of the original data matrix. (ii) Transformation of the original data matrix by means of an "oblique " (promax) rotation principal components analysis. (iii) The input of the derived principal component scores into a "step-wise" regression analysis with a 10% confidence interval. (iv) The calculation of the mean safety indices from the residuals given by the "step-wise" regression analysis. (v) The input of the mean safety indices into a trend surface analysis to study the spatial variations.

Parallel to this analysis the temporal analysis used the variables designated by the step-wise regression (from the original data matrix) and input them into a separate regression model in order to derive the β (standardised) coefficients, and the level of explanation (R^2) for the years 1966 to 1970. At the same time the absolute increase in road accident casualties, and the growth rate (1966 - 1970) in road accident casualties were studied according to their initial road accident casualty level.

This then was the methodological approach used and the consequent results are given in the following sections of this report.

Chapter 9. THE ANALYSIS AND PRESENTATION OF RESULTS

9.1 INTRODUCTION

The results obtained in this research, using the methodological approach outlined in the previous section of this report are presented in this chapter. In order to facilitate ease of comprehension, the presentation of these results will be disaggregated, dealing with different sections of the approach individually. Once such a presentation has been effected, it will be possible to draw these results into a suitable form in order to reach some meaningful conclusions, as regards this piece of research.

9.2 THE PRINCIPAL COMPONENTS ANALYSIS

Although the principal components analysis was initially introduced as a means of reducing the multi-colinearity amongst the explanatory variables, through use of this technique several other advantages were attained as well as further interesting information, as regards the inter-relationship of the various sub-systems of the road accident system. For example, if the component scores are introduced into the regression equation, not only is the problem of multi-colinearity reduced, but also these scores are standardised with a mean of zero, and variance of one. Additionally these component scores have the further advantage of being almost normal. Thus by this means most of the limitations previously mentioned, inherent in multiple regression analysis, have been satisfactorily removed.

The principal components analysis was carried out on the 18[1] initial explanatory variables. The importance of the rotated components in explaining the total variation of the explanatory variables is shown in Table (9.2.1) which lists the components according to their variances in declining order. (1970). As a simple means of reference this listed ranking of components will be known as the "Variance Ranking."

It will be noted from the variance ranking that only 6 components have an eigenvalue exceeding 1.0. However, it should also be noted that these six components also account for 81.75% of the total variation within the list of explanatory variables. This in fact indicates one of the major properties of P.C.A. The variances (eigenvalues) of the higher number components are very small and such components are often disregarded as non-significant, especially when they are used in a two-stage regression analysis. In some instances meanwhile, it may be totally fallacious to exclude components from the regression analysis for this reason. It could, for example, be true that a component which may be non-significant when judged by the proportion of the total variance it explains, may nonetheless still be significant when used in a regression constructed to explain variation within some other variable. Accordingly, all the derived components within this

1 The four explanatory variables; Police prosecution index, Police Severity Index, Rainfall Index and Road Density Index, were not included in this initial PCA because they were not principally related to the urban structure of the 100 sampled urban areas, or in the case of the Road Density Index, because of lack of data available for the 49 non-county borough areas

/contd...

COMPONENT NUMBER	COMPONENT (VARIANCE) EIGENVALVE	CUMULATIVE EXPL. VARIANCE (%)
1	5.863	32.570
2	2.733	47.751
3	2.019	58.968
4	1.777	68.838
5	1.298	76.048
6	1.034	81.795
7	0.731	85.856
8	0.593	89.151
9	0.464	91.731
10	0.399	93.949
11	0.327	95.767
12	0.287	97.360
13	0.216	98.562
14	0.137	99.325
15	0.091	99.829
16	0.016	99.917
17	0.011	99.976
18	0.004	100.000

TABLE 9.2.1. THE VARIANCE RANKING FOR THE 100 URBAN AREAS 1970.

analysis were treated as potential independent variables in the regression analysis.

Because the results of the P.C.A. were rotated using a Promax rotation with k being assigned the value of 4, the derived components were not orthogonal, as is usually the case in a simple P.C.A. Instead there was a degree of correlation noted between the components, and these correlations are shown in Table (9.2.2.). (1970). From this table it can be seen that there are only 9 significant correlations where the level of correlation is above the value $r = \pm 0.50$. There is however one correlation which reaches the high value of $r = -0.76$, but since this is between the last two derived components (17 and 18), the restrictions due to this fact are only marginal, especially when matched against the increased feasibility of interpretation of the derived components.

The importance of the interpretation (and therefore rotation of) the derived components in a piece of research cannot be overemphasised. If any meaningful interpretation is to be given to the results of the regression analysis, then it is obvious that one must be able to identify the various inputs into the model, which in this instance are the independent variables. In most instances, the P.C.A. without rotation is unable to satisfy this need, and as Kendal (1957) states, "In many cases our principal

---

1 contd. included within the sample. However when the multiple regression model was used, these indices were included as explanatory variables. In order to equate them to the Factor Scores used from the P.C.A., these variables were also standardised using the formula,

$$Z_i = \frac{X_i - \bar{X}}{\sigma X}$$

/contd....

FACTORS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.	1.00																	
2.	0.03	1.00																
3.	-0.23	0.16	1.00															
4.	-0.18	0.01	0.01	1.00														
5.	-0.16	0.10	0.30	-0.06	1.00													
6.	0.05	0.06	-0.30	0.18	-0.13	1.00												
7.	-0.05	0.18	0.12	-0.17	0.32	-0.02	1.00											
8.	-0.10	-0.48	0.01	-0.07	0.04	-0.09	-0.00	1.00										
9.	-0.54	-0.00	0.16	0.13	0.31	-0.02	0.14	0.21	1.00									
10.	-0.02	0.13	0.13	0.59	0.09	0.08	-0.11	-0.14	0.08	1.00								
11.	-0.07	-0.52	0.08	-0.04	-0.22	0.01	-0.30	0.37	0.04	-0.02	1.00							
12.	0.44	-0.06	-0.39	-0.09	-0.35	0.35	0.16	-0.11	-0.46	-0.10	-0.06	1.00						
13.	0.16	0.37	-0.58	0.07	-0.20	0.19	0.01	-0.11	-0.08	-0.09	-0.40	0.23	1.00					
14.	0.13	0.53	0.19	0.21	-0.11	0.06	-0.19	-0.56	-0.19	0.14	-0.33	-0.01	0.05	1.00				
15.	-0.14	0.01	-0.13	0.06	-0.13	0.40	0.25	-0.14	-0.10	-0.10	-0.05	0.49	0.16	-0.08	1.00			
16.	-0.17	-0.09	-0.18	-0.04	-0.09	0.16	0.15	-0.03	-0.11	0.02	0.00	0.42	-0.01	-0.08	0.03	1.00		
17.	-0.62	-0.24	-0.04	0.04	0.06	0.26	0.21	0.09	0.46	-0.09	0.07	0.12	-0.09	-0.25	0.38	0.36	1.00	
18.	0.59	0.07	0.01	-0.10	-0.10	-0.26	-0.31	-0.04	-0.34	0.01	0.03	-0.13	0.06	0.18	0.47	-0.42	-0.76	1.00

TABLE 9.2.2. CORRELATION VALUES BETWEEN THE DERIVED PROMAX FACTORS (k = 4) 1970.

components do not have an identifiable separate existence and are to be regarded as convenient mathematical artefacts." Dhrymes (1970) also makes the same point when he argues, "in general, one does not or cannot interpret principal components in an intuitively meaningful way. They are best regarded as the outcome of a statistical technique that handles certain types of problems efficiently."

In this research similar problems were encountered as those described above. However, following the use of the Promax rotation, it was possible to put some meaningful interpretation on almost all of the derived components and these are described below.

When one is trying to put some interpretation upon a derived component there are two sources of information:-

(i) FACTOR PATTERN MATRIX - to factor analyse means to express a variable as a linear combination of independent variables, either defined or inferred. The Factor pattern matrix contains the regression weights of the common factors and therefore, tells the researcher the composition of a variable in terms of hypothetical factors. The oblique factor pattern matrix obtained for 1970 is given in Table (9.2.3).

---

1 contd.. where Z = the standard deviate score for area i,
σ_x = the standard deviation of variable X,
$\bar{X}$ = the mean of variable X,

such that each variable had a transformed mean value of zero and a standard deviation of 1.

/contd...

FACTORS

VARIABLES	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.	1.09	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	-0.15
2.	0.84	–	–	–	–	–	–	–	–	–	–	–	–	–	-0.38	–	–	–
3.	0.04	–	–	–	–	–	0.25	–	–	–	–	0.89	–	–	–	–	–	–
4.	1.01	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.23	–	–
5.	0.94	–	–	–	–	–	–	–	–	–	–	–	–	–	0.25	–	-0.10	–
6.	0.14	–	–	–	–	–	–	–	-0.92	–	–	–	–	–	–	–	–	–
7.	–	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
8.	–	–	-0.10	–	–	–	–	–	–	–	–	–	0.93	–	–	–	–	–
9.	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
10.	–	–	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–	–	–
11.	–	–	–	–	–	–	–	–	–	0.99	–	–	–	–	–	–	–	–
12.	–	–	–	–	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–
13.	–	–	–	–	–	–	0.99	–	–	–	–	0.32	–	–	–	–	–	–
14.	–	-0.14	–	–	–	–	–	–	–	–	–	–	–	-0.86	–	–	–	–
15.	–	–	–	–	–	–	–	-0.99	–	–	–	–	–	–	–	–	–	–
16.	1.09	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.15	–
17.	–	–	–	–	–	–	–	–	–	–	0.99	–	–	–	–	–	–	–
18.	–	–	–	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–	–

TABLE 9.2.3. OBLIQUE FACTOR PATTERN MATRIX (100 AREAS k = 4) 1970 - VALUFS>± 0.1

(ii) FACTOR STRUCTURE MATRIX - this second source of information relating to the composition of the derived components is simply a matrix of correlation coefficients between each variable and each factor. In many ways this factor structure is more useful than the pattern matrix because not only does it indicate the important associations and their direction, but it also gives some indication of the strength of the association between components and variables. The correlation between a component and its constituent variables becomes weaker as the components are taken seriatim in variance ranking. The effect of the rotation of the factors can also be seen by the fact that the resultant correlations of any significance on each component, is limited to very few variables. The Factor structure matrix is shown in Table (9.2.4). (1970).

From this structure matrix it can be seen that for components 1 - 11, there is always some variable which has a high correlation coefficient ($\geq \pm$ 0.60 = r). However, for higher order components, the ease of interpretation is drastically reduced with the result that the maximum correlation coefficients for the 17th and 18th components are r = 0.07 and r = -0.08 respectively. Except for

---

1 contd . Similarly, the level of correlation between these standardised variables and the derived components was not sufficient to invoke the problem of multi-colinearity which had been one of the major reasons for using the P.C.A. in the first instance.

FACTORS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.	0.47	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.08
2.	0.37	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.20	-	-	-
3.	0.44	-	-	-	-	-	0.19	-	-	-	-	0.38	-	-	-	-	-	-
4.	0.41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.14	-	-
5.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.14	-	-0.05	-
6.	-	-	-	-	-	-	-	-	-0.60	-	-	-	-	-	-	-	-	-
7.	-	-	0.62	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8.	-	-	-	-	-	-	-	-	-	-	-	-	0.56	-	-	-	-	-
9.	-	0.56	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.	-	-	-	0.68	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11.	-	-	-	-	-	-	-	-	-	0.71	-	-	-	-	-	-	-	-
12.	-	-	-	-	-	0.78	-	-	-	-	-	-	-	-	-	-	-	-
13.	-	-	-	-	-	-	0.74	-	-	-	-	-0.13	-	-	-	-	-	-
14.	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.55	-	-	-	-
15.	-	-	-	-	-	-	-	-0.72	-	-	-	-	-	-	-	-	-	-
16.	0.47	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.07	-
17.	-	-	-	-	-	-	-	-	-	-	0.71	-	-	-	-	-	-	-
18.	-	-	-	-	0.79	-	-	-	-	-	-	-	-	-	-	-	-	-

TABLE 9.2.4. OBLIQUE FACTOR STRUCTURE MATRIX FOR 100 URBAN AREAS 1970 ($k = 4$) - VALUES>±0.1

the first component therefore, the interpretation of the derived components was based on the highest recorded 'r' value, or a combination of 'r' values in the higher components. In the instance of the first component, however, it was found that there were five variables with high correlation coefficients, these were; Population, Area, Total Traffic Flow, Peak Traffic Flow, and Total Dwelling Index. It can be seen that these variables are all related to the size and traffic flow level within any urban area. For this reason this component was interpreted as a "Size/Interaction" dimension of the urban areas, accounting for approximately 32% of the total variation within the original data matrix. Other interpretations are given in Table (9.2.5).

The importance of the promax rotation can once again be readily appreciated if the figures in Table (9.2.6) are studied. This table shows the major correlation coefficients between variables and derived components, following the initial P.C.A. That is, before any rotation had taken place, and using the original factor loadings matrix. From this table two conclusions can immediately be made:-

(i) Most of the derived components had reasonably high correlations with a fair number of original variables. Following rotation of the components however, this multiplicity of correlation is replaced by a severely reduced "set" of significant variables, usually with higher individual correlation coefficients.

(ii) Before rotation, the interpretation of the derived components is inhibited due to the regular appearance of certain variables

FACTOR	INTERPRETATION	MAXIMUM CORRELATION WITH VARIABLES
1	INTERACTION VARIABLE	0.472
2	SRI	0.558
3	COW	0.619
4	TCI	0.682
5	CHSI	0.790
6	NEG.DOS	-0.782
7	SHI	0.741
8	NEG.DVI	0.719
9	NEG.ETT	-0.603
10	CI	0.715
11	HTI	0.707
12	DEN	0.375
13	MSI	0.556
14	NEG.AVI	-0.546
15	NEG.AR/TTF	-0.203/0.137
16	PTF	0.143
17	NEG.TTF/TDI	-0.046/0.073
18	NEG.POP	-0.082
19	PPI	N.A.
20	PSI	N.A.
21	RI	N.A.
22	RDI	N.A.

TABLE 9.2.5. INTERPRETATION OF DERIVED FACTORS (1970)

FACTORS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.	0.97	-0.07	0.16	0.05	0.13	-0.02	-0.06	0.04	0.04	-0.06	-0.00	0.02	0.03	-0.00	0.05	-0.01	0.05	-0.05
2.	0.84	-0.09	0.31	0.17	0.05	-0.30	-0.12	0.05	0.06	-0.06	-0.00	-0.08	0.07	-0.03	-0.17	-0.05	-0.04	-0.01
3.	0.59	-0.02	-0.23	-0.31	0.15	0.53	0.32	0.01	-0.03	0.03	0.11	-0.28	-0.04	-0.01	-0.07	-0.03	-0.01	-0.00
4.	0.97	-0.08	0.14	0.05	0.10	-0.01	-0.02	0.01	0.04	-0.02	-0.03	0.02	-0.03	-0.01	-0.09	0.11	-0.00	0.01
5.	0.96	-0.03	0.12	0.04	0.09	0.03	-0.02	0.05	0.00	-0.02	0.05	0.02	0.11	0.01	0.19	0.00	-0.06	0.00
6.	0.74	-0.01	0.01	0.10	-0.26	0.15	0.25	-0.17	-0.31	0.28	0.21	-0.20	-0.05	0.01	-0.01	0.01	0.01	0.00
7.	-0.38	0.10	0.70	0.31	-0.11	0.26	0.01	0.27	-0.18	0.00	-0.03	0.03	0.24	0.11	-0.06	0.00	0.00	-0.00
8.	0.31	0.31	-0.58	-0.42	0.10	-0.42	0.01	0.12	-0.01	0.12	0.13	0.07	0.13	0.19	-0.05	0.00	0.00	-0.00
9.	0.08	0.84	0.13	-0.16	-0.09	-0.06	-0.12	0.33	0.03	0.26	0.02	0.05	0.00	-0.21	0.01	-0.00	0.01	0.00
10.	-0.19	0.31	-0.37	0.60	0.42	-0.06	0.20	0.04	-0.15	-0.24	0.25	0.07	0.06	-0.11	-0.02	0.00	0.00	0.00
11.	-0.07	0.32	-0.07	0.67	0.51	-0.04	0.16	-0.01	0.16	0.28	-0.21	-0.07	-0.06	0.09	0.02	-0.00	0.00	0.00
12.	-0.15	-0.13	0.55	0.01	-0.25	-0.52	0.53	-0.07	0.08	0.01	0.00	0.18	-0.01	-0.02	0.03	0.00	0.00	0.00
13.	-0.26	0.15	0.43	-0.52	0.43	0.22	0.23	0.18	0.21	-0.08	0.21	-0.19	-0.10	0.05	0.01	0.00	0.00	0.00
14.	-0.18	-0.79	-0.09	-0.25	0.35	-0.06	0.11	-0.07	0.06	0.15	-0.06	-0.12	0.26	-0.15	-0.04	-0.00	0.00	-0.00
15.	0.19	0.73	0.01	0.07	-0.28	0.20	0.04	-0.39	0.35	-0.05	0.07	-0.03	0.17	-0.01	-0.04	-0.00	-0.00	-0.00
16.	0.95	-0.09	0.16	0.05	0.12	-0.03	-0.07	0.03	0.05	-0.08	-0.01	0.02	0.04	0.01	0.04	-0.04	0.05	0.04
17.	-0.13	-0.71	-0.07	0.40	-0.17	0.16	-0.15	0.14	0.29	0.20	0.28	0.15	-0.04	0.02	-0.01	0.00	0.00	-0.00
18.	-0.30	0.13	0.55	-0.18	0.48	-0.05	-0.31	-0.38	-0.15	0.14	0.12	0.19	-0.02	0.01	-0.00	0.00	-0.00	0.00

TABLE 9.2.6. ORIGINAL CORRELATION MATRIX BETWEEN VARIABLES AND PRINCIPAL COMPONENTS -PRIMARY FACTOR HEADINGS MATRIX.

of varying significance (r value) in different components. That is, the same variables appeared in nearly all the factors, only in different ranked orders. After rotation however, this problem was removed by the fact that in general, every oblique factor was only correlated highly with one variable, and also each variable was only correlated strongly with one oblique factor, except in the higher variance rankings where there were very few correlations of any significance at all.

A further point of interest related to the P.C.A. was the degree of stability of the derived components for the time period 1966 - 1970. An independent P.C.A. was conducted on the original data matrices for each year 1966 - 1970. Yet as can be seen from Table (9.2.7), the derived components for each year were almost identical as regards; the derived eigenvalues, the percentage of total explanation and ranking of interpreted components. The degree of association between the annual interpretation of components was tested using Spearman's Rank Correlation Coefficient, which is given by the equation,

$$r_s = 1 - \frac{6\sum_{i=1}^{n} D_i^2}{N(N^2 - 1)}$$

where,

D_i^2 = the sum of squares of difference between any pair of rankings

N = the number of components = 18

FACTOR NUMBER	YEAR				
	1970	1969	1968	1967	1966
1	INTACT	INTACT	INTACT	INTACT	INTACT
2	SRI	SRI	NSRI	NSRI	NDVI
3	COW	COW	COW	NTCI	NCOW
4	TCI	CI	CI	CCW	NSRI
5	CHSI	SHI	SHI	SHI	CI
6	NDOS	NDOS	CHSI	DOS	NDOS
7	SHI	CHSI	NDOS	CHSI	CHSI
8	NDVI	DEN	DEN	DEN	DEN
9	NETT	DVI	NDVI	NDVI	NSHT
10	CI	NHTI	HTI	NCI	HTI
11	HTI	NTCI	TCI	HTI	NTCI
12	DEN	ETT	NETT	ETT	ETT
13	MSI	AVI	AVI	AVI	AVI
14	NAVI	MSI	MSI	MSI	MSI
15	NAR/TTF	NAR/TTF	NAR/TTF	NAR/TTF	NAR
16	PTF	PTF	PTF	NPTF	PTF
17	TDI	TDI	TDI	POP	NTTF
18	NPOP	NPOP	NPOP	TDI	NPOP
19	PPI	PPI	PPI	PPI	PPI
20	PSI	PSI	PSI	PSI	PSI
21	RI	RI	RI	RI	RI
22	RDI	RDI	RDI	RDI	RDI

INTACT = POP/AR/PTF/TTF/TDI

TABLE 9.2.7. INTERPRETATION OF DERIVED FACTORS 1966-1970.

This correlation coefficient has a range of +1.0 to -1.0, being +1.0 when the rankings are in perfect agreement, and -1.0 if they are in perfect disagreement, and zero if there is no relationship whatsoever. The correlations so calculated are given below and it can readily be seen that the lowest value is $r_s = +0.85$ between 1970 and 1966.

	1970	1969	1968	1967	1966
1970	1.00	0.87	0.88	0.96	0.85
1969		1.00	0.99	0.90	0.89
1968			1.00	0.89	0.88
1967				1.00	0.90
1966					1.00

Although it is only to be expected that the correlation between the two marginal years will be the lowest value calculated, the stability of the relationship can be seen by the fact that if one tests the null hypothesis that there is no linear relationship in the population, i.e. $p = 0$, using an F test such that,

$$F(1, N - 2) = \frac{r^2}{1 - r^2}(N - 2)$$

where F = the tabled F value with 1 and (N - 2) degrees of freedom, at some specified α level.

r = the correlation coefficient.

N = the number of observations.

the test is found to be significant even at the 99.9% level.[2]

The significance of this stability of the various components over the time period 1966 - 70 is important in practical terms as well as in the above mentioned statistical terms. The implications are that, the urban structure, as measured by these 18 variables, has altered very little since 1966, with the same significant "dimensions" being present, and just as significant, at each time period. Once again, therefore, the resistance of the urban structure to rapid and essential change appears to have been highlighted by this P.C.A. In terms of this research however, the static nature of the urban structure has proved very useful, since it could consequently be argued when involved in the temporal analysis, that any variation in the significance of the various independent variables, could not be due to significant changes in the variables themself, but rather to interaction with other factors within the road accident system, or a temporal fluctuation in the significance of that variable itself as regards road accident causation.

The 100 sampled areas can subjectively, be disaggregated into two groups, county borough areas (large burghs in Scotland) and non-county borough areas (small burghs). It can be argued that these two groups are distinct according to the degree of responsib-

---

2 $F(1, 16: \alpha = 0.001) = \underline{16.2}$

$$\frac{r^2}{1 - r^2} (N - 2) = \frac{0.723}{0.277} \cdot 16 = 41.76$$

Therefore since the calculated value is greater than the tabled value the H_0 can be rejected and the hypothesis accepted that $P \neq 0$ at $\alpha = 0.001$ level.

51 COUNTY BOROUGH AREAS.

1.	BARNSLEY
2.	BATH
3.	BIRMINGHAM
4.	BLACKBURN
5.	BLACKPOOL
6.	BOLTON
7.	BOURNEMOUTH
8.	BRADFORD
9.	BRIGHTON
10.	BRISTOL
11.	BURNLEY
12.	CANTERBURY*
13.	CARDIFF
14.	CARLISLE
15.	CHESTER
16.	COVENTRY
17.	DERBY
18.	DEWSBURY
19.	DUNDEE
20.	EDINBURGH
21.	EXETER
22.	GLOUCESTER
23.	GRIMSBY
24.	HASTINGS
25.	HUDDERSFIELD
26.	IPSWICH
27.	KIRKCALDY
28.	LEEDS
29.	LEICESTER
30.	LINCOLN
31.	LIVERPOOL
32.	LUTON
33.	MANCHESTER
34.	NORWICH
35.	NOTTINGHAM
36.	OLDHAM
37.	OXFORD
38.	PLYMOUTH
39.	PORTSMOUTH
40.	PRESTON
41.	READING
42.	ROCHDALE
44.	SALFORD
44.	SHEFFIELD
45.	SOUTHAMPTON
46.	SOUTHEND
47.	STOKE
48.	SUNDERLAND
49.	SWANSEA
50.	WOLVERHAMPTON
51.	YORK

TABLE 9.2.8. LIST OF 51 COUNTY BOROUGH AREAS WITHIN THE SAMPLE 100 AREAS.

ility they hold over their own environment. Thus in non-county boroughs, the authority which is the responsible highway authority is the relevant County Council office. For this reason, therefore, an individual P.C.A. was carried out upon the 51 county borough areas, for each of the years 1966 - 1970, in order to assess any differing response which may arise from the isolation of these areas. A list of the 51 areas used is shown in Table (9.2.8), whilst the basic statistics for the three groups of urban areas are shown in Table (9.2.9). variable by variable for 1970.

Before moving on to study the results of the County Borough P.C.A., it would be relevant to observe Table (9.2.9) in some more detail in order to see where the three groups differ markedly.

The first impressions from this table are the marked differences in the means of the size variables for the County Borough areas (CBs) and the Non-County Borough areas (NCBs). Thus just taking the population variable it can be seen that the difference between the mean values is 163,318.00, which is over three times larger than the original mean value (49,998.73 = $\overline{X}$) for the NCBs.

However, besides this expected variation in the size variables, there are significant differences in some of the other variable statistics, which could explain any different "response rates" for the various groups of urban areas. Using the "t" statistic it was possible to show that there was a significant difference between the $\overline{X}_{CB}$ and $\overline{X}_{NCB}$ for 12 variables at the $\alpha = 0.01$ prob-

VARS.	ALL		CBS		NCBS	
	$\bar{X}$	6	$\bar{X}$	6	$\bar{X}$	6
1. POP *	133199.11	160016.73	213316.73	191821.25	49998.73	25031.84
2. AR*	10810.32	9268.16	14874.10	10644.63	6580.67	4829.91
3. DEN*	11.56	4.58	13.78	4.17	9.25	3.80
4. PTF*	6269.83	3970.16	8655.49	4263.11	3786.80	1112.28
5. TTF*	405194.04	530277.68	653619.57	649785.24	146628.69	79553.56
6. ETT*	21.71	16.69	29.25	17.72	13.87	11.14
7. MST*	2.59	0.16	2.64	0.13	2.54	0.17
8. SRI	1.94	0.05	1.93	0.05	1.94	0.06
9. TCI	0.18	0.06	0.18	0.06	0.19	0.06
10. CI	0.90	0.07	0.89	0.06	0.90	0.07
11. DOS	8.11	8.24	6.69	4.25	9.58	10.81
12. SHI*	9.06	24.28	2.32	0.94	16.09	33.41
13. AVI	0.61	0.06	0.61	0.04	0.61	0.07
14. DVI	0.18	0.03	0.18	0.03	0.18	0.03
15. TDI*	42997.58	50611.27	68756.59	60212.20	16187.18	8231.25
16. HTI	2.51	2.22	2.42	1.46	2.59	2.83
17. CHSI*	11.29	1.71	10.84	1.03	11.75	2.11
18. PPI	20.83	6.08	21.63	6.79	19.99	5.17
19. PSI	0.86	13.16	2.00	15.29	- 0.32	10.51
20. RI	31.07	6.70	31.65	7.23	30.46	6.13
21. RDI*	18.03	6.51	20.72	6.02	15.23	5.84
22. COW*	0.48	0.21	0.42	0.09	0.54	0.13
23. POP	4.93	0.39	5.20	0.32	4.65	0.21
24. SRI	0.29	0.01	0.29	0.01	0.29	0.01
25. CHSI	1.05	0.06	1.03	0.04	1.06	0.07
26. SHI	0.57	0.44	0.33	0.18	0.81	0.49
27. DOS	0.76	0.44	0.76	0.22	0.76	0.59
28. DVI	-0.75	0.07	-0.75	0.06	-0.76	0.08
29. AVI	-0.22	0.04	-0.22	0.03	-0.22	0.05
30. HTI	0.26	0.40	0.33	0.22	0.18	0.52
31. PPI	1.30	0.12	1.32	0.12	1.29	0.12
32. PSI	0.36	0.56	0.92	0.61	0.33	0.51

* INDICATES SIGNIFICANT DIFFERENCE BETWEEN $\bar{X}_{CB}$ and $\bar{X}_{NCB}$
$\alpha = 0.01$ (USING 'T' STATISTIC)

TABLE 9.2.9. THE BASIC STATISTICS FOR THE 22 INDEPENDENT VARIABLES DISAGGREGATED ACCORDING TO TYPE OF URBAN AREA, AND INCLUDING THE TRANSFORMED (LOG10) VARIABLES.

ability level.[3] Of these variables perhaps the more important ones are, Density, Degree of Through Traffic, Mean Social Index and Car Ownership.

The variable with the greatest statistically significant difference between $\bar{X}_{CB}$ and $\bar{X}_{NCB}$ was the density variable. As can be seen from the table $\bar{X}_{CB} = 13.78$ compared to 9.25 for $\bar{X}_{NCB}$, giving a calculated t value of $t_{(98)\propto = 0.01} = 7.12$. Therefore, it can be statistically assumed that non-county boroughs have a much lower density of population than the county borough areas. Now this is of some relevance, since this research has tried to limit its observations and analysis to solely "urban areas," because the difference between rural accident patterns and urban accident patterns has often been emphasised in the relevant literature. For this reason it is necessary to observe the degree of response for this variable between the different groups, when introduced into the regression model, so as not to bias or invalidate the findings of the major regression analyses, when using the full 100 sampled urban areas.

A further variable which exhibits a similar divergence between the two groups is the Degree of Car Ownership variable, ($t_{(98)\propto = 0.01} = 5.40$). From the table it can be seen that the NCBs have a significantly higher level of car ownership than the CBs which would presumably affect the level of traffic and vehicle mileage within these groups. Thus once again this variable has to be considered carefully when introduced into the regression model.

---

3 The significance value for $t_{(98)\propto = 0.01}$ is 2.62.

Another variable which has to be treated with considerable caru is the one of Estimated Through Traffic ($t_{(93)} \alpha = 0.01 = 5.23$). It will be noted that the level of through traffic in the NCBs is considerably lower than in the CB areas. Obviously if the impact of through traffic upon the level of road accidents in any urban area, is as important as hypothesised, then once again this difference between the two groups should be appreciated and the results modified accordingly.

Before leaving this line of argument one further variable which ought to be elucidated further is the Change in Housing Stock Index. Once again this variable is shown to be statistically higher for NCBs ($t_{(98)} \alpha = 0.01 = 2.72$) and this could be important in one or two ways, especially as regards the temporal analysis. This higher value for the NCBs could imply one of two hypotheses,

1) That the NCB urban areas are growing at a more rapid rate than the CB areas.

2) That because of the lower densities shown to exist in these NCB areas, these areas can modify their urban structure to meet present needs, at a much faster rate. That is, the response in these NCB areas as regards their urban structure system, is much higher and quicker in meeting the deteriorating standards of the road traffic and accident systems, than in the CB areas. If this is true one would expect the NCB areas to have a slower road accident growth rate, than the CB areas, if all other variables are held constant. This point will be left until later however,

when the results of the temporal analysis are presented in some more detail.

Before the original data matrix for the CB areas was modified through the P.C.A., three additional variables were added which were only available for these areas. These variables were; an Ethnic Index, Vehicle Registration Index, and the Road Density Index.

1) ETHNIC INDEX

Familiarity with an area is a factor generally accepted as helping to reduce road accidents. Where unfamiliarity exists, and especially where it is exaggerated by different habits and customs, this can often significantly affect the level of road accidents in an urban area. Several studies in the USA have tentatively concluded that as there is a different road accident pattern between the sexes, so there may be between the races, (e.g. Colton et al 1968). Therefore it was thought reasonable to introduce an ethnic index which would attempt to relate the level of the immigrant[4] population within an urban area, to the total population of that area. The data for this index was obtained from the 1966 Sample Census Returns (Migration Tables). The index is defined as,

$$E_i = \frac{I_{mi}}{(P_i/1000)}$$

where, E_i = the ethnic index for area i

4 An immigrant is defined as being any person living in the British Isles who was not born there.

I_{mi} = the immigrant population for area i.

P_i = the total population of area i.

2) VEHICLE REGISTRATION INDEX

This index was introduced simply as a supplement to the traffic variables and the degree of car ownership. The absolute level of the number of motor-vehicles registered for each urban area was the value used. The data was obtained from the annual publication of the Ministry of Transport, "Highway Statistics."

3) ROAD DENSITY INDEX[5]

This index is the same one as described previously in the Data Chapter. (See page 124).

The mean and standard deviation of these variables for the 51 CB areas are given below for 1970.

	$\bar{X}$	σx
EI	33.34	17.17
VRI	52510.00	43022.93
RDI	20.72	6.02

When the original data matrix for the CB areas was subjected to the P.C.A., there were 51 cases and 21 variables. The results of this analysis are shown in a comparable form to those for the full 100 urban areas, in Tables (9.2.10) to (9.2.14).

5 It was possible later to include this variable for all areas following the personal provision of data from the relevant Government Offices.

COMPONENT NUMBER	COMPONENT EIGENVALUE	CUMULATIVE EXPLANATION
1	6.879	32.76
2	3.572	49.76
3	2.534	61.83
4	1.596	69.43
5	1.419	76.19
6	1.033	81.11
7	0.866	85.24
8	0.681	88.48
9	0.506	90.89
10	0.466	93.11
11	0.428	95.14
12	0.330	96.71
13	0.242	97.87
14	0.168	98.67
15	0.115	99.22
16	0.068	99.54
17	0.057	99.81
18	0.020	99.91
19	0.013	99.97
20	0.004	99.99
21	0.003	100.00

TABLE 9.2.10 THE VARIANCE RANKING FOR THE 51 COUNTY BOROUGH AREAS 1970.

	FACTORS																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1.	1.00																				
2.	-0.04	1.00																			
3.	0.08	-0.11	1.00																		
4.	-0.12	-0.13	-0.09	1.00																	
5.	-0.11	-0.24	0.16	0.05	1.00																
6.	0.21	0.06	-0.18	-0.14	-0.27	1.00															
7.	-0.04	0.00	-0.27	0.20	-0.02	-0.06	1.00														
8.	-0.07	-0.34	-0.30	-.06	0.01	-0.03	0.19	1.00													
9.	0.16	-0.07	0.30	-0.46	0.13	-0.10	-0.30	-0.09	1.00												
10.	-0.51	0.10	-0.21	0.17	0.23	-0.32	0.32	0.02	-0.13	1.00											
11.	-0.22	0.62	-0.09	-0.02	-0.23	-0.20	-0.04	-0.27	0.04	0.15	1.00										
12.	0.30	0.34	-0.13	0.01	-0.64	0.09	-0.01	-0.12	-0.12	-0.20	0.32	1.00									
13.	-0.23	-0.31	0.19	0.45	0.40	0.05	0.05	0.16	-0.33	0.08	-0.44	-0.51	1.00								
14.	0.09	0.10	0.08	-0.54	-0.08	-0.02	-0.20	-0.22	0.53	-0.03	0.32	0.18	-0.51	1.00							
15.	-0.49	0.04	-0.33	0.27	0. 24	-0.17	-0.02	0.17	-0.12	0.60	0.22	-0.19	0.11	-0.12	1.00						
16.	-0.03	-0.33	0.03	0.50	0.12	-0.44	0.18	0.33	-0.29	0.10	-0.11	-0.18	0.31	-0.57	0.22	1.00					
17.	0.10	-0.08	0.40	-0.15	-0.18	0.15	-0.17	-0.27	0.01	-0.42	0.00	-0.03	-0.05	0.20	-0.58	-0.18	1.00				
18.	-0.34	0.12	0.08	0.19	-0.02	0.07	0.13	-0.02	-0.63	0.03	0.05	-0.03	0.28	-0.36	-0.08	0.03	0.23	1.00			
19.	0.38	0.07	-0.49	0.04	-0.09	0.17	-0.12	0.11	-0.01	-0.03	0.10	0.15	-0.28	0.13	0.30	0.03	-0.12	-034	1.00		
20.	-0.62	0.01	0.36	-0.00	0.10	-0.30	0.09	-0.19	-0.08	0.22	0.07	-0.30	0.20	-0.06	-0.13	0.00	0.42	0.53	-0.60	1.00	
21.	-0.63	-0.12	0.40	0.12	0.11	-0.24	0.13	-0.17	-0.09	0.33	-0.04	-0.33	0.29	-0.10	-0.08	0.10	0.30	0. 40	-0.58	0.84	1.00

TABLE 9.2.11 COUNTY BOROUGH AREAS FACTOR CORRELATION MATRIX.

FACTORS

VARIABLES	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1.	1.11	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.20	–
2.	0.96	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	-0.27	–	–	–	–
3.	–	–	0.56	–	–	–	–	–	–	–	–	–	–	–	-0.65	–	–	–	–	–	–
4.	1.06	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	-0.21	–	–
5.	0.90	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.32	–	–	–	–
6.	–	–	–	–	–	–	–	–	–	-0.96	–	–	–	–	–	–	–	–	–	–	–
7.	–	–	–	–	–	–	–	–	-0.99	–	–	–	–	–	–	–	–	–	–	–	–
8.	–	–	–	–	–	–	–	–	–	–	–	–	–	0.99	–	–	–	–	–	–	–
9.	–	–	–	–	–	–	–	–	–	–	–	0.54	–	–	–	-0.73	–	–	–	–	–
10.	–	–	–	–	–	–	–	–	–	–	0.94	–	–	–	–	–	–	–	–	–	–
11.	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
12.	–	–	–	–	–	–	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–	–
13.	–	–	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
14.	–	–	–	–	–	–	–	–	–	–	–	0.98	–	–	–	–	–	–	–	–	–
15.	–	–	–	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
16.	1.11	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.22	–
17.	–	–	–	–	–	–	–	–	–	–	–	–	0.98	–	–	–	–	–	–	–	–
18.	–	–	–	–	–	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–	–	–
19.	–	–	1.01	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
20.	1.06	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.27	–	–	–
21.	–	–	–	–	–	0.99	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

TABLE 9.2.12 C.B.S. Oblique Factor Pattern Matrix.

VARIABLES	FACTORS 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1.	0.34	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.07	–
2.	0.29	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–0.14	–	–	–	–
3.	–	–	0.27	–	–	–	–	–	–	–	–	–	–	–	–0.23	–	–	–	–	–	–
4.	0.33	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–0.10	–	–
5.	0.28	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.16	–	–	–	–
6.	–	–	–	–	–	–	–	–	–	–0.54	–	–	–	–	–	–	–	–	–	–	–
7.	–	–	–	–	–	–	–	–	–0.48	–	–	–	–	–	–	–	–	–	–	–	–
8.	–	–	–	–	–	–	–	–	–	–	–	–	–	0.53	–	–	–	–	–	–	–
9.	–	–	–	–	–	–	–	–	–	–	–	0.29	–	–	–	–0.35	–	–	–	–	–
10.	–	–	–	–	–	–	–	–	–	–	0.51	–	–	–	–	–	–	–	–	–	–
11.	–	0.60	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
12.	–	–	–	–	–	–	–	0.72	–	–	–	–	–	–	–	–	–	–	–	–	–
13.	–	–	–	0.63	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
14.	–	–	–	–	–	–	–	–	–	–	–	–0.53	–	–	–	–	–	–	–	–	–
15.	–	–	–	–	0.62	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
16.	0.34	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.07
17.	–	–	–	–	–	–	–	–	–	–	–	0.55	–	–	–	–	–	–	–	–	–
18.	–	–	–	–	–	–	0.76	–	–	–	–	–	–	–	–	–	–	–	–	–	–
19.	–	–	0.48	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
20.	0.33	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
21.	–	–	–	–	–	0.59	–	–	–	–	–	–	–	–	–	–	–	0.14	–	–	–

TABLE 9.2.13 CBS OBLIQUE REFERENCE STRUCTURE MATRIX.

FACTOR NUMBER	YEAR 1970	1969	1968	1967	1966
1	INTACT	INTACT	INTACT	INTACT	INTACT
2	CI	NAVI	DVI	NAVI	NAVI
3	RDI	NCCW	NCOW	NCOW	NCHSI
4	SHI	NRDI	RDI	RDI	NRDI
5	NDVI	CI	NDOS	CI	CI/TCI
6	EI	EI	EI	EI	EI
7	CHSI	DOS	CHSI	CHSI	NSHI
8	DOS	CHSI	CI	DOS	NETT
9	NCOW	ETT	NETT	ETT	NCOW
10	NETT	SHI	SHI	SHI	NDOS
11	TCI	MSI	MSI	MSI	MSI
12	NAVI	HTI	AVI	NDVI	SRI
13	HTI	DVI	HTI	HTI	HTI
14	MSI	SRI	DEN	NSRI	TCI
15	NDEN	TCI	NTCI	NTCI	DVI
16	NSRI	NDEN	SRI	NDEN	NAR/TTF
17	NAR/TTF	NAR/TTF	NAR/TTF	NAR/TTF	DEN
18	VRI	VRI	VRI	VRI	VRI
19	NPTF	NPTF	NPTF	NPTF	PTF
20	POP	TDI	POP	POP	NTTF
21	TDI	NPOP	TDI	TDI	NPOP
22	PPI	PPI	PPI	PPI	PPI
23	PSI	PSI	PSI	PSI	PSI
24	RI	RI	RI	RI	RI

TABLE 9.2.14 INTERPRETATION OF DERIVED FACTORS 1966-1970 FOR COUNTY BOROUGH AREAS.

It will be seen from these tables that the two sets of results are very similar in many details. Once again only six distinct "dimensions" are recognised with an "eigenvalue" of more than one, accounting similarly for approximately 81% of the total variation within the total explanatory variables. One difference that can be noted, however, is that interpretation of the higher order derived components is much more difficult with only the first 9 components having high factor loadings ($r \geqslant \pm 0.60$). Similarly those low order components which can be interpreted produced a different set of variables (Table 9.2.14), although this was only to be expected considering the differences between the original variables discussed above. It will be shown however, that these differences did not seemingly alter the results of the regression analysis, when the derived components in both groups were used as independent variables.

As regards the nature of these derived components over time, using the same technique as employed in the 100 area P.C.A. (Spearman's Rank), it can be seen below, that although the correlations between the years was lower as a whole than for the 100 areas, they were all still significant ($F_{1,19}\ \alpha = 0.001$) at the 99.9% level[6]. This indicates that the static nature noticed in the full sample

	1970	1969	1968	1967	1966
1970	1.00	0.82	0.90	0.86	0.82
1969		1.00	0.87	0.91	0.88
1968			1.00	0.89	0.84
1967				1.00	0.85
1966					1.00

6 The table value for $F_{1,19}\ \alpha = 0.001 = 15.08$. Therefore any r value $\geqslant 0.67$ is significant at the 99.9% level.

selection appears to remain true when only the CB areas are studied. In fact it would seem reasonable to conclude that the results achieved by the two P.C.A. produced results which vary only marginally, although the importance of any such differences will only show itself in the regression analysis, which is described in the following section of this report.

9.3 MULTIPLE REGRESSION ANALYSIS

Using the factor scores and derived components from the two groups of P.C.A., as the input variables in the regression model 120 regression equations were derived. Of these, 90 related to the 45 dependent variables described in the data section of this report for the years 1969 and 1970. The remainder consisted of 19 regressions relating only to County Borough areas (1966/70); 7 regressions for 1970 with transformed dependent variables; and 4 regressions relating to the years 1966 and 1961. This composition of variable regressions is shown in Table (9.3.1).

Because of the magnitude of these regressions it is obviously impossible to describe each result in detail, even though all results are presented in Appendix (3). Therefore, the aim of this section of the report will be to present an overall view of the result of these regression models, and also deal with several of the dependent variables in some more detail, whilst reserving most attention for the 8 variables listed below:-

1) Total Casualties.
2) Total Accidents

REGRESSION Number		REGRESSION DEPENDENT Variables	
1 - 45		45 Basic regressions 1970	
46 - 90		45 " " 1969	
91		LOG. TC70	Transformed Regressions
92		LOG. TA70	
93		LOG. TPED70	
94		LOG. TFS70	
95		LOG. TCVM70	
96		LOG. TAVM70	
97		LOG. TCPP70	
98 - 100	County Borough Areas Only	Total Casualties 1966, 1969, 1970	
101 - 102		Total Accidents 1969, 1970	
103 - 104		Total Pedestrian Casualties 1969, 1970	
105 - 107		Total Casualties per 10^3 Population 1966, 1969, 1970	
108 - 109		Total Accidents Per 10^3 Population 1969, 1970	
110 - 111		Total Ped. Cas. " " " " "	
112 - 114		Total Casualties Per 10^6 Vehicle Miles, 1966, 1969, 1970	
115 - 116		" Accidents per 10^6 Vehicle Miles, 1966, 1969, 1970	
117		Total Casualties Per 10^3 Population 1966	
118		Total Casualties 1966	
119		Total Casualties Per 10^6 Vehicle Miles 1966	
120		Fatal and Serious Casualties 1966	

TABLE 9.3.1. COMPOSITION OF THE 120 REGRESSION EQUATIONS.

3) Total Casualties/000 population

4) Total Accidents/000 population

5) Total Pedestrian Casualties

6) Total Casualties/Vehicle mileage

7) Total Accidents/Vehicle mileage

8) Total Pedestrian Casualties/000 population

As a simple approximation the results of the regression analyses can be studied in two ways:- 1) Significance and level of explanation of each derived regression equation, 2) Analysis of the variables entered into each regression equation at the 10% confidence level ($F_{(\alpha = 0.1)} = 2.79$).

Taking the first analysis method and observing the results of the 45 basic regression equations, it can be seen in Table (9.3.2) that there is a large variation in the level of explanation and "goodness of fit" of individual regressions. These two measures are indicated by the coefficient of multiple determination (R^2), and the estimated standard deviation about the regression expressed as a percentage of the mean response, respectively. This second statistic perhaps needs some further explanations. In any regression calculation the residual mean square S^2 is an estimate of $\sigma^2_{Y.X}$, the variance about the regression. The square root of this value therefore (S) is obviously equivalent to the standard deviation about the regression. Examination of this statistic indicates that the smaller it is, the better; that is, the more precise will be the predictions. This statistic is generally referred to as the "Standard error of Estimate." When, comparing different

Yo	CODING	R^2	$\bar{Y}$	S.D.	S.E.E.	$\bar{Y}$% of S.E.E.
1	TC70	0.986	804.43	1015.21	123.32	15.3
2	TA70	0.987	636.72	808.17	98.06	15.4
3	TFS70	0.960	223.75	281.57	57.57	25.7
4	TPED70	0.988	257.41	369.55	41.69	16.2
5	TMC70	0.901	72.80	84.11	27.48	37.7
6	TPCY70	0.641	86.73	81.03	49.50	57.1
7	TDRC70	0.959	329.04	362.64	76.26	23.2
8	TCPED70	0.967	120.75	177.06	33.01	27.3
9	TYDR70	0.938	166.04	180.16	45.56	28.0
10	TJUN70	0.961	431.16	557.98	115.62	26.8
11	TPDCR70	0.878	18.93	33.81	12.12	64.0
12	TTRN70	0.907	156.42	145.77	45.82	29.3
13	TRT70	0.851	21.06	26.72	10.54	50.0
14	TTJN70	0.950	227.86	269.18	61.82	27.1
15	TYJN70	0.731	14.30	19.90	10.48	73.3
16	TXJN70	0.954	124.10	209.16	46.63	37.6
17	TVCRD70	0.972	290.50	376.20	63.68	21.9
18	TCRD70	0.921	347.63	440.91	129.03	37.1
19	TCFP70	0.292	5.74	1.102	0.95	16.6
20	TAPP70	0.285	4.49	0.879	0.78	17.4
21	TFSPP70	0.162	1.59	0.530	0.49	30.9
22	TCVM70	0.784	18.48	13.140	6.38	34.5
23	TAVM70	0.739	14.47	10.450	5.67	39.2
24	TFSVM70	0.594	5.21	4.410	2.89	55.4
25	SR70	0.059	0.28	0.077	0.08	29.0
26	CPEDCR70	0.569	0.48	0.099	0.07	14.7
27	YDRCR70	0.202	0.51	0.064	0.06	11.9
28	TUNAR70	0.408	0.27	0.075	0.06	22.0
29	PDCRAR70	0.262	0.03	0.040	0.04	142.9
30	TAMUC70	0.414	1.70	0.607	0.47	27.7
31	TAMC70	6.356	7.75	3.220	2.64	34.1
32	RTAR70	0.259	0.05	0.037	0.04	78.4
33	TJNAR70	0.348	0.57	0.101	0.08	13.9
34	YJNAR70	0.131	0.04	0.028	0.03	78.9
35	XJNAR70	0.544	0.23	0.098	0.07	30.8
36	TPEDPP70	0.640	1.67	0.444	0.28	16.8
37	PEDCR70	0.581	0.29	0.068	0.04	13.7
38	PCYWT70	0.407	31.06	22.740	17.97	57.9
39	PEDWT70	0.599	1.92	0.686	0.45	23.5
40	PCYAR70	0.302	0.18	0.115	0.10	55.6
41	MCAR70	0.301	0.12	0.042	0.04	32.5
42	TDRVM70	0.718	8.73	6.770	3.71	42.5
43	DRCR70	0.547	0.45	0.084	0.06	13.4
44	TCTA70	0.423	1.27	0.085	0.07	5.51
45	TJNAR70	0.474	0.65	0.093	0.07	10.8

TABLE 9.3.2. RESULTS OF THE BASIC 45 REGRESSION EQUATIONS 1970.

regression equations the actual value of S is obviously related to the metric, and size of the dependent variable. Therefore in order to have some means of comparing such regression equations, the standard error of estimate is often related to the mean value of the dependent variable, in terms of a percentage value. The lower this percentage value, the better the goodness of fit, and the lower the deviations from the actual regression plane.

From the table (9.3.2) it can be seen that according to the derived R^2 values, there would appear to be two groups of regression equations. The first group involves the absolute casualty figures disaggregated according to location or the mode of the persons involved. Of these 18 regression equations all have an R^2 value greater than 0.50 and only 4 have a value of less than 0.90. Also as is true for all 45 regression equations, all regressions are significant at the 95% confidence level, as indicated by the calculated value of the F test statistic. That is, one can say that these derived regression equations and relationships could not have occurred by chance at the 95% confidence limit. Unfortunately, however, when the deviations from the regression planes are studied it can be seen that several regression equations have very high percentage values, with the smallest value being 15.3% for TC7Ø. The indications are therefore that, whilst this group of regression equations account for a large amount of the total sum of squares, the actual variance about the regression planes is relatively large.

The second group of equations meanwhile has a much more varied set of results. As regards the R^2 values it can be seen that these range between 0.059(SR7Ø), and 0.784(TCVM7Ø) with a predominence of values at the lower levels. It is interesting to note, however, that the exposure dependent variables have most of the high R^2 values. Thus, TCVM7Ø = 0.784, TAVM7Ø = 0.739, TFSVM7Ø = 0.594, PCYWT7Ø = 0.407, PEDWT7Ø = 0.599, and TDRVM7Ø = 0.718. In terms of understanding the road accident situation from these regression models, this previous fact is of very great importance since it means that both of Katz's families of dependent variables show high explanation levels when dealing with structural variables, and therefore the importance of these variables are noted accordingly.

Turning to the deviation of this second group of regression equations, once again the general level (%) is higher than would seem desirable with a minimum value of 5.51% for TCTA7Ø and a maximum value of 142.9% for PDCRAR7Ø. The variation within this range would also appear random, with no seeming relationship between this value and any other statistic related to the regression. It is important to point out meanwhile that the percentage values for the 6 "exposure" regressions all appear fairly high with PCYWT7Ø having a value of 57.9%. Thus because of this high degree of deviation, allied with reasonable levels of explanation, these exposure regressions are deserving of further analysis, which will be completed later in this report.

Within this second group of regressions there are several other dependent variables which produce certain interesting results. For example within this group there is a sub-set of 4 regressions which have as their dependent variable some statistic related to the population level (10^3) of each urban area. These four variables are TCPP7Ø, TAPP7Ø, TFSPP7Ø and TPEDPP7Ø. What is of interest here is that the first three seem to exhibit different tendencies from the fourth regression. Thus whilst the first three exhibit low R^2 values (0.292, 0.285, and 0.162 respectively), the regression explaining the variation within the number of pedestrian casualties per 10^3 population, produces the reasonable R^2 value of 0.640. The implication from this would seem to be that the total number of casualties or accidents in any urban area, are influenced by factors other than those describing that urban structure. Therefore it could be argued that perhaps the number of pedestrian casualties per 10^3 population, is a better indicator of the performance of any urban area than any one of the other three variables. Conversely the other three variables would seem to be good indicators of how any urban area deals with its external factors such as the level of through traffic, and regional commercial attraction.

The second method of generalising about the results of the regression analysis, mentioned previously, is by regarding the variables introduced into the various regressions. The variables included in all 45 regressions (1970) are given in Table (9.3.3). From this table it can be seen that the most important variable is X_1 which was interpreted as an interaction variable, and which appears in 30 out of the 45 regression equations. Although the significance of this variable will be dealt with later when individual regressions are studied, some explanation should be made at this point.

One of the fundamental questions within this research is whether it is possible to reduce road accidents within urban

FACTORS	REGRESSIONS WHERE SIGNIFICANT	TOTAL NUMBER OF REGRESSIONS
INTACT	1: 2: 3: 4: 5: 6: 7: 8: 9: 10: 11: 12: 13: 14: 15: 16: 17: 18: 22: 23: 24: 28: 30: 33: 35: 36: 39: 40: 42: 43	30
SRI	1: 2: 10: 16: 18: 26: 34: 35: 39: 45:	10
COW	4: 5: 7: 9: 19: 36: 37: 38: 41: 42: 43	11
TCI	38: 39: 41:	3
CHSI	20: 22: 27: 28: 31: 38: 39: 40: 41: 43: 44: 45	12
NDOS	26: 29: 32: 35: 39:41: 43: 44: 45:	9
SHI	5: 9: 19: 26: 32: 33: 35: 36: 41: 44:	10
NDVI	1: 2: 3: 10: 12: 26: 32:	7
NETT	3: 10: 19: 20: 26: 27: 29: 35:	8
CI	1: 2: 4: 8: 10: 14: 16: 22: 23: 24: 29: 36: 42:	13
NTI	23: 29: 30: 35: 36: 37:	6
DEN	6: 11: 22: 31: 34: 35: 37: 43: 45:	9
MSI	8: 26:	2
NAVI	23: 27: 34:	3
NAR	1: 2: 4: 5: 7: 8: 9: 10: 13: 14: 15: 16: 17: 18: 22: 25: 38:	17
PTF	3: 4: 5: 7: 9: 10: 11: 12: 14: 18: 19: 20: 22: 23: 24: 30: 33: 36: 39: 40: 42: 44:	22
TDI	1: 2: 4: 5: 6: 7: 8: 9: 10: 11: 12: 14: 16: 18: 29: 39:	16
NPOP	1: 2: 4: 5: 8: 9: 11: 13: 16: 17: 19: 20: 21: 22: 23: 24: 31: 36: 42:	20
PPI	1: 2: 7: 10: 12: 14: 15: 16:	8
PSI	- -	-
RI	6: 13: 17: 21: 27: 28: 30: 32: 33: 37: 38: 40: 43: 44: 45:	15
RDI	22: 23: 24: 31: 36: 42:	6

TABLE 9.3.3. FACTORS INCLUDED WITHIN THE 45 REGRESSION EQUATIONS 1970

areas through the modification of one or more structural variables? That is, is the road accident problem predominantly an engineering problem or predominantly a psychological problem? Although it is obvious that the level of road accidents will increase with the increased size of an urban area, and also the level of interaction within each area, one would not expect variables which describe such conditions to dominate the regression equation, if the road accident problem was predominantly an engineering problem. Therefore whenever such variables do occur in a regression both their statistical significance and the size of their b coefficient should be studied. In a general summary such a study cannot be accurately completed, however it is very significant that the variable X_1 should appear in 30 regressions. Similarly three other variables, which are related to the peak traffic flow (PTF) number of dwellings (TDI) and the population (NPOP) of each urban area, also occur in 22, 16, and 20 respectively, of the regression equations.

Although these four variables appear most often within the regression equations, other variables appear fairly often, and shed further light on the road accident problem. Accordingly these variables will be discussed below.

(1) X_2 : SRI7Ø -

This variable which is related to the relative presence of males and females within an urban area is present within ten of the regression equations. In 8 of these equations the b coefficient is negatively signed, and therefore it would seem that as the proportion of males increases, the response of these dependent variables is reduced. Since this variable is significant within both regressions, TC7Ø and TA7Ø it can be maintained that both total casualties and total accidents will decrease where the proportion of males increases. In some ways, this inverse relationship is somewhat surprising since it is generally considered that females are safer with respect to both driving and walking. However, this seeming discrepancy can perhaps be explained with regard to relative exposure. That is to say, women are more exposed to traffic conflicts during the hours whilst the majority of menfolk are at work, both as regards vehicle and pedestrian conflicts. This could be the result of more and more women having the use of the family car during these hours for both recreation and shopping activities. Similarly the activities of women as pedestrians, especially on shopping expeditions, also exposes them to increased danger. Perhaps a further explanation could be that an excess of males in any population will tend to reduce the number of young inhabitants associated with that area. Since the accidents to persons under the age of 15 years accounted for some 47% of all pedestrian accidents in 1970, this could be a very important feature. Of these two basic possible explanations the first one seems to be a little lacking since if it was true that increased exposure for women leads to more accidents, then one would also expect SRI7Ø to be significant both in total pedestrian casualties and also total driver and rider casualties. Since this is not so, one must consider this explanation with some degree of uncertainty. However, it should be noted here that SRI7Ø does appear significant in PEDWT7∅ which relates pedestrian casualties to the number of pedestrian work trips in each urban area, but even with this qualification, the conclusion reached above would seem to remain valid. The second possible explanation meanwhile seems to be invalidated by the fact that SRI7∅ is positively related to CPEDCR7∅ in that regression. Thus as the relationship between child pedestrian casualties and total pedestrian casualties increases, (i.e. child pedestrian casualties account for a larger proportion of total pedestrian casualties) the SRI7∅ also increases. If the postulated explanation was to have been correct then this relationship would have to be inverted. Consequently all that can be really stated about this variable is that there is an inverse relationship between the sex ratio index and the number of total casualties and total accidents in any urban area. The exact reason for this can only be discerned by further research and study.

(2) X_3: COW7∅

This variable is significant in 11 of the derived regression equations. Its importance seems to lie more in the fact that it disaggregates between the various types of accidents rather than affecting the absolute total number of accidents and/or casualties. Thus, COW7∅ seems to be positively related to those accidents involving driver and rider casualties and motor-cycle casualties, and negatively related to those casualties involving pedestrian and pedal cycle casualties. The implications here are that the general level of accidents can be fairly well predetermined, but the disaggregation of the types of road accidents can only be determined with respect to various internal factors such as COW7Ø.

(3) X_5: CHSI7Ø

Since this variable gives some indication of the amount of change within an urban area, it can also indicate what effect this change has had upon the road accident situation. In some ways it could be argued that it should also point out the impact of modern comprehensive planning. With these thoughts in mind it is interesting to note that the variable is found to be significant in 12 of the derived regression equations, not one of which appears amongst the 18 absolute dependent variables. Therefore once again it would seem that CHSI7Ø has not helped in reducing the total number of casualties and/or accidents, but has determined how they are distributed about the total population.

Of the 12 significant regressions 8 have inverse relationships with CHSI7Ø, including the traffic dependent variables such as, PCYAR7Ø, MCAR7Ø, DRCR7Ø and YDRCR7Ø. Other dependent variables with inverse relationships are also of interest. For example, TCTA7Ø is inversely related, thereby indicating either that the car occupancy rate is lower in these areas, and therefore most accidents involve only one casualty, perhaps one pedestrian. Alternatively because of the improved environment, more people could be travelling on foot, and therefore when involved in an accident, fewer people are involved in each conflict situation. This argument would seem to be supported by the direct relationship between CHSI7Ø and PEDWT7Ø. This relationship would seem to suggest that more pedestrians are exposed to conflict situations in newer or more planned areas. Two other interesting regressions with direct relationships with CHSI7Ø are TCVM7Ø and TAMC7Ø, both of which suggest that the improved road networks in these areas have also tended to increase the number of accidents per degree of exposure.

Summarising therefore, this variable seems to indicate 6 conclusions which may be drawn, as regards these newer, more changed, urban areas:-

1) The CHSI7Ø variable has no effect upon the overall level of road accidents.

2) The CHSI7Ø variable can influence the distribution of road accidents between the type of accidents.

3) Any increase in the CHSI7Ø variable can increase the number of pedestrian casualties when related to degrees of exposure.

4) Where there has been extensive redevelopment the number of casualties per degree of exposure on the roads has also increased, possibly as a result of an increase in the number of pedestrians or alternatively because of a reduction in the mileage of unclassified roads in these areas.

5) Redevelopment has tended to reduce the number of accidents

caused by turning manoeuvres, and those located at road junctions. (See TUNAR7Ø and TJNAR7Ø).

6) Where the value of CHSI7Ø increases the number of casualties related to the number of accidents (i.e. the casualty rate per accident) is reduced.

(4) X_8: NDVI7Ø

Since this factor is negatively related to the variable DVI7Ø all relationships derived from the regression equations (i.e. signs of the b coefficients) have to be inverted to arrive at the correct relationship. For this reason all relationships will be presented here in terms of the original DVI7Ø variable.

The DVI7Ø variable is significant in 7 of the derived regression equations, of which 5 relate to absolute dependent variables, all of which have a positive relationship. Thus the number of people in any urban area, in the age group 15 years - 25 years, directly influences the level of road accidents, casualties etc. in that urban area. It should be noted however, that this variable is not significant amongst any of the exposure dependent variables. The implication here therefore is that the increase in the total number of accidents etc. is not due simply to the absolute number of people in this age group, but more likely to the fact that this group has a higher exposure rate. This seems to vindicate the presumed conclusions of recent research, which has tended to suggest that this age group is no more potentially vulnerable or likely to be involved in road accidents, than any other age group.

The only dependent variable with a negative (inverse) relationship is CPEDCR/7Ø which in many ways is only to be expected. The reason being that as DVI7Ø increases, other age groups must be relatively reduced. Therefore, the absolute number of people within the child age group must also be substantially reduced, and there should be a related reduction in the number of accidents occurring to people in this age group.

(5) X_9: NETT7Ø

Once again since this factor is negatively related to the variable ETT7Ø the arguments and conventions put forward above are once again used in relation to this factor.

Although this variable ETT7Ø is found to be significant in 8 of the derived regression equations, because of the varying direction of the relationships, it proves very difficult to interpret within the road accident framework. However, the most important result is that there is a positive relationship between ETT7Ø and both TCPP7Ø and TAPP7Ø. It would seem from these two regressions that the low levels of explanation obtained with these regressions, ($R^2 = 0.292$ and 0.285 respectively) could be explained by the underestimation of the impact of the amount of through traffic. The absolute number of casualties etc. is related to the situation present within an urban area at any one point in time and with no explicit regard to external factors. When the number of casualties/accidents per thousand of population is regarded however, the impact of external agents, in the guise of through traffic, obviously must play some very important part. If this is so, then the impact of through traffic upon the road accident situation should not be underestimated.

Other relationships between ETT7Ø and the significant dependent variables are very difficult to interpret, with any degree of certainty. For example, there is an inverse relationship between ETT7Ø and TFS7Ø. In other words as the degree of through traffic increases then the total number of fatal and serious casualties decreases. At first such a relationship may appear to be the reverse of what would be expected. However, this result can perhaps be explained by the fact that absolute levels of through traffic are usually associated positively with the population size of each urban area. That is, the amount of through traffic is a function of the nodality of any urban area; the level of trade and regional commercialism is a function of the nodality of any urban area; the growth of an urban area is a function of the commercial and trading capacity of that urban area due to its focal nodality; therefore, the present size of any urban area is a function of the nodality, expressed in terms of the amount of through traffic. Extending this reasoning a stage further the size of an urban area has a positive relationship with the degree of traffic congestion, and therefore traffic speeds. Thus if the absolute number of fatal and serious casualties is a positive function of traffic speed, then it can be seen just how there is an inverse relationship between ETT7Ø and TFS7Ø.

Another interesting feature of this variable is the negative relationship between ETT7Ø and the number of child pedestrian casualties as a proportion of all pedestrian casualties. The most reasonable explanation for this relationship would seem to be that as the volume of through traffic increases further modifications are made to accommodate this unwanted traffic. (Ring roads etc.). As a result unwanted traffic is removed from pedestrian and living areas of the urban environment.

(6) X_{10}: CI7Ø

This compactness index was found to be significant in 13 of the derived regression equations within which all relationships were found to be inverse. That is, as the shape of the urban area became more compact (circular) there was a corresponding decrease in the set of dependent variables. Since this set of dependent variables included 7 of the absolute variables, and TC7Ø, TA7Ø and TPED7Ø in particular, the importance of this variable becomes obvious. In fact of the 22 variables used in this research, CI7Ø was one of the few which has a negative response upon the absolute level of road accidents. More important CI7Ø along with PPI7Ø were the only two variables which can actively be manipulated to reduce the accident level. In the instance of CI7Ø it may prove that the link between urban shape and the level of road accidents is not a direct link. For example it may prove that the link goes through traffic movement and traffic congestion values, since it is usually presumed that circular towns have radial network systems, which thereby tend to increase congestion as one moves towards the urban centre. Whatever the reason however, the fact remains that from the road accident situation, the more compact an urban area the lower the level of road accidents, casualties etc.

Further justification for these conclusions is also given by the other regressions within which CI7Ø is also significant. These include the 4 exposure variables related to the vehicle mileage, (TCVM7Ø, TAVM7Ø, TFSVM7Ø, and TDRVM7Ø) indicating that not only are the absolute levels of road accident conflicts reduced, but also that the level of conflicts related to degree of exposure is also reduced.

The discussion of this variable is very important when one considers the 7 New Town Areas, included in this research. These areas and their calculated CI7Ø values are given below, along

URBAN AREA	CI7Ø	DEVIATION	FACTOR SCORE	DEVIATION
CUMBERNAULD DC	0.876	-0.328	-0.306	-0.31
EAST KILBRIDE LB	0.950	+0.776	0.791	0.79
GLENROTHES DC	0.689	-3.119	-3.103	-3.10
HARLOW	0.896	-0.030	-0.025	-0.03
HEMEL HEMPSTEAD	0.945	+0.701	0.692	0.69
STEVENAGE	0.912	+0.209	0.652	0.65
WELWYN GC	0.970	+1.075	1.046	1.05
100 AREAS MEAN	0.898	-	0.000	
100AREAS SD	0.067	-	1.000	
CUMBERNAULD NT	0.800	-1.463	-1.448	-1.45
EAST KILBRIDE NT	0.858	-0.597	-0.568	-0.57
GLENROTHES NT	0.955	+0.851	0.852	0.85

with the derived factor scores, and their deviations from the overall means. The significance of this table lies in the fact that of these new towns the three which are most recent, Cumbernauld, E. Kilbride and Glenrothes, tend to have a value below the overall mean of the 100 sampled areas. For Cumbernauld this remains true independent of the authority unit used. (Local government or New Town). However, with both East Kilbride and Glenrothes whether the deviation value is positive or negative depends upon the unit used. Thus East Kilbride is negative if the new town area is used and positive if the local authority area is used. As it has been pointed out that the higher the CI7Ø value, the lower the level of road accidents, it would appear that these new towns are not making the optimum use of the shape of the urban area as regards road accidents. It should be remembered however that road safety is only one aspect of new town planning proposals.

(7) X_{11}:HTI7Ø

This variable which measures the relative type of housing tenure between private and public, was found to be significant in 6 of the derived regression equations. The index was so calculated that high values of HTI7Ø indicates a predominance of private tenure and construction, whilst the opposite indicates a predominance of local authority development.

The results of these relevent regression equations once again provided some interesting observations. Of the 6 regressions, 4 proved to have an inverse relationship; these were, TAVM7Ø, TAMUC7Ø, TPEDPP7Ø and PEDCR7Ø. Thus as the predominance of private ownership increased, so the volume and values of these dependent variables decreased. Although this seems to indicate at first glance that private development is "safer" than public developments, this may not be so once certain other factors have been considered.

In the first instance it should be noted that only accident rates are affected, and not casualty rates. Therefore if there has been an increase in the accident rate, there must have been some compensating reduction in the number of casualties per accident, thereby indicating that there had been an increase in one casualty accidents, which is the usual feature of pedestrian accidents. This assumption seems to be supported by the fact that as HTI7Ø increases, there is a reduction in the two pedestrian casualty rates. The conclusions to be drawn here therefore are that HTI7Ø seems to affect road accidents by means of the number of pedestrian casualties, and that public developments have higher pedestrian casualty rates. Meanwhile, it should not be concluded that this relationship is causal in nature. For example, the types of development and ownership contained within HTI7Ø are usually associated with different types of social class, which in turn have different accident response rates, especially amongst child pedestrian casualties.[7] Therefore the relationship between housing tenure and road accident levels may work through the social class variable, which has not been included within the causal link framework. The only way in which this relationship, and the impact of HTI7Ø in general, can be correctly understood is by further detailed local study into statistically comparable areas.

---

7 See X_{13}: MSI7Ø

(8) X_{12}: DEN7Ø

Once again this variable, (which was significant in 9 of the derived regression equations) seems to be one which allocates the number of accidents and casualties, within any urban area, into groups indicating the type of accident, rather than the volume. Thus as DEN7Ø increases, so does the pedestrian casualty rate, with a compensating reduction in the Driver and Rider casualty rate, and pedal cycle accidents. Similarly, DEN7Ø seems to affect the location of accidents, with positive relationships with TAMC7Ø and TJNAR7Ø.

(9) X_{13}: MSI7Ø

Although this variable was significant in only 2 of the derived equations its impact in these was very strong. Both regressions were related to child pedestrian casualties; TCPED7Ø and CPEDCR7Ø.

The MSI7Ø variable was calculated such that as the value increases the social class composition of that area, decreases. In both regressions the derived relationships were positive, thereby indicating that as social class increases, the number of child pedestrian casualties is reduced. The reason for this relationship could be due to many factors such as degree of supervision, access to play areas, including spacious gardens, and as was mentioned previously, type and standard of housing development. Whatever the reason[8] however, this relationship

8 If one studies the two regression equations to find out which other variables are significant, TCPED7Ø reveals very little, whilst CPEDCR7Ø also includes the variables NDOS7Ø,SRI7Ø,SHI7Ø,NDVI7Ø and NETT7Ø, of which the most important appears to be SHI7Ø.

is obviously an integral feature of the overall road accident situation.

(10) X_{19} PPI7Ø

Two variables were included in this research which gave some indication of police activity as regards motor law enforcement, PPI7Ø and PSI7Ø. Of these PPI7Ø was found to be significant in 8 of the derived regressions, whilst PSI7Ø did not appear in any.

The importance of knowing the actual impact of the police enforcement is well illustrated by the following quote which is trying to answer the question, "What can be done?"

"However good our present or future safety legislation may be, it will not save lives if it remains pigeonholed and unenforced. It seems likely that if detection was increased by increasing the police available, if punishments were administered more speedily, and were more effective, so that the fear of them became a deterrent to traffic violators, then even with our present legislation, about 30% of the present road deaths and casualties could be prevented. This would mean that about 2,400 lives and nearly £90 million a year could be saved." (Cohen and Preston 1968).

Without substantiation however, such statements, of which many abound, can have little meaning. Consequently, the fact that PPI7Ø has an inverse relationship with all the dependent variables, where significant, is of the utmost importance. Similarly, since the list of significant regressions include both

TC7Ø and TA7Ø it would seem that perhaps the savings in both life and money mentioned by Preston (above) may be totally substantiated. Thus it would appear that the results of this research, and the national casualty trends since the introduction of the 1967 Road Safety Act[9] (Legislation on Drink and Driving) both indicate the advantages to be attained by increased police activity within the field of road safety. The problem however, is that with the continued shortage of manpower within the police forces, what is desirable and what is practical, are two entirely different features. Perhaps the only solution here is to increase the power of some subsidiary force, such as the present traffic wardens, but even this would be only satisfactory if this new body commanded the same influence, both on the streets and in the courts, as the police do at present.

(11) X_{21}: RI7Ø

One of the major conclusions from the work by Recht (1965) was that "the major portion of the differences in accident experience is associated with differences in the non-program factors....particularly vehicle density, kind of travel, age of population and weather." (Recht op.cit). In this research this rainfall index was representative of this last group of non-program factors. Recht used as his rainfall index, the amount of precipitation in inches, and derived a strong negative relationship

9 According to the information produced by the Department of the Environment in Road Accidents, 1969, following the 1967 Road Safety Act, the level of total casualties was down approximately 10% on pre-Act level despite a 4% rise in the volume of motor traffic.

between this variable and his "Total death Ratio." However, in the present research of the 15 significant regressions, 9 had positive relationships, and the only absolute dependent variables included were TPCY7Ø and TUCRD7Ø. Basically the results from this variable were all totally inconclusive with only some brief generalisations being possible. This could be due to the fact that the variation noted within this variable is mainly due to spatial variation and should therefore be really studied in context, with the overall spatial variation in the volume of road accidents, casualties etc. However, what tentative conclusions that can be made, are given in the following:-

(i) The volume of pedal cycle accidents is positively related to the amount of rainfall in any one area. That is where there is a higher rainfall incidence both TPCY7Ø and PCYAR7Ø increase proportionally.

(ii) There is a positive increase in the volume of road accidents occurring on unclassified roads relative to a rise in the amount of precipitation. This could be indicative of the relative standard of surfacing on these roads, or alternatively due to the reduction in visibility in these less well lit roads.

(iii) There would seem to be a positive relationship between the level of rainfall and the number of fatal and serious accidents occurring within a given volume of population, (TFSPP7Ø).

(iv) Since there is a positive relationship between RI7Ø and YDRCR7Ø it would seem to suggest that in inclement weather conditions the inexperience of some young drivers and riders can account for the variation in road accidents in any given urban area. One further variable which could perhaps be relevent here is the attendent speed in such accidents.

(v) Finally the incidence of all driver and rider casualties as a proportion of all casualties is positively related to the level of precipitation.

(12) X_{22}: RDI70

This varieble which describes the road density in any area and which was pointed out earlier to be highly correlated with junction density was found significant in 6 of the regression equations. Of these 4 related to the exposure variables, TCVM70, TAVM70, TFSVM70 and TDRVM70. However, it is interesting to note that RDI70 was not significant in any of the junction dependent variables, which suggests that whilst road density increases the relative accident/casualty response, this increase does not occur at road junctions themselves. Perhaps this could be explained by increased psychological frustration due to the slower progress rates caused by greater junction "stopping and starting."

Of the remaining variables which have not been explicitly examined here, either no distinct relationship can be found or alternatively they appear to be related to the impact of other variables which are more significant or powerful. For example, SHI70 would appear to be highly similar in influence to CHSI70 and NAVI70 to NDVI70. For these results one is referred back to Table (9.3.3) and Appendix 3.

Having given so far a generalised account of the results of 45 (1970) regression equations it now becomes necessary to look at some of the regression equations in more detail. This will be completed commencing with the major 8 regressions listed previously.[10]

1). TOTAL CASUALTIES 1970 (TC7Ø)

During the discussion on the results of the principal components analysis, reference was made to the "Variance Ranking" which listed the components according to their variances in declining order (Table 9.2.1). A comparable ranking system is the "Regression Ranking." For each dependent variable there is a possibility of three different models, the composition of which is determined by the three significance levels (F values) used as entry limits for the independent variables. The three F levels for inclusion in the regression model are 0.1, 2.79 and 4.00 respectively, and the three models derived for TC7Ø are shown in Table (9.3.4).

The most obvious result from this table is the difference between the two ranking systems, thus showing that whilst a component may account for a large portion of the variance in the original data matrix, it does not have to account for a significant variable when attempting to explain the variation in some dependent variable. Thus in the regression ranking for TC7Ø only one of the first five variables has a variance rank of less than 10 and that is the first variable itself.

10 - See page 201.

Table 9.3.4 Regression Ranking for TC70.

MODEL ONE			MODEL TWO			MODEL THREE		
VARIANCE RANK NO.	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION
1	0.98	96.65	1	0.98	96.65	1	0.98	96.65
18	0.49	97.77	18	0.49	97.77	18	0.49	97.77
19	-0.18	98.04	19	-0.18	98.04	19	-0.18	98.04
10	-0.07	98.20	10	-0.07	98.20	10	-0.07	98.20
17	-0.58	98.39	17	-0.58	98.39	17	-0.58	98.39
15	-0.04	98.55	15	-0.04	98.55	15	-0.04	98.55
2	0.03	98.59	2	0.03	98.59			
8	-0.11	98.64	8	-0.11	98.64			
16	-0.12	98.67						
7	0.06	98.69						
14	0.11	98.70						
9	-0.57	98.71						
20	-0.08	98.72						
5	-0.17	98.73						
22	-0.29	98.73						
12	0.51	98.75						
11	-0.11	98.76						
3	-0.23	98.76						
6	0.08	98.77						
13	0.17	98.77						
TOTAL R^2 VALUE		98.77	TOTAL R^2 VALUE		98.64	TOTAL R^2 VALUE		98.55

When looking at the regression equation and attempting to explain them, the second model is the one used, since this includes variables, which are significant at the 10% confidence level. In the instance of TC7∅ this model included 8 independent variables and is given by the equation;

$$Y_i = 779.1 + 1011.3X_1 - 36.2X_2 - 26.67X_8 - 44.02X_{10} - 45.62X_{15}$$

$$- 92.94X_{17} - 157.9X_{18} - 38.4X_{19} \qquad (1970)$$

This model accounts for 99% of the variation in the dependent variable ($r^2 = 0.986$), with a Standard Error of Estimate (S.E.E.) expressed as a percentage of the mean response of 15.3%

The importance of the individual variables in the regression equation can be assessed in two different ways. In the first instance one can observe the increase in the regression percentage explanation associated with each independent variable. Secondly, since the independent variables are standardised with a mean of zero and a standard deviation of one, the b coefficients can be simply related to the beta (β) coefficients and therefore compared directly according to their magnitude. Thus the larger the coefficient the greater that variable's influence upon the dependent variable.

A further piece of information which is important in each regression is the statistical significance of the value of the b coefficient. Although this information is implicit in the step-wise regression procedure, an independent assessment is of great practical value.

Using the Student's t test it is possible to test the null hypothesis (H_o) that $b_i = b_{io}$, where b_{io} is a specified value which can be Ø or any other value, against the test hypothesis (H_1) that b_i is different from b_{io}. ($H_o : b_i = b_{io}$ versus $H_1 : b_i \neq b_{io}$). If H_o has to be accepted then it must be assumed that the b value could have occurred by chance, and the independent variable may not significantly affect the dependent variable. The test statistic is given by,

$$t = \frac{(b_i - b_{io})}{\text{est.se}\ (b_i)}$$

$$= \frac{(b_i - b_{io}) \left\{ \Sigma(X_i - \overline{X})^2 \right\}^{\frac{1}{2}}}{S}$$

where,

S = Standard Error of Estimate and the confidence interval is given by $t(n - 2; 1 - \frac{1}{2}\alpha)$

The calculated t values are also given in Table (9.3.4).

As regards the regression equation TC7Ø, it becomes obvious that the entire model is dominated by the one variable X_1: INTACT 7Ø. This one variable accounts for 96.7% of the variation in TC7Ø and has a β value of +1011.3. That is, since these are partial regression coefficients, a change of 1 in the variable X_1 will increase the value of the dependent variable by 1011.3, if the remaining independent variables are held constant. The only other variable which increases the percentage explanation by more than one per cent is X_{18}: NPOP7Ø (1.12%).

This dominance of INTACT7Ø is in fact a feature of all the regression equations dealing with absolute dependent variables, and can be interpreted in one of two ways. In the first instance it could simply be stating that as urban areas increase in size, there is a corresponding increase in the number of road accidents. Alternatively it could be argued that the level of road accidents is determined less by the urban structure of an area but more by the actual level of interaction, within that area. In other words it would seem to be suggesting that the road accident problem is a psychological one rather than an engineering one. If this is true then it would seem to support the conclusions of Cohen (1963), that there is a need for a better understanding of the human driver and his psychological attitudes. This is further supported by the significance of PPI7Ø.

Although the variation in TC7Ø was highly accounted for, the actual variance about the regression plane was relatively high as given by the S.E.E. Therefore in order to look further into this problem three additional observations were made.

i) The residuals from the predicted regression plane were plotted both against the dependent variable, and each of the independent variables, in order to search for any non-random irregularities.

ii) The variable TC7Ø was regressed against the independent variables relating only to the County Borough areas (51).

iii) The dependent variable (TC7Ø) was transformed where necessary and regressed once again against the independent variables.

The plotting of the residuals against the total variables used in the regression is a simple method of observing anomalies, and discovering unsuitability within the regression model used. The diagrams in Fig. (9.3.5) show three of the more common abnormalities revealed by residual plots. The interpretation of the abnormalities is dependent upon the variables being used in the plotting. However, when plotting against the independent variables the three anomalies shown in Fig. (9.3.5) can be interpreted as follows:-

1) Variance not constant; need for weighted least squares or a preliminary transformation on the Y's.

2) Error in calculations; linear effect of X_j not removed.

3) Need for extra terms, for example, a quadratic term in X_j in the model or a transformation on the Y's.

When the residuals are plotted against the Y_i meanwhile the interpretation for the three same anomalies is slightly different.

1) Variance not constant as assumed; need for weighted least squares or a transformation on the observations Y_i before making a regression analysis.

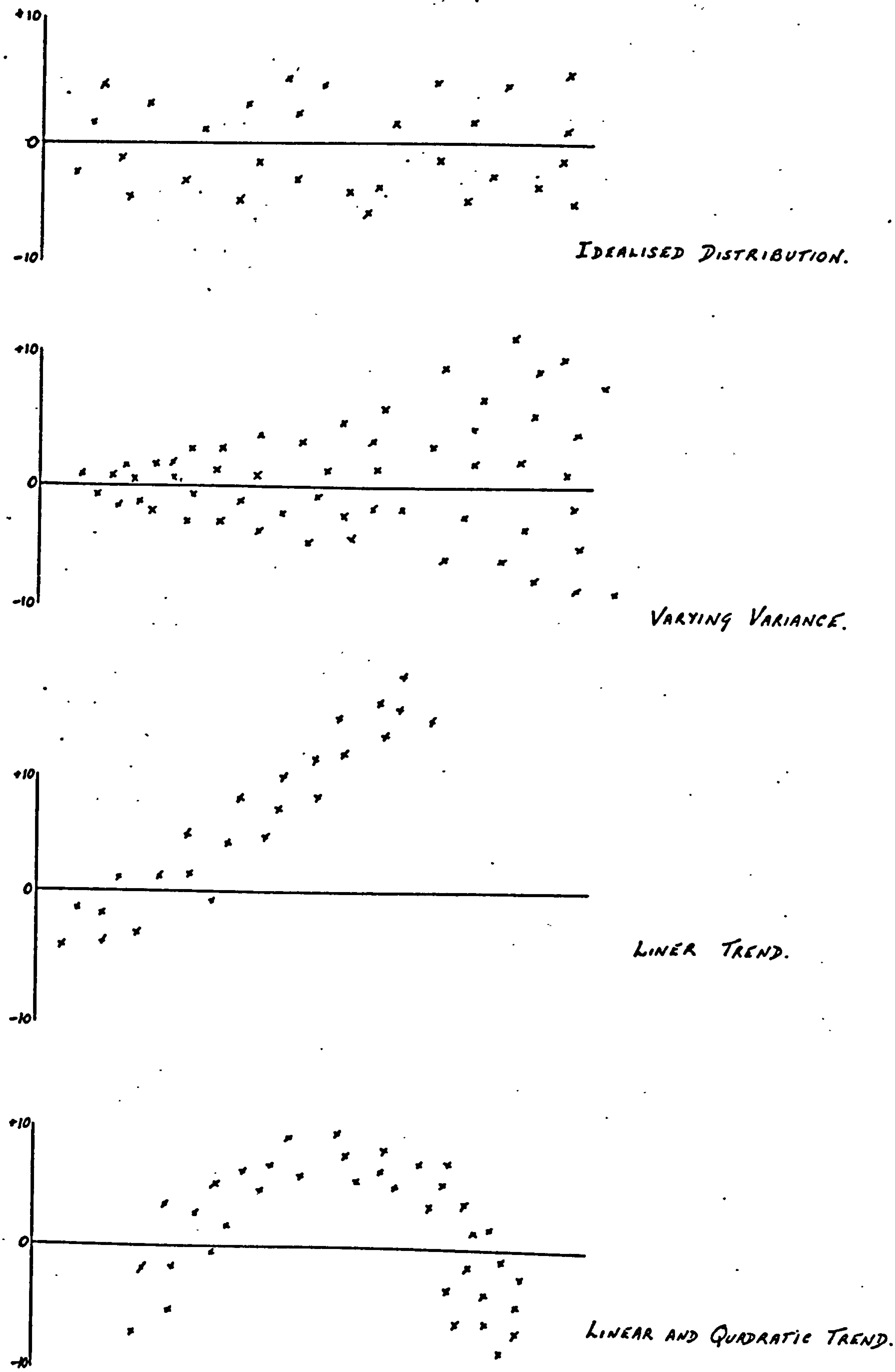


Fig. 9.3.5. THREE POSSIBLE RESIDUAL GRAPH PLOT ABNORMALITIES.

2) Error in analysis; the departure from the fitted equation is systematic. Negative residuals correspond to low Y_s and positive residuals to high Y_s. The effect can also be caused by omitting the β_o term in the model.

3) Model inadequate - need for extra terms in the model (e.g. square or cross-product terms) or need for a transformation on the observations Y_i before analysis.
(After Draper et al 1966).

In most of the regressions studied there were very few which proved to have abnormal plots, and therefore for this reason, these residual plots will only be mentioned where relevant in the following accounts.

As regards TC7Ø the only plot which appears anything like abnormal is the one between the residuals and X_9: NETT. (Fig. 9.3.6). Remembering that this factor is negatively related to the amount of through traffic variable, it can be seen from this graph that the regression equation is overestimating the number of casualties when the volume of through traffic is high, and underestimating the number of casualties when the volume of through traffic is at, and around its mean value. The indication here once again is that the actual impact of this variable is still not understood sufficiently nor modelled correctly. It would seem therefore that the relationship between TC7Ø and ETT7Ø is either logarithmic or some polynomial function, rather then the linear function postulated in this model. As a result, the influence of the amount of through traffic, which is not

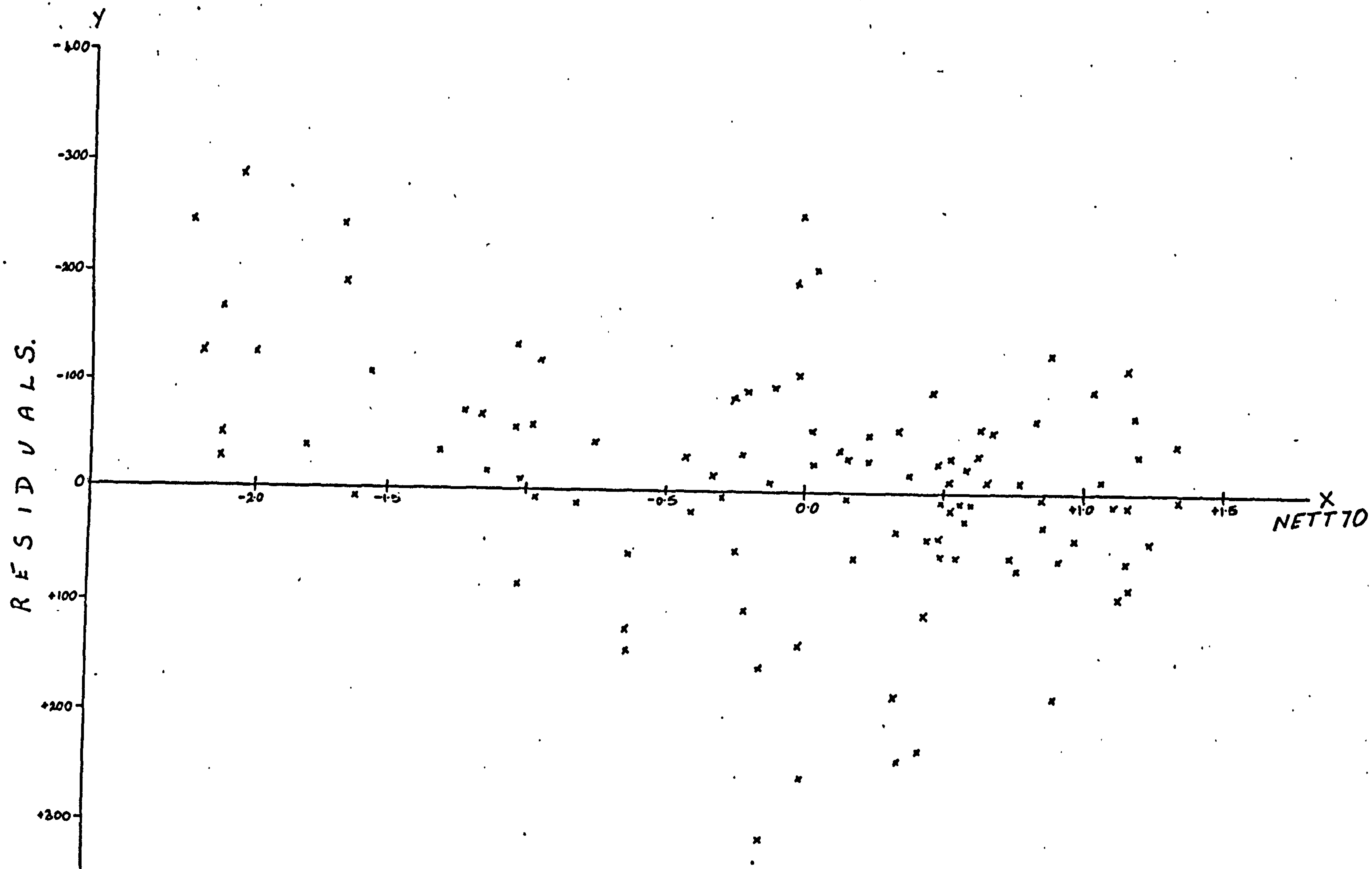


Fig. 9.3.6. PLOT OF RESIDUALS FROM TC70 AGAINST NETT.

significant in TC7Ø, could have been seriously underestimated, as has been suggested from other results in this research. Consequently the dependent variable TC7Ø was logarithmically transformed and a new regression equation derived.

The results of this new regression equation are given in Appendix (3) and the variables found to be significant are shown in Table (9.3.7). Although there was a similar number of variables in the new equation, the actual variables differed, and the new equation was as below:

$$\text{Log } Y_1 = 2.67 + 0.32X_1 + 0.02X_3 - 0.02X_4 - 0.05X_7 - 0.02X_9 - 0.02X_{11} + 0.23X_{16} + 0.10X_{18}$$

From this it can be seen that whilst all the size and traffic variables are positively related to LNTC7Ø, there is a degree of limitation imposed by the inverse relationships between LNTC7Ø and the three significant structural variables, TCI7Ø, SHI7Ø and HTI7Ø. Thus in this form the dependent variable LNTC7Ø would appear to be influenced by both human and structural variables, with the latter being able to reduce or restrict the level of log. road casualties. However, it should be also noted that only three variables substantially increase the percentage level of explanation, and all of these are size or interaction variables. (Table 9.3.7).

A further feature of this regression equation for LNTC7Ø was that whilst the variation about the regression plane was

FACTORS		LOG.TC	TC	LOG TC/VM	TC/VM	LOG TA	TA	LOG PED.	PED	LOG FS	FS
INTACT	1.	x	x	x	x	x	x	x x	x	x	x
SR1	2.		x				x				
COW	3.	x						x	x		
TC1	4.	x						x			
CHS1	5.			x	x	x					
NDOS	6.										
SH1	7.	x						x			
NDV1	8.		x			x	x				x
NETT	9.	x									x
C1	10.		x	x	x		x		x		
HTI	11.	x						x			
DEN	12.			x	x						
MSI	13.										
NAV1	14.										
NAR	15.		x	x	x		x	x	x		
PTF.	16	x		x	x	x		x	x	x	x
TDI	17.		x				x		x		
NPOP	18.	x	x	x	x	x	x	x	x	x	
PPI	19.		x	x			x				
PSI	20.							x			
R1	21.					x				x	
RD1	22			x	x			x			
R^2		0.963	0.986	0.937	0.784	0.950	0.987	0.965	0.988	0.907	0.960
$\bar{Y}$%		3.2%	15.3%	7.8%	34.5%	3.9%	15.4%	4.3%	16.2%	6.59%	25.73%

TABLE 9.3.7. SIGNIFICANT FACTORS WITHIN REGRESSIONS AND THEIR COMPARABLE TRANSFORMED VARIABLES.

reduced (S.E.E.% = 3.2%), the overall regression explanation level was also reduced to $r^2 = 0.963$. Thus whilst certain features have been improved by this transformation, other features have been restricted and it might well be asked what is the advantage of a further complication of the regression model through the addition of a logarithmic transformation, when the overall explanation level is reduced. The only real justification is if the model is wanted for accurate predictive purposes. In this research this is not so, although in terms of future research such models could be useful.

When the dependent variable TC7Ø was regressed against the independent variables relating only to the 51 county boroughs areas, the two sets of results proved very similar, and the derived regression equation was given as:

$$Y_{CB_i} = 1306.7 + 1248.4X_1 - 87.6X_2 + 19.20X_{17} - 48.0X_{19} + 244.1X_{20} - 124.5X_{21} - 55.7X_{22}$$ [11]

The comparable significant variables are given in Table (9.3.8).

The removal of the non-county borough areas, which was to see if there was a disaggregated dichotomous response between the list of sampled areas, had the effect of maintaining the percentage explanation level ($r^2 = 0.987$) whilst at the same time reducing the S.E.E.% response (11.6%). Thus it can be concluded that the non-county borough areas whilst having a similar casualty pattern as county borough areas, also have a wider variation in their deviations from the regression plane. In some way this is only

11 - The X_1 variables relate to those for County Borough areas only, given in Table (9.2.14).

FACTORS		TC		TA		E PED		TC/000		TA/000		PED/000		TC/Vm		TA/Vm	
		ALL	CBS	ALL	CBS	ALL	CBS	ALL	CBS	ALL	CBS	ALL	CBS	ALL	CBS	ALL	CBS
INTACT	1.	x	x	x	x	x	x					x		x	x	x	x
SRI	2.	x		x											x		x
COW	3.					x		x	x			x					
TC1	4.																x
CHSI	5.									x				x			
NDOS	6.																
SHI	7.						x	x				x					
NDVI	8.	x		x													x
NETT	9.							x	x	x	x						
C1	10.	x	x	x	x	x	x		x		x	x	x	x		x	
HTI	11.											x	x			x	
DEN	12.						x							x			
MSI	13.																
NAVI	14.															x	x
NAR	15.	x	x	x	x	x	x							x			
PTF	16.					x		x		x		x		x	x	x	x
TDI	17.	x	x	x	x	x											
NPOP	18.	x	x	x	x	x		x		x		x		x	x	x	x
PPI	19.	x	x	x	xx										x		x
PSI	20.																
RI	21.																
RDI	22.		x									x	x	x	x	x	x

TABLE 9.3.8. SIGNIFICANT FACTORS WITHIN REGRESSIONS AND THEIR COMPARABLE COUNTY BOROUGH REGRESSIONS.

to be expected from the varying nature of these non-county borough areas, as well as their greater size range.

Summarising therefore, TC7Ø seems to be mainly explained by the level of interaction within any urban area, as well as the actual size of that urban area. Similarly although the introduction of the transformed variable LNTC7Ø reduces the variation about the regression plane, and therefore enables more accurate prediction, it also reduces the overall level of explanation.

Finally when the county-borough areas only are considered very little difference is noted in both the regression equation variables and also the level of explanation, although the variation about the regression plane is reduced slightly. Consequently the relationships noted above would appear to remain consistent for all types of area, and therefore practically efficient for spatial comparisons.

2) TOTAL ACCIDENTS (1970) TA7Ø

Most of the results obtained for this dependent variable are directly comparable to those obtained for TC7Ø and therefore discussion of this variable is somewhat limited.

The regression ranking for TA7Ø is given in Table (9.3.9) and the eventual derived regression equation was as below:

$$Y_1 = 616.8 + 812.2X_1 - 31.7X_2 - 26.9X_8 - 32.6X_{10} + 34.6X_{15}$$
$$-68.5X_{17} - 135.1X_{18} - 31.3X_{19}.$$

Table 9.3.9 Regression Ranking for TA70.

MODEL ONE			MODEL TWO			MODEL THREE		
VARIANCE RANK NO.	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION
1	0.98	96.41	1	0.98	96.41	1	0.98	96.41
18	0.48	97.80	18	0.48	97.80	18	0.48	97.80
19	-0.18	98.08	19	-0.18	98.08	19	-0.18	98.08
15	-0.03	98.23	15	-0.03	98.23	15	-0.03	98.23
17	-0.57	98.37	17	-0.57	98.37	17	-0.57	98.37
10	-0.07	98.52	10	-0.07	98.52	10	-0.07	98.52
16	-0.11	98.57	16	-0.11	98.57			
2	0.03	98.61	2	0.03	98.61			
8	-0.12	98.68	8	-0.12	98.68			
7	-0.06	98.70						
22	-0.30	98.71						
12	0.52	98.73						
3	-0.24	98.74						
11	-0.10	98.76						
6	-0.09	98.77						
20	-0.07	98.78						
9	-0.57	98.78						
4	-0.21	98.79						
TOTAL R^2 VALUE		98.79	TOTAL R^2 VALUE		98.68	TOTAL R^2 VALUE		98.52

The equation accounted for over 98% of the total variation within the dependent variable and the deviation about the regression plane was approximately 15% (S.E.E. = 15.4%). Once again as was the case with TC7Ø, over 96% of the overall explanation level is accounted for by the single variable X_1:INTACT7Ø, with only the one other variable X_{18}: NPOP7Ø accounting for more than 1% of the explanation.

Since a similar graph plot was noted for the residuals from TA7Ø, and NETT7Ø, the dependent variable was once again transformed logarithmically. The regression equation thus derived is given as,

$$\text{Log } Y_1 = 2.56 + 0.27X_1 - 0.03X_5 + 0.04X_8 + 0.22X_{16} + 0.13X_{18} + 0.04X_{21}$$

and once again the r^2 value ($r^2 = 0.950$) was reduced at the same time as the regression deviation (S.E.E. = 3.9%). However, there is one basic difference between these results and that for LNTC7Ø in that in this equation there are none of the structural variables which had been present in the LNTC7Ø regression, and once again the emphasis is retained with the psychological and human aspects of the road accident situation. This comparison of significant variables is shown in Table (9.3.7).

When the original dependent variable was regressed against the County Borough areas (C.B.S.) the situation was again shown to be static over space, with the explanation level remaining almost constant ($r^2 = 0.985$) and the regression deviation being slightly reduced (S.E.E. = 12.1%). The derived CBS regression equation is the one given below, and it can be seen from this and Table (9.3.8) the similarity between the sets of significant variables for both of

$$Y_1 = 1040.8 + 1009.2X_1 - 65.8X_2 + 54.9X_{17} + 37.0X_{18}$$
$$+ 206.4X_{20} - 79.7X_{21} - 45.9X_{22}.$$

the dependent variables TC7Ø and TA7Ø.

3) TOTAL CASUALTIES PER 10^3 POPULATION 1970 (TCPP7Ø)

It was indicated in the discussion upon the regression equation derived for the dependent variable TC7Ø, that the dominance of the independent variable INTACT7Ø in explaining the variation in TC7Ø could simply be related to the fact that the volume of road casualties increases as does the size of the urban area being studied. If this assumption was correct then one would also expect TCPP7Ø to be virtually constant for all the urban areas studied. Yet the mean value for 1970 was 5.744 with a standard deviation of 1.102. The relevant question therefore is once the population factor has been held constant (admittedly in a linear manner) what other factors account for the remaining variation. Theoretically it would appear that whatever factors are found to be significant, would be the structural ones which can perhaps be modified so as to allow beneficial changes to take place. Unfortunately the derived regression equation for TCPP7Ø was found to account for only approximately 30% of the

Table 9.3.10 Regression Ranking for TCPP70.

MODEL ONE			MODEL TWO			MODEL THREE		
VARIANCE RANK NO.	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION
9(R)	-0.36	12.90	9	-0.36	12.90	9	-0.36	12.90
7	-0.32	22.20	7	-0.32	22.20	7	-0.32	22.20
18	0.33	24.72	18	0.33	24.72			
16	0.05	27.10	16	0.05	27.10			
3	0.06	29.24	3	0.06	29.24			
10(R)	-0.09	30.94						
11	-0.04	32.56						
1(R)	0.27	33.39						
2	-0.03	34.04						
14	0.24	35.37						
17	-0.33	36.42						
4	-0.16	37.44						
13	0.00	39.07						
6	-0.07	40.24						
19	-0.20	40.72						
21	0.17	41.10						
20	0.03	41.59						
15(R)	-0.15	41.80						
5	-0.17	41.98						
12	0.10	42.10						
1	0.27	42.44						
22	0.05	42.44						
TOTAL R^2 VALUE		42.44	TOTAL R^2 VALUE		29.24	TOTAL R^2 VALUE		22.20

variation in the dependent variable ($r^2 = 0.292$), whilst the regression variance was also relatively high. (S.E.E. = 16.6%). However, the regression as a whole with an F value of 7.77 and 5 and 94 degrees of freedom was significant even at the 0.001 probability level. The derived regression was given as,

$$Y_i = 5.68 + 0.167X_3 - 0.416X_7 - 0.292X_9 + 0.222X_{16} + 0.279X_{18}.$$

Comparing this equation with the regression ranking for TCPP7Ø given in Table (9.3.10) it can be seen that the two most important variables are X_7 : SHI7Ø and X_9 : NETT7Ø, with X_7 accounting for almost 9% of the total variation and X_9 almost 13%. The importance of the level of through traffic would once again appear to have been demonstrated, whilst SHI7Ø would also seem to indicate the advantages to be obtained from modernisation and perhaps redevelopment. It should however, be remembered that only 30% of the variation in the dependent variable has been accounted for by this regression.

A study of the residual graph plots complicates the matter further since as can be seen in Fig. (9.3.11), the plot between e_i and TCPP7Ø reveals a systematic error. The derived regression equation is obviously over-predicting the lower casualty rates and under-predicting higher rates. Consequently the dependent variable was given a logarithmic transformation to see if this error could be removed and whether the overall explanation level increased. The results along with the significant variables are given in Table (9.3.7). From this it can be seen that there is still virtually

Fig. 9.3.11. PLOT OF TCPP70 AGAINST THE REGRESSION RESIDUALS

no change in the explanation level with $r^2 = 0.300$ and the regression equation being defined as below,

$$\text{Log } Y_i = 0.75 - 0.02X_4 - 0.04X_7 - 0.03X_9 + 0.01X_{14}$$

Similarly since the systematic error still remains, one must conclude that the structural variables on their own still do not explain this variation and in the context of the road accident system one must look at the other sub-systems in order to explain this variation. Once again the sub-system which would seem to be the most relevant in this context would be the human sub-system and the different psychological response related to different levels of congestion and interaction, with stress etc. being much higher in those more congested (larger) urban areas.

However, in such a situation one would expect that the level of explanation associated solely with CBS areas would be higher. Unfortunately the derived regression equation for these areas given below, whilst having a lower variation about the

$$Y_{CBi} = 6.00 - 0.225X_2 - 0.204X_9 - 0.251X_{10}$$

regression plane (S.E.E. = 12.1%) does not have a better explanation level with $r^2 = 0.231$. Meanwhile it is relevant that the X_0 value for this regression is considerably higher than for all the 100 sampled areas: 6.00 compared to 5.68.

The conclusion that must be drawn therefore concerning this regression equation is that given the present level of information we are unable to account for the observed variation with any certainty, but it is postulated that perhaps the solution could be found within the human psychological sub-system.

4) TOTAL ACCIDENTS PER 10^3 POPULATION 1970 (TAPP7Ø)

Once again the results for this regression are very similar to the equivalent casualty variable described above, and consequently the arguments put forward for that variable (TCPP7Ø) are also applicable within this regression.

The explanation level for TAPP7Ø was once again found to be very low ($r^2 = 0.285$) and with a high regression variance. (S.E.E. = 17.46). The derived regression equation is given as below.

$$Y_1 = 4.47 - 0.20X_5 - 0.22X_9 + 0.21X_{16} + 0.23X_{18}$$

If this equation is compared with the equivalent regression for TCPP7Ø it will be seen that the only difference in variable inclusion is that this regression (TAPP7Ø) has replaced variable X_7: SHI7Ø, with a variable X_5: CHSI7Ø. However, as can be seen from the regression ranking given in Table (9.3.11) the matter is further confused by the fact that it would appear that these two variables are interchangeable. As a result the inclusion of one seems to lead to the exclusion of the other and vice-versa.

Due to the similarity between TAPP7Ø and TCPP7Ø it was deemed unnecessary to test this regression with any transformation, even though the systematic error noted above was also present in the TAPP7Ø graph plots. There was however, a comparison with

Table 9.3.11 Regression Ranking for TAPP70.

MODEL ONE			MODEL TWO			MODEL THREE		
VARIANCE RANK NO.	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION
9(R)	-0.42	17.57	9	-0.42	17.57	9	-0.42	17.57
5	-0.34	22.58	5(R)	-0.34	22.58	5	-0.34	22.58
7	-0.23	24.85	7(R)	-0.23	24.85			
16	0.18	27.30	16	0.18	27.30			
18	0.25	30.44	18	0.25	28.11			
8	-0.18	31.70	5	-0.34	30.44			
10	-0.12	33.27						
2	-0.08	34.54						
17	-0.25	36.16						
3	-0.05	37.53						
12	0.29	39.75						
11	-0.05	41.71						
6	0.04	43.19						
4	-0.15	43.91						
1	0.28	44.40						
20	0.05	44.72						
21	0.19	45.15						
13	0.02	45.48						
19	-0.26	45.76						
14	0.14	45.89						
TOTAL R^2 VALUE		45.89	TOTAL R^2 VALUE		28.52	TOTAL R^2 VALUE		22.58

the CBS areas. The derived regression equation is given below.

$$Y_{CBi} = 4.79 - 0.152X_2 - 0.209X_{10} + 0.199X_{18}$$

but as was expected there was little improvement in the overall suitability of the model. The level of explanation was in fact reduced to $r^2 = 0.225$ although the variation about the regression plane was also reduced to, S.E.E. = 13.0%. Basically however, especially since the overall regression significance level was only low (F = 4.54), it must once again be stressed that little confidence should or can be placed on either of these last two sets of regression equations.

5) TOTAL PEDESTRIAN CASUALTIES 1970 (TPED7Ø).

The level of pedestrian casualties in any urban area was found once again to be highly related to the size and degree of interaction within that area. It can be seen from the derived equation given below that

$$Y_i = 248.4 + 375.8X_1 - 19.62X_3 - 12.8X_{10} + 22.7X_{15} - 14.3X_{16} - 12.0X_{17} - 55.9X_{18}$$

the only significant variables outside these size variables were X_3 : COW7Ø and X_{10}: CI7Ø, of which it would seem only the latter may be modified in order to control the value of TPED7Ø. The significance of COW7Ø is of little surprise since the inverse relationship noted simply shows that as the level of car ownership increases, the number of pedestrian casualties decreases. The simple result therefore of varying pedestrian exposure.

The regression ranking for TPED7Ø in Table (9.3.12) shows the impact once again of X_1: INTACT7Ø. This variable accounts for 96% of the total variation in TPED7Ø, with only X_{18}:NPOP accounting for a percentage $> 1\%$.

The overall explanation level of the regression equation is approximately 99% ($r^2 = 0.988$) and the variation about the regression plane was 16.2%. Thus although the variation in this dependent variable was well explained by this regression equation there was a fair amount of regression deviation. However, a study of the residual plots for TPED7Ø revealed no abnormalities, and when a logarithmic transformation was applied there was a reduction in the level of explanation ($r^2 = 0.965$) as well as a reduction in the regression variation (4.3%). For these reasons therefore this line of study was not persued any further.

When TPED7Ø was looked at in relation to the CBS areas, it was found that an even better relationship was derived with the level of explanation being raised to $r^2 = 0.992$, and the regression variance reduced to 10.1%. The regression equation derived is given below and the close comparison with the original regression

$$Y_{CB_1} = 438.2 + 490.47X_1 - 22.3X_2 - 20.1X_4 + 24.7X_{15} + 41.9X_{17} - 33.6X_{18} + 107.3X_{20}$$

equation is obvious, and can be seen in Table (9.3.8). There is one significant addition in this equation, and that is the introduction of X_4: SHI7Ø with an inverse relationship with TPED7Ø.

Table 9.3.12 Regression Ranking for TPED70.

MODEL ONE			MODEL TWO			MODEL THREE		
VARIANCE RANK NO.	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION
1	0.98	96.10	1	0.98	96.10	1	0.98	96.10
18	0.47	97.87	18	0.47	97.87	18	0.47	97.87
15	-0.01	98.34	15	-0.01	98.34	15	-0.01	98.34
3	-0.29	98.54	3	-0.29	98.54	3	-0.29	98.54
16	-0.16	98.68	16	-0.16	98.68	16	-0.16	98.68
10	-0.07	98.78	10	-0.07	98.78	10	-0.07	98.78
17	-0.54	98.82	17	-0.54	98.82			
22	-0.31	98.85						
7	0.07	98.86						
11	-0.12	98.88						
12	0.52	98.90						
2	0.03	98.91						
19	-0.14	98.92						
20	-0.07	98.93						
9	-0.54	98.93						
14	0.09	98.93						
13	0.21	98.94						
5	-0.17	98.94						
6	0.11	98.95						
TOTAL R^2 VALUE		98.95	TOTAL R^2 VALUE		98.82	TOTAL R^2 VALUE		98.78

Although the actual impact of SHI7Ø upon the level of pedestrian casualties is very limited, it does indicate the marginal impact of such structural variables. Its absence from the original 100 area regression equation was in some ways surprising, but it would seem with hindsight that the impact of SHI7Ø is limited in the smaller less urbanised non-county boroughs, and consequently was deemed insignificant in the total regression situation.

Summarising therefore it would seem that TPED7Ø is mainly explained by the one variable INTACT 7Ø, as was the case with both TA7Ø and TC7Ø. However, it would also seem that the actual level of pedestrian accidents is also affected by the two variables CI7Ø and COW7Ø, with the third variable SHI7Ø being significant when the influence of non-county borough areas is removed from the regression.

Although the level of explanation for all the regression equations related to TPED7Ø was consistently high, there still remained some regression variance which needs further study. With this thought in mind the impact of the size variables was once again held constant, and the remaining variation studied in the following regression analysis.

6). TOTAL PEDESTRIAN CASUALTIES PER 10^3 POPULATION 1970 (TPEDPP7Ø)

The most striking feature of this regression is that whilst the level of explanation for this dependent variable is not as high as those for the absolute dependent variables, it is much higher than those for the comparable variables TCPP7Ø and TAPP7Ø.

Table 9.3.13 Regression Ranking for TPEDPP70.

MODEL ONE			MODEL TWO			MODEL THREE		
VARIANCE RANK NO.	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION
3	-0.57	32.19	3	-0.57	32.19	3	-0.57	32.19
1	0.50	47.08	1	0.50	47.08	1	0.50	47.08
16	0.25	52.40	16	0.25	52.40	16	0.25	52.40
5	-0.35	54.83	5(R)	-0.35	54.83	5	-0.35	54.83
11	-0.22	57.42	11	-0.22	57.42	11	-0.22	57.42
18	0.24	58.96	18	0.24	58.96			
22	-0.45	61.14	22	-0.45	61.14			
7	-0.09	62.97	7	-0.09	62.97			
4	-0.18	64.52	10	-0.16	63.96			
15	0.09	65.34						
21	-0.20	66.12						
13	0.44	66.47						
2	0.00	66.73						
14	0.07	67.41						
20	0.01	67.68						
19	-0.22	67.90						
9	-0.42	68.00						
8	-0.11	68.07						
12	0.56	68.11						
TOTAL R^2 VALUE		68.11	TOTAL R^2 VALUE		63.96	TOTAL R^2 VALUE		57.42

The r^2 value obtained in this regression was 0.640, and the regression variance was 16.8%

The regression ranking for TPEDPP7Ø is given in Table (9.3.13) and from this and also the derived regression equation given below, it can be seen that the three size variables X_1, X_{16}, X_{18} are all

$$Y_1 = 1.64 + 0.11X_1 - 0.13X_3 - 0.14X_7 - 0.06X_{10} - 0.11X_{11}$$

$$+ 0.10X_{16} + 0.09X_{18} - 0.13X_{22}$$

amongst the most important variables, with the remainder mainly being physical structural variables.

Looking at these significant variables it can be seen that the one which accounts for the greatest amount of variation in TPEDPP7Ø is COW7Ø (32.2%). As was true in the previous regression, the importance of this variable lies in its determination of pedestrian exposure, which must be related inversely.

The second variable in importance in terms of level of additional explanation is X_1: INTACT7Ø which indicates that the level of pedestrian casualties is not solely related to population size, but also the degree of interaction, which has not been fully accounted for by holding the population levels of different urban areas constant.

The remaining variables in this regression meanwhile include such variables as SHI7Ø, PDI7Ø, HTI7Ø etc. all of which have a negative relationship with the dependent variable, and which all account approximately for 2% of the variation.

The results of this regression imply that although the number of pedestrian casualties increases with the population size of that urban area, the actual relationship is much more complex with several structural variables being able to modify this simple increase. When one compares these conclusions with those for TCPP7Ø and TAPP7Ø, the most obvious difference revolves around the fact that TPEDPP7Ø seems to be affected mainly by internal structural factors, whilst the other two variables seem to be more affected by external influences such as the amount of through traffic. The implication here therefore is that what may be suitable for one type of casualty may not suit the other type of casualty, and in some cases these two groups could be in direct conflict, with each other. Basically this difference can perhaps be categorised by the fact that one needs a macro approach whilst the other needs a micro approach, which must be implemented through the direct study of, and application of local problems and remedies.

When TPEDPP7Ø is studied in relation to the CBS areas only, the overall level of explanation is slightly reduced to $r^2 = 0.528$ but at the same time the variation about the regression plane is reduced to 14.0%.

The variables that are found to be significant in this CBS regression are very similar to those found in the original regression, although as the derived regression shows below, there were some interesting modifications. For example the variable COW7Ø was

$$Y_{CB_i} = 1.92 - 0.12X_2 + 0.19X_3 + 0.08X_6 - 0.15X_{13} - 0.10X_{18}$$

replaced by the variable X_{18} : VRI7Ø, which however, measures virtually the same thing. Similarly included in this equation is the Ethnic index (Ei7Ø) which was applied to all the CBS regressions. Although this variable is only just statistically significant, the recorded positive relationship with TPEDPP7Ø does perhaps emphasise the particular problems of these immigrant sections of the population, and the advantages which may accrue from further attention in this area of road safety propaganda and education.

(7) <u>TOTAL CASUALTIES PER 10^6 VEHICLE-MILES 1970</u> (TCVM7Ø)

This variable which relates total casualties to vehicle miles travelled is the simplest exposure dependent variable used in this research. As was pointed out earlier in this report, the inclusion of some measure of exposure has continually been proved to be vital in research of this kind. For example, without exposure it was possible to conclude that young drivers are more probable to have an accident due to inexperience, than any other group of drivers. Yet recent research has shown that perhaps if exposure rates are included this group of drivers is no more vulnerable than any other group. Accordingly, therefore this variable was looked upon as being very important in this research.

The regression ranking for TCVM7Ø is shown in Table (9.3.14) and the derived regression equation is presented below.

Table 9.3.14 Regression Ranking for TCVM70.

MODEL ONE			MODEL TWO			MODEL THREE		
VARIANCE RANK NO.	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION
1	-0.56	33.32	1	-0.56	33.32	1	-0.56	33.32
16	-0.46	65.21	16	-0.46	65.21	16	-0.46	65.21
18	-0.31	69.40	18	-0.31	69.40	18	-0.31	69.40
10	-0.17	72.35	10	-0.17	72.35	10	-0.17	72.35
22	0.46	75.02	22	0.46	75.02	22	0.46	75.02
5	0.28	75.98	5	0.28	75.98			
12	-0.51	76.85	12	-0.51	76.85			
15	0.09	78.37	15	0.09	78.37			
4	0.10	79.00						
20	-0.02	79.30						
7	0,00	79.63						
19	0.14	79.86						
14	-0.06	80.01						
3	0.29	80.14						
8	0.09	80.31						
21	-0.01	80.56						
9	0.44	80.70						
13	-0.10	80.82						
6	-0.09	80.93						
2	0.02	80.99						
17	0.27	81.09						
TOTAL R^2 VALUE		81.09	TOTAL R^2 VALUE		78.37	TOTAL R^2 VALUE		75.02

$$Y_i = 18.83 - 8.02X_1 + 1.70X_5 - 2.47X_{10} + 5.44X_{12} - 2.61X_{15}$$

$$- 9.33X_{16} - 4.56X_{18} + 3.96X_{22}$$

The level of explanation produced by this equation was reasonable accounting for almost 79% of the total variation in TCVM7Ø. However, when the degree of regression variance was calculated it was found that S.E.E. expressed as a percentage of the mean response, equalled the high value of 34.5%. In any situation such a value would always cast considerable doubt upon the actual model used in this regression. However, before moving onto this problem, it would be useful to have a quick detailed look at the variables introduced into this regression.

The most important variable in explaining the variation in TCVM7Ø was once again X_1: INTACT7Ø which explained 38.3% of the variation. This time however, the relationship derived between INTACT7Ø and the dependent variable was negative. That is, as the degree of interaction increases the volume of road casualties per vehicle-mile is decreased. Thus one comes to the situation that although the larger urban areas are more "exposed" (in terms of vehicle-miles travelled) than smaller urban areas, these larger areas are relatively more safe. This point can also be seen in that $\overline{Y}$ for all 100 sampled areas = 18.481 yet $\overline{Y}$ for only CBS areas = 10.08! The logic behind this pronouncement cannot be proved one way or the other at present, but it can be hypothesised that the lower TCVM7Ø values are associated with the larger urban areas because of, 1) the higher levels of congestion and therefore lower speeds associated with these areas and, 2) the higher levels of income associated with these areas with which improvements can be made, both in terms of reconstruction and major network improvements, and also in terms of stringent local solutions and degrees of traffic management.

Against the impact of INTACT7Ø one must meanwhile balance the influence of two other variables X_{10} and X_{22} which both account for approximately 3% of the variation in TCVM70. The former is the shape variable CI7Ø which has its usual inverse relationship with the dependent variable, whilst the latter is the road density index (RDI7Ø) which is positively associated with TCVM7Ø. Therefore since the mean value of RDI7Ø for CBS areas is higher than that for all areas, this variable is affecting the value of TCVM7Ø in the opposite direction to INTACT7Ø, and suitable modification should take place.

It was argued above that because of the high degree of variation about the regression plane for this dependent variable, considerable doubt should be cast upon the viability of the model used. To collect further information about this problem, the residual graph plots were studied in some detail, with especial attention given to the plot between the derived residuals (e_i) and the dependent variable TCVM7Ø. This plot is shown in Fig. (9.3.15), and from this it can be seen that there is some distinct relationship between the two variables which can be described simply as parabolic. Thus the derived linear regression equation is underestimating the value of TCVM7Ø at both low and high values,

Fig. 9.3.15 PLOTT OF TCVM70 AGAINST THE REGRESSION RESIDUALS

whilst overestimating at medium values. The probability therefore is that the relationship between TCVM7Ø and the various independent variables is polynomial in nature rather than simple linear.

For these reasons it was decided to transform the dependent variable TCVM7Ø. The eventual transformation used was logarithmic, and the result is given below;

$$\text{Log } Y_1 = 1.16 - 0.27X_1 + 0.03X_5 - 0.04X_{10} + 0.11X_{12} - 0.04X_{15}$$
$$-0.18X_{16} - 0.11X_{18} - 0.02X_{19} + 0.11X_{22}$$

It can be seen from this equation and also Table (9.3.7), that there is very little difference between the two sets of significant variables for the transformed and non-transformed equations. As regards the regression equation itself however, there is considerable difference. In the first instance the level of explanation rises to almost 94% ($r^2 = 0.937$) and the variation about the regression plane is reduced to 7.8%. Similarly when the residual plot for this transformed regression is observed the parabolic trend noted earlier has been almost totally removed, and the plot as shown in Fig. (9.3.16) is almost random. In attempting to explain this improvement one should immediately look to the independent variable X_1: INTACT7Ø since this variable now accounts for 66.3% of the total variation in LNTCVM7Ø. Thus it would seem that the inverse relationship noted previously between TCVM7Ø and INTACT7Ø, is logarithmic in nature, suggesting that the improvements abounding in the larger towns are not as good as was originally predicted, but which still do accrue.

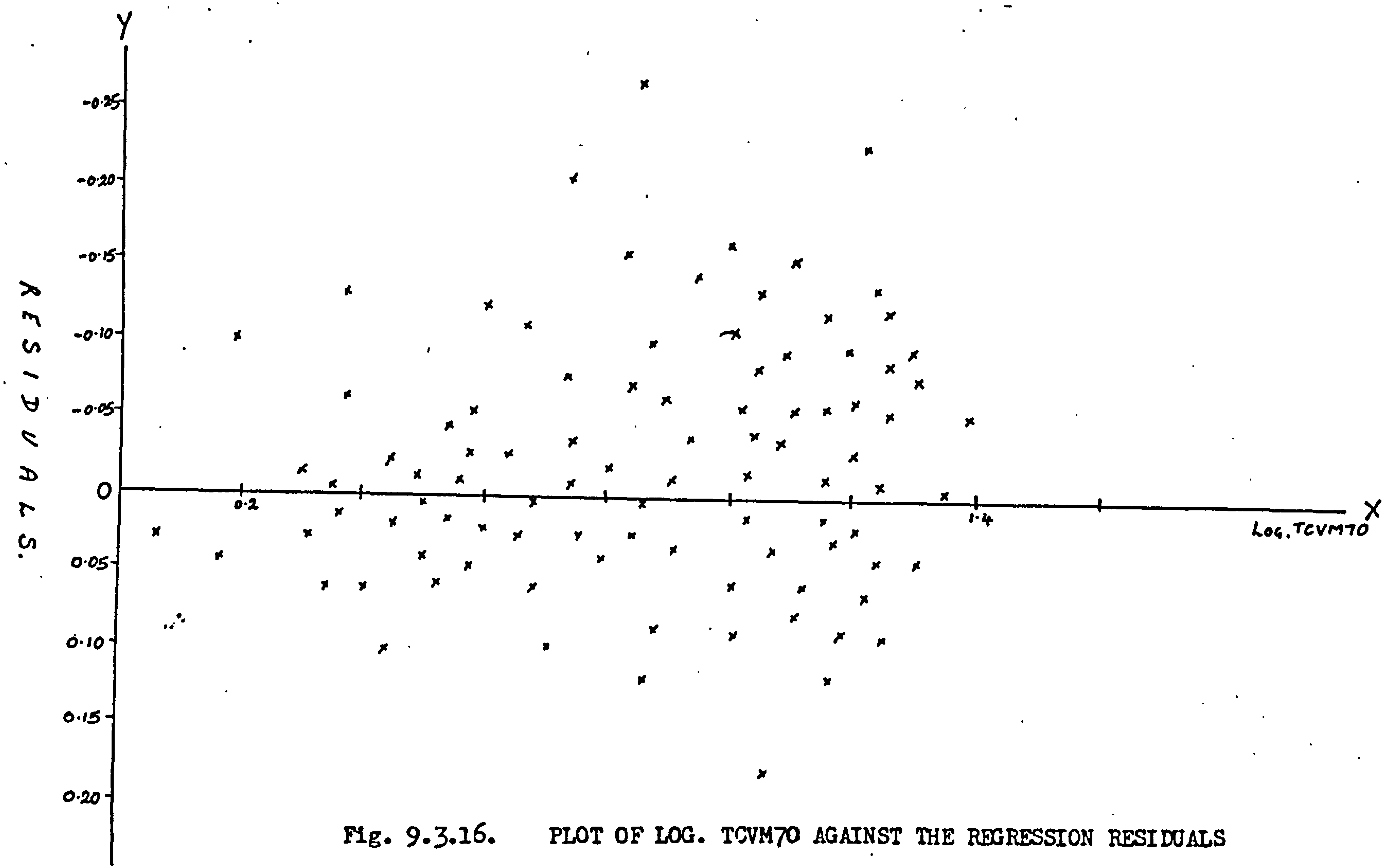


Fig. 9.3.16. PLOT OF LOG. TCVM70 AGAINST THE REGRESSION RESIDUALS

Before leaving this transformed regression, it should be noted that in this form the police variable PPI7Ø (X_{19}) once again becomes significant, indicating the influence of police action upon road safety levels, even though the influence may only be marginal.

Turning to the regression for TCVM7Ø for only CBS areas, the derived equation is given below:

$$Y_{CB_i} = 10.08 - 2.75X_1 - 2.45X_3 + 0.86X_{16} + 3.46X_{19}$$

$$+ 4.26X_{20} - 1.13X_{22}$$

As can be seen the non-transformed regression version was employed here, and although the goodness of fit left much to be desired (S.E.E. = 23.8%), the level of explanation remained fairly high ($r^2 = 0.875$). From this regression there are perhaps three conclusions and comments which should be made:-

1) Although the variation about the regression plane remains high, the parabolic relationship noted for the full 100 sample regression no longer is obvious, indicating that the relationship for these CBS areas is almost linear, and the logarithmic relationship is only necessary when the smaller non-county borough areas are included.

2) The relationship between RDI7Ø and TCVM7Ø for only CBS areas is negative and not positive as was the case for the 100 sample areas. Thus the impact of road density would seem to have

some kind of threshold value, above which congestion etc. seems to invert the relationship, such that at high values of RDI7Ø the impact upon TCVM7Ø is such that values of TCVM7Ø decrease as RDI7Ø increases.

3) Once again the police variable PPI7Ø is significant in the CBS areas perhaps indicating that the influence of police action is more successful in the larger urban areas, where more manpower and instrumental devices are available. However, it should also be pointed out that these larger areas also provide a higher probability of escaping detection, which may reduce the impact of PPI7Ø.

8) TOTAL ACCIDENTS PER 10^6 VEHICLE MILES. 1970 (TAVM70)

This regression is once again virtually identical to the preceding variable, and accordingly the arguments pertaining to TCVM7Ø also apply here. The overall level of explanation for the regression shown below was $r^2 = 0.739$, with a regression variance (S.E.E.) of 39.2%. The equation form was;

$$Y_i = 14.89 - 4.17X_1 - 1.94X_{10} - 1.47X_{11} - 1.59X_{14} - 5.74X_{16}$$
$$- 3.80X_{18} + 1.99X_{22}$$

The regression ranking for this variable is shown in Table (9.3.17) from which it can be seen that the set of significant variables is very similar to those derived for TCVM7Ø. This similarity also extends to the residual plots and therefore the presumed logarithmic relationship between the independent variables and TAVM7Ø.

Table 9.3.17 Regression Ranking for TAVM70.

MODEL ONE			MODEL TWO			MODEL THREE		
VARIANCE RANK NO.	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION	VARIANCE RANK NO	CORRELATION WITH DEP VAR	CUMULATIVE EXPLANATION
1	-0.57	32.42	1	-0.57	32.42	1	-0.57	32.42
16	-0.42	59.91	16	-0.42	59.91	16	-0.42	59.91
18	-0.34	65.45	18	-0.34	65.45	18	-0.34	65.45
10	-0.18	68.74	10	-0.18	68.74	10	-0.18	68.74
22	0.42	71.36	22	0.42	71.36	22	0.42	71.36
14	-0.14	72.30	14	-0.14	72.30			
11	-0.05	73.94	11	-0.05	73.94			
13	-0.10	74.41						
4	0.08	74.93						
21	-0.08	75.50						
12	-0.48	76.13						
15	0.10	77.19						
8	0.10	77.69						
3	0.22	78.06						
5	0.27	78.22						
20	0.03	78.29						
9	0.45	78.36						
2	-0.05	78.41						
17	0.31	78.44						
TOTAL R^2 VALUE		78.44	TOTAL R^2 VALUE		73.94	TOTAL R^2 VALUE		71.36

As regards the regression against the CBS areas only, the level of explanation was increased to 90% ($r^2 = 0.895$) and the variance about the regression plane was reduced to 21.7%. The regression is given below;

$$Y_{CB_1} = 7.94 - 1.58X_1 - 2.29X_3 + 0.62X_{11} - 0.95X_{12} + 0.68X_{16}$$

$$+ 2.04X_{19} + 3.08X_{20} - 0.90X_{22}$$

Once again the similarity with TCVM7Ø is striking, with just two extra variables in the equation, TCI7Ø and NAVI7Ø, which relate to the shape of the road network and the age composition of the population structure. It is interesting that the variable introduced here was NAVI7Ø and not DVI7Ø which would have been an indication that young drivers are more vulnerable to accidents than other age groups. As it stands, because of the introduction of NAVI7Ø, this test hypothesis definitely cannot be accepted on the present data, and research, and one must presume that when exposure variables are introduced there is no difference between any of the age groups in the population.

Besides these eight regressions described in some detail above, other regressions provided some interesting results and are deserving of some further attention. This will be done in the next section of this report by grouping these variables where possible, into homogeneous units.

(i) FATAL AND SERIOUS REGRESSIONS

This group of regressions includes the four dependent variables, TFS7Ø, TFSPP7Ø, TFSVM7Ø and SR7Ø, and are all trying to look at the incidence of the more serious casualties in different urban areas. The full list of results for these and other regressions is given in Appendix (3) and for the detailed regressions, the reader is referred to that section of this report.

The absolute dependent variable TFS7Ø exhibited characteristics similar to those described for other absolute variables. That is, the level of explanation was high ($r^2 = 0.960$) but so was the regression variance (S.E.E. = 25.7%). Examination of the residual plots showed that the derived equation was over-predicting at the higher values of TFS7Ø. However, when the variable was transformed logarithmically, certain gains were made in that the regression variance was reduced to 6.59%. However, the overall gains were limited due to the fact that (i) the level of explanation was reduced to 91% ($r^2 = 0.907$) and, (ii) the residual plots still showed considerable over-prediction at most levels.

Of the variables that were introduced into these regressions, DVI7Ø and RI7Ø both increased with higher values of TFS7Ø. That is when the incidence of young drivers within a population was high, there was a higher incidence of fatal and serious casualties, which could be a result of the postulated higher speeds of these age groups.

As regards the other three regressions in this group, none of the levels of explanation were high with $r^2 = 0.594$ for TFSVM7Ø being the maximum value, and yet all the regression variances were high when expressed as a percentage of the mean response. Therefore the main conclusions which could be reached for this group of regressions was that the variation amongst the number of fatal and serious casualties for different urban areas cannot really be accounted for by the structural variables. Accordingly, more detailed local work is necessary if the true relationships for these variables is to be found.

(ii) PEDESTRIAN CASUALTY REGRESSIONS

If the two pedestrian regressions mentioned earlier are discounted, this group also involves four dependent variables, TCPED7Ø, PEDCR7Ø, PEDWT7Ø and CPEDCR7Ø.

As can be seen from the results of these regressions, the level of explanation tends to be on the low side and accordingly very little can be said definitely about these results. It would seem, however, that the significant variables for these regressions relate to the social standing of the various urban areas, as well as the standard of housing etc. noted in these urban areas. Thus when dealing with the child pedestrian casualties, such variables as MSI7Ø, SHI7Ø and HTI7Ø are found to be significant. Even so it would seem that the major causes for pedestrian casualties are once again related to micro features rather than the macro features dealt with in this research. However, because of the additional importance of ETT7Ø, it is interesting to look at these pedestrian variables in a spatial context, as will be described later.

(iii) DRIVER AND RIDER CASUALTIES

The two absolute variables within this group TDRC7Ø and TYDR7Ø both exhibit very similar results. The levels of explanation are very high, being $r^2 = 0.959$ and 0.938 respectively, whilst the regression variances also prove to be relatively high (S.E.E. = 23.2% and 28.0%), indicating some degree of lack of fit. One of the interesting differences between the two however, is the appearance of PPI7Ø in the regression for TDRC and not for TYDR7Ø. In the first place this obviously emphasises the importance of police control in relation to driver casualties, whilst at the same time perhaps indicating that the impact of police intervention is much more marked amongst the groups of older driver and riders, than amongst young drivers. However, this statement must be qualified by the fact that the variable PPI69 is significant in the respective regression for 1969 (TYDR69), therefore one must decide upon one of two conclusions:- 1) Either PPI is not significant in 1970 as a result of a true temporal relationship, or 2) The significance of the variable PPI is understated in 1970.

Of these two hypotheses the former seems the more plausible especially when one considers that the influence of the 1967 Road Safety Act was beginning to have a reduced influence by this period.

The other variables in these two regressions of any significance were INTACT7Ø which accounted for the major portion of the explanation in each regression, and COW7Ø and DEN7Ø. The former had a positive relationship with the dependent variable and was a measure of driver exposure, whilst DEN7Ø had an inverse relationship perhaps indicating a reduction in accident levels with an increase in congestion (and indirectly a decrease in speeds).

Two other regressions which can be included in this group are TDRVM7Ø and DRCR7Ø. In both of these models the level of explanation is only moderate with values for r^2 of 0.718 and 0.547 respectively. Similarly the regression deviations are 42.5% and 17.4%, the larger of which could be explained by the parabolic residual plot noted when the residuals are plotted against TDRVM7Ø, which indicates that a logarithmic relationship is more appropriate, as was the case when dealing with the other "vehicle-mile" exposure variables.

Of these two regressions the more interesting involves DRCR7Ø since most of the relationships noted in TDRVM7Ø approximate very closely to those already described in TCVM7Ø and TAVM7Ø.

As regards DRCR7Ø the list of significant variables is of considerable interest. Besides the expected variables such as INTACT7Ø and CON7Ø, there are also four other variables included. The variable CHSI7Ø was found to have an inverse relationship within this regression. Thus as the amount of new development increases the number of driver and rider casualties as a proportion of all casualties, is reduced. This is consistent with the former observation that there was a positive relationship between CHSI7Ø and PEDST7Ø. The reasons for this are either that new development improves the road network for drivers at the expense of pedestrian safety, or alternatively induces more people to travel around on foot because of the improved environment, thereby increasing pedestrian exposure.

Also important within this regression are the two variables DOS7Ø and RI7Ø. The former has a negative relationship with the dependent variable perhaps implying that as the amount of open space increases there are relatively more pedestrian and/or pedal cycle injuries, due to increased movement. The RI7Ø variable meanwhile has a high positive relationship with DRCR70, which could be an indirect implication of including the riders of motor cycles and motor bikes[12] in this dependent variable, as well as the direct result of increased road hazards due to inclement weather.

(iv) MISCELLANEOUS

Of the remaining regressions, three will be described under this section. Two of these TJUN7Ø and TTRN7Ø describe virtually the same manoeuvre and therefore will be discussed together. For both of these the r^2 value was high being 0.961 and 0.907 respectively, however the regression deviations were also reasonably high being 29.3% and 26.8%. Within the two sets of significant variables two appear of some interest. Once again the police variable PPI7Ø appears highly significant in both regressions, showing that police action can control both the absolute level of road casualties, and also that the fruits of their work are shown greatest in stopping manoeuvres which are hazardous. The other important variable was

12 - This would seem to be substantiated by the fact that RI7Ø is significant in both motor cycle variables 1969: TMC69 and MCAR69, although not for 1970

DVI70 which had a positive relationship with the dependent variable. Thus it would seem that if young drivers are more accident prone than other age groups, then this proneness is shown most in risk taking manoeuvres and at junctions. However, since this variable is not significant in the absolute variable regression, it could be that whilst young drivers are more accident prone at "risk" situations, they may equalise the situation by responding better in other potential accident situations.

The final regression to be mentioned is the one which tries to explain the reason for the number of casualties per accident in each urban area, TCTA70. Basically however, the results of this regression do little to help in this explanation, since the level of explanation is only $r^2 = 0.423$, although the regression variance is very good with S.E.E. = 5.51%.

The most important variable in this regression is CHS170 which is positively related to TCTA70. Thus the redevelopment of an area would seem to increase the incidence of driver accidents and now also the number of casualties per accident. This could be explained by the fact that either in such areas the incidence of car occupancy is higher, or alternatively the type of accidents are such that more people are injured in any accident. Thus it could be that the incidence of two, three or four car accidents is much higher in these areas instead of one car accidents which normally result in less casualties. Similar reasons could also be put forward to explain the positive relationship between DOS70 and TCTA70, with perhaps the additional factor of higher speeds associated with such areas, being included.

Having looked at the various regression equations derived during this stage of the analysis, it now becomes necessary to study the variation in accident and casualty levels, within a spatial context. It was explained in the methodological section of this report that this spatial analysis was conducted through the use of the residuals from the regression equations modified into a Mean Safety Index, and introduced into a trend surface analysis. The following section of this report will therefore analyse the results of this spatial analysis.

9.4 THE SPATIAL VARIATION ANALYSIS

It would have been possible to subject every regression equation to a trend surface analysis using the derived residuals as the input into the trend surface model.[13] However, since this would have proved both wasteful and inpractical, the number of regressions subjected to the analysis was reduced to twenty, which were considered to be of the greater importance. The list of the 20 regressions used in the trend surface analysis is given in Table (9.4.1).

---

13 Fairburn and Robinson point out that the actual calculation of the trend surface polynomials for irregularly spaced data does not refer to the complete trend as defined by Grant (1957), but only to some unspecified portion of the complete trend. Accordingly these surfaces should be referred to as "partial trend surfaces." However, for convenience in this report such a distinction will be assumed to be understood, thereby needing no modification.

1. TOTAL CASUALTIES
2. TOTAL ACCIDENTS
3. " FATAL AND SERIOUS CASUALTIES
4. " PEDESTRIAN CASUALTIES
5. 1 " DRIVER AND RIDER CASUALTIES
6. " CHILD PEDESTRIAN CASUALTIES
7. " YOUNG DRIVER AND RIDER CASUALTIES
8. " ACCIDENTS ON UNCLASSIFIED ROADS
9. " ACCIDENTS ON CLASSIFIED ROADS
10. " CASUALTIES PER 10^3 POPULATION
11. " ACCIDENTS " " "
12. CASUALTY RATE PER 10^6 VEHICLE MILES
13. ACCIDENT " " " " "
14. TOTAL ACCIDENTS ON CLASSIFIED ROADS PER MILE OF CLASSIFIED ROAD
15. " " " UNCLASSIFIED " " " " UNCLASSIFIED ROAD
16. TOTAL PEDESTRIAN CASUALTIES PER 10^3 POPULATION
17. TOTAL PEDAL CYCLE ACCIDENTS PER 10^3 PEDAL CYCLE WORK TRIPS
18. PEDESTRIAN CASUALTIES PER 10^2 WALK WORK TRIPS
19. TOTAL DRIVER AND RIDER CASUALTIES PER 10^6 VEHICLE MILES
20. TOTAL CASUALTIES/TOTAL ACCIDENTS.

TABLE 9.4.1. LIST OF THE 20 VARIABLES USED IN THE TREND SURFACE ANALYSIS.

Assessment of the resultant trend surfaces can take several forms;

1) Observation of the various surfaces produced, in order to discover the importance of independent variables, which have previously been ignored.

2) Observe the deviations from the trend surface, for the same reasons given above.

3) Observe the reduction in the sum of squares and therefore the percentage level of explanation. The significance of this reduction can further be tested using an informal analysis of variance test, as proposed by Anderson & Bancroft. (1952).

4) Test the significance of the overall surface following the addition of each individual surface, through the construction of confidence levels, based on cumulative F - tables.

It is essential to understand the difference between these last two tests since the former represents the strength of the trend, whilst the latter which say, is represented by a high confidence for a linear surface, is interpreted as meaning, ".....some flexibility in its angle and direction of dip but without sufficient freedom so that its direction of inclination could be completely reversed." (Krumbein 1963). In other words this last test is a test of the "reality" of the derived surface.

The percentage explanation level, due to the reduction in the sum of squares for the twenty regression variables, is given

			EXPLANATION LEVEL (%)					
TSM.	1ST ORDER	2ND ORDER	3RD ORDER	4TH ORDER	5th ORDER	6TH ORDER	7TH ORDER	8TH ORDER
1	0.67	1.59	12.22	14.90	16.47	22.30	24.12	–
2	0.51	2.80	13.00	16.37	18.26	21.90	22.68	–
3	4.28	6.74	8.83	14.35	17.41	23.98	28.53	–
4	1.66	2.15	6.19	6.90	12.59	17.04	20.09	27.74
5	1.20	8.06	21.71	27.29	32.53	39.36	42.88	–
6	0.05	7.21	9.15	11.20	19.47	24.81	26.87	–
7	0.58	2.70	14.00	16.57	27.35	34.56	48.41	–
8	0.88	1.78	4.17	6.73	11.04	17.66	19.17	–
9	0.88	4.14	8.42	10.65	19.80	29.55	41.84	–
10	2.45	6.73	12.82	16.30	19.91	33.81	35.96	–
11	6.66	14.64	25.34	27.56	33.89	42.63	43.14	43.99
12	0.20	3.09	3.37	7,65	10.92	19.81	24.82	–
13	0.76	2.48	3.45	7.85	11.82	18.91	29.26	–
14	0.89	1.50	2.78	13.84	25.85	30.85	39.42	–
15	0.26	1.56	4.23	9.84	17.54	22.39	42.50	–
16	1.10	2.06	4.67	9.10	11.48	20.23	24.90	43.21
17	3.03	5.17	13.73	15.13	20.61	26.26	30.27	31.64
18	4.28	13.65	20.70	23.23	33.79	36.23	36.49	–
19	0.30	2.33	3.00	6.74	9.99	16.62	24.63	52.70
20	1.13	3.16	6.06	8.48	13.29	21.80	39.39	–

TABLE 9.4.2. THE EXPLANATION LEVEL FOR EACH ORDER SURFACE OF THE 20 TREND SURFACE ANALYSES.

TSM	1st Reg.	2nd Reg.	3rd Reg.	4th Reg.	5th Reg.	6th Reg.	Tot. Reg.	Best Surface	F. Value	Sig. Level
1.	0.33	0.29	2.73	0.54	0.25	0.77	0.77	3	1.39	775%
2.	0.25	0.74	2.64	0.68	0.31	0.48	0.75	3	1.49	775%
3.	2.17	0.83	0.51	1.10	0.49	0.89	0.84	1	2.17	775%
4.	0.59	0.16	0.97	0.13	0.86	0.55	0.55	-	-	-
5.	0.59	2.34	3.92	1.30	1.02	1.16	1.73	6	1.73	795%
6.	0.03	2.42	0.48	0.39	1.35	0.73	0.88	2	1.46	775%
1.	0.29	0.68	02.96	0.52	1.95	1.13	1.41	5	1.49	775%
8.	0.44	0.29	0.56	0.47	0.64	0.83	0.57	-	-	-
9.	0.32	1.06	1.05	0.42	1.50	1.42	1.12	6	1.12	775%
10.	1.22	1.43	1.57	0.71	0.59	2.15	1.36	6	1.36	775%
11.	3.47	2.93	3.23	0.52	1.26	1.57	1.98	6	1.98	795%
12.	0.10	0.94	0.06	0.79	0.48	1.14	0.66	-	-	-
13.	0.26	0.55	0.23	0.81	0.59	0.90	0.62	-	-	-
14.	0.44	0.19	0.30	2.18	2.13	0.74	1.19	5	1.38	775%
15.	0.13	0.41	0.63	1.06	1.23	0.64	0.77	-	-	-
16.	0.55	0.31	0.61	0.83	0.35	1.13	0.68	-	-	-
17.	1.52	0.71	2.23	0.28	0.91	0.79	0.95	3	1.59	775%
18.	3.30	3.40	2.00	0.56	2.10	0.39	1.51	5	2.01	795%
19.	0.15	0.65	0.16	0.68	0.48	0.82	0.53	-	-	-
20.	0.55	0.64	0.73	0.43	0.74	1.13	0.74	-	-	-

TABLE 9.4.3. SIGNIFICANCE TESTS FOR THE INDIVIDUAL SURFACES OF THE TREND SURFACE ANALYSIS (F. VALVES)

TSM	1st Surf.	2nd Surf.	3rd Surf.	4th Surf.	5th Surf.	6th Surf.
1.	0.33	0.30	1.39	1.06	0.78	0.77
2.	0.25	0.54	1.49	1.19	0.88	0.75
3.	2.17	1.36	0.97	1.02	0.83	0.84
4.	0.59	0.41	0.66	0.45	0.57	0.55
5.	0.59	1.65	2.77	2.28	1.90	1.73
6.	0.03	1.46	1.01	0.77	0.96	0.88
7.	0.29	0.53	1.63	1.21	1.49	1.41
8.	0.44	0.34	0.44	0.44	0.49	0.57
9.	0.32	0.81	0.92	0.72	0.98	1.12
10.	1.22	1.36	1.47	1.18	0.98	1.36
11.	3.47	3.22	3.40	2.31	2.03	1.98
12.	0.10	0.60	0.35	0.50	0.49	0.66
13.	0.26	0.48	0.36	0.52	0.53	0.62
14.	0.44	0.29	0.29	1.38	1.38	1.19
15.	0.13	0.30	0.44	0.66	0.84	0.77
16.	0.55	0.40	0.49	0.61	0.51	0.68
17.	1.52	1.03	1.59	1.59	1.03	0.95
18.	3.30	2.97	2.61	1.84	2.01	1.51
19.	0.15	0.45	0.31	0.44	0.44	0.53
20.	0.55	0.59	0.64	0.57	0.61	0.74
D.F.	2.97	5.94	9.90	14.85	20.79	27.72

TABLE 9.4.4. SIGNIFICANCE TESTS FOR THE COMPOUND SURFACES OF THE TREND SURFACE ANALYSIS (F. VALUES)

in Table (9.4.2) whilst the results of the two sets of significance tests extended as far as the sixth-order surfaces, are given in Tables (9.4.3) and (9.4.4) respectively. From these tables it can be seen that whilst eight of these regression variables fail to produce any surface which is significant at any level, the remaining twelve have only one surface which fails to exceed the 75% significance level. In order to describe and analyse these spatial variations, therefore, each variable will be described independently where necessary, with special emphasis being placed upon the variable TC7Ø.

(i) TOTAL CASUALTIES TREND SURFACE ANALYSIS (1970)

It has already been shown that the regression equation accounted for a high level of explanation within this dependent variable, and therefore except for one or two values the range of the residuals for this variable were not very high. However, in terms of the Mean Safety Index, the range was calculated between 300.00[14] and 46.93 for Buxton and Ryde respectively. When this variation was analysed within the trend surface analysis (TSM) 22.30% of the variation was explained by the time the sixth-order surface had been reached. However, from the relevant tables it can be seen that the cumulative surfaces are only significant (75 < F < 90%) at the third-order surface, which accounts for only 12.22% of the total variation. None of the surfaces added after this surface

14 Although some calculated values for the Mean Safety Index (MSI) were higher than 300.00, in order to keep calculations within mathematical bounds a maximum limit of 300.00 was imposed in these instances.

prove statistically significant, and are therefore possibly unreal surfaces, even though the strength of the surfaces are increasing progressively. Therefore, for this variable, three surfaces are shown in Figs. (9.4.5., 9.4.6. and 9.4.7) giving 1st, 3rd and 6th order trends, of which the third surface should be the most important.

The linear trend which accounts for only 0.67 of the variation, and is insignificant at all levels, merely indicates an increasing trend in road casualties in a general SW - NE direction. Obviously however, because of the low strength and reality features little confidence should be placed in this result at all.

The third order surface meanwhile, whilst accounting for 12.22% of the variation, also has an individual surface significant ($90 < F < 95\%$), and a cumulative surface significance above the $F = 75\%$ confidence level.[15]

Observing this surface, where the contours are marked in deviations of 5 from the 100.00 mean safety value, it can be seen that the "minimax"[16] form derived in the second order surface, has been modified into a compound minimax solution, with the two central cross-points centred at central North England, and the "gateway" to the South-west peninsula (Fig. 9.4.6). Consequently it is possible to

---

15 Another way in which to find the best fit surface is also available by finding the minimum residual mean sum of squares

16 The "minimax" surface (Davis 1956) is one wherein two axes cross at right-angles at a point which is the maximum value along one and the minimum along the other.

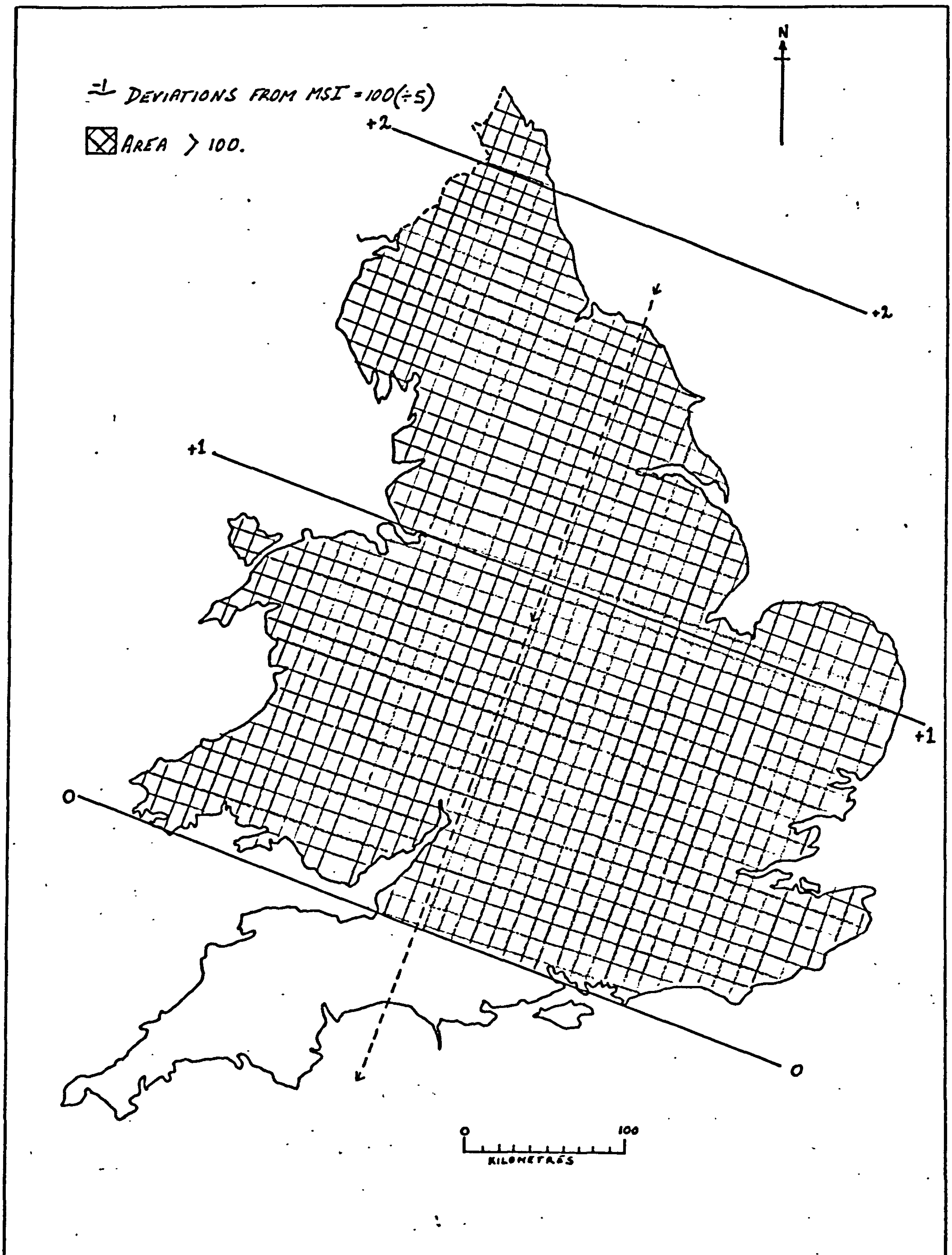


Fig. 9.4.5. LINEAR TREND SURFACE FOR TC70

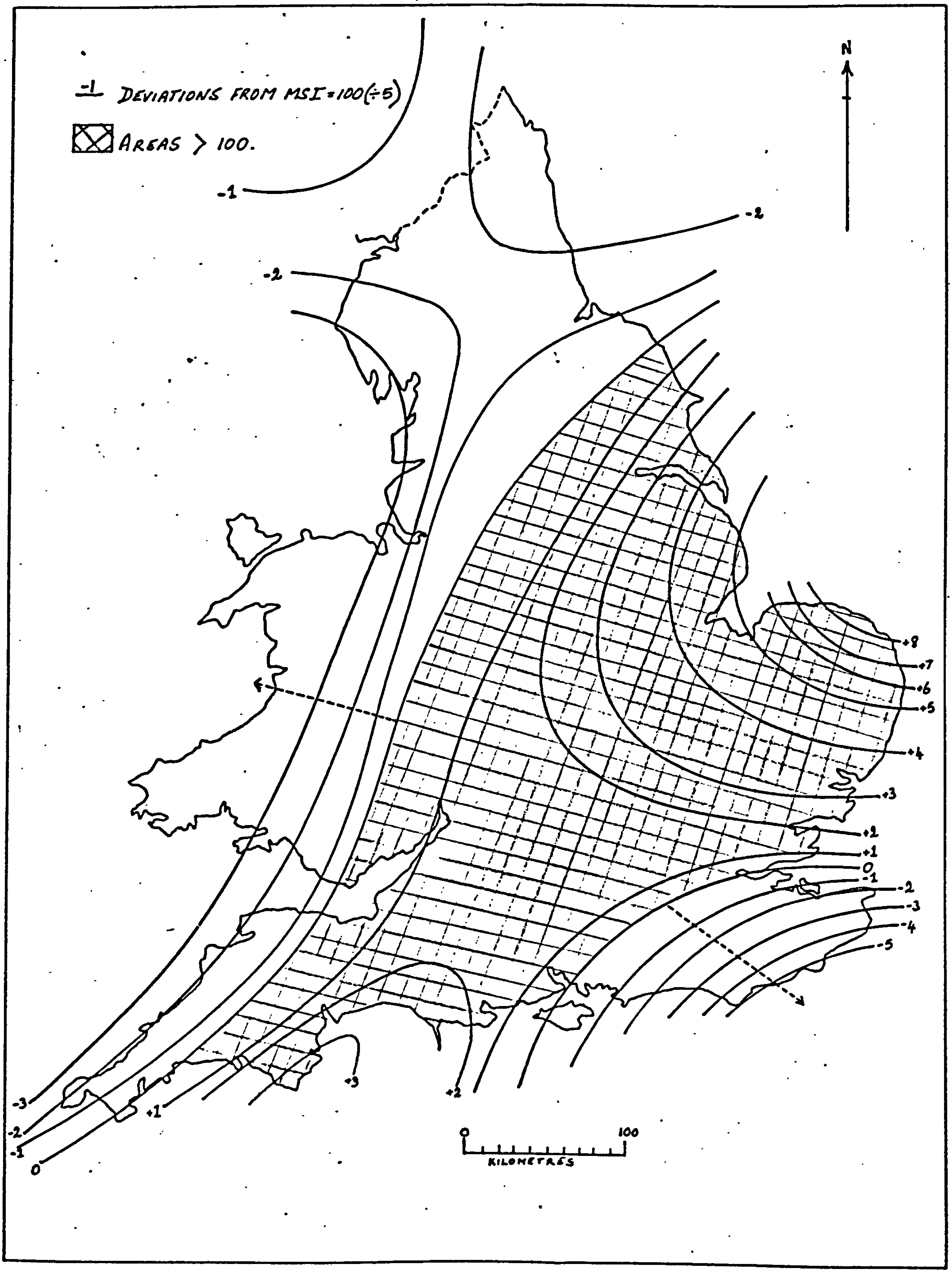


Fig. 9.4.6. CUBIC TREND SURFACE FOR TC70

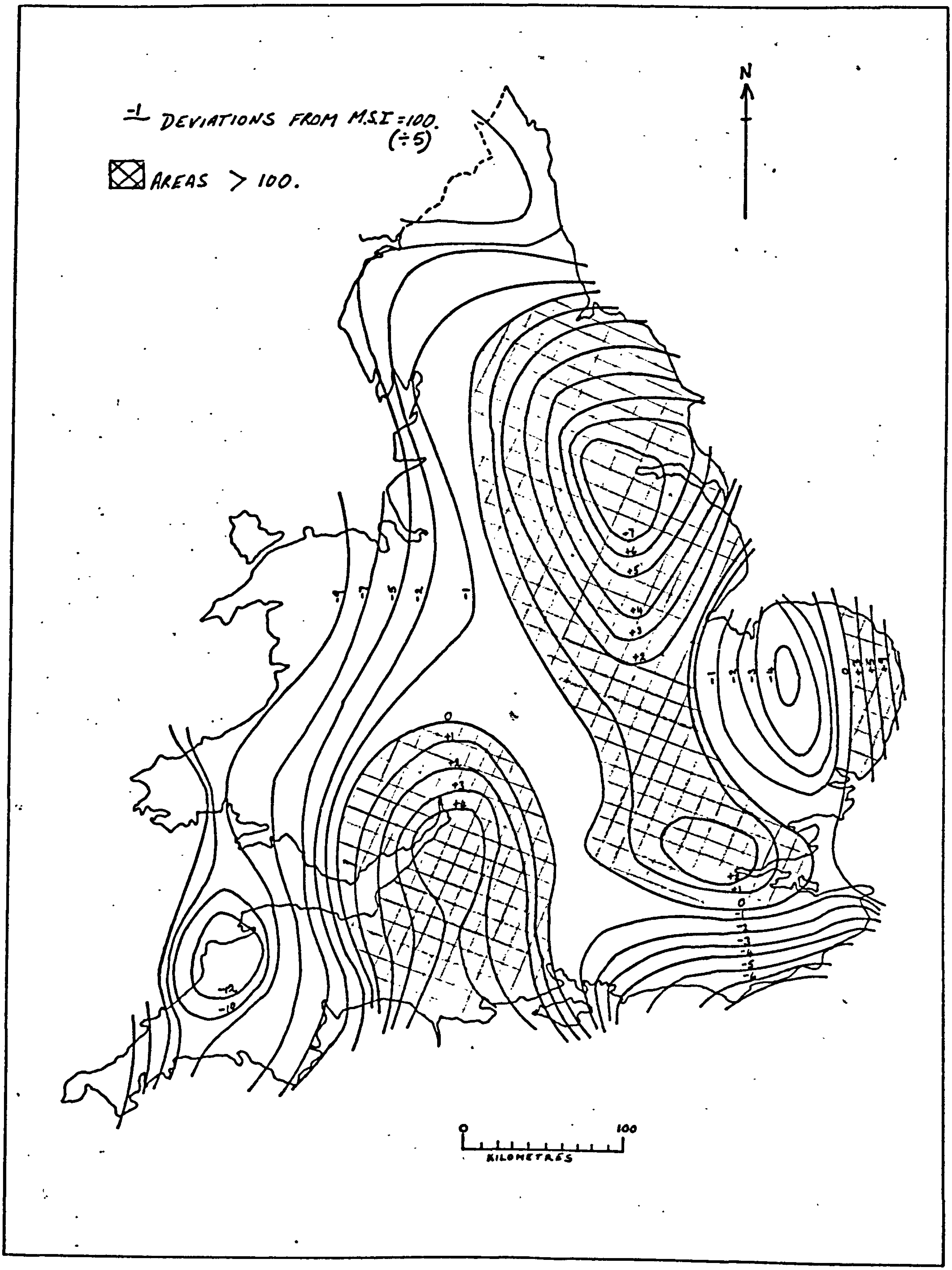


Fig. 9.4.7. SIXTH ORDER SURFACE FOR TC70

recognise four basic sub-regions:-

1) An area in the S.E. of England with a declining mean safety value as one moves in a S.E. direction.

2) An area along the east coast and extending inland as far as Sheffield and Coventry, which has a rapidly increasing set of mean safety values, reaching a peak in the area of E. Anglia.

3) A gently increasing surface rising from the west and Welsh coasts into the edge of the industrial regions of the North and Midlands.

4) A rapidly rising surface from the west to the east coast of S.W. England, with a levelling off of the surface as it moves into the minimax centroid around Bath and Bristol.

In a geographical sense this derived cubic surface is very familiar on two counts. In the first instance it isolates five of the major sub-regions of England, the border lands of the North and Wales, the S.W. Peninsula, the S.E. region and the East Coast area. Similarly this surface also seems to follow somewhat, the transport axis idealisation. That is, the four "legs" of the modified minimax solution seem to follow almost the same lines as the theoretical "industrial axes" of England, which run in a N.E.- S.W. N.W. - S.E. direction, along which the major lines of communication and transport seem to flow.

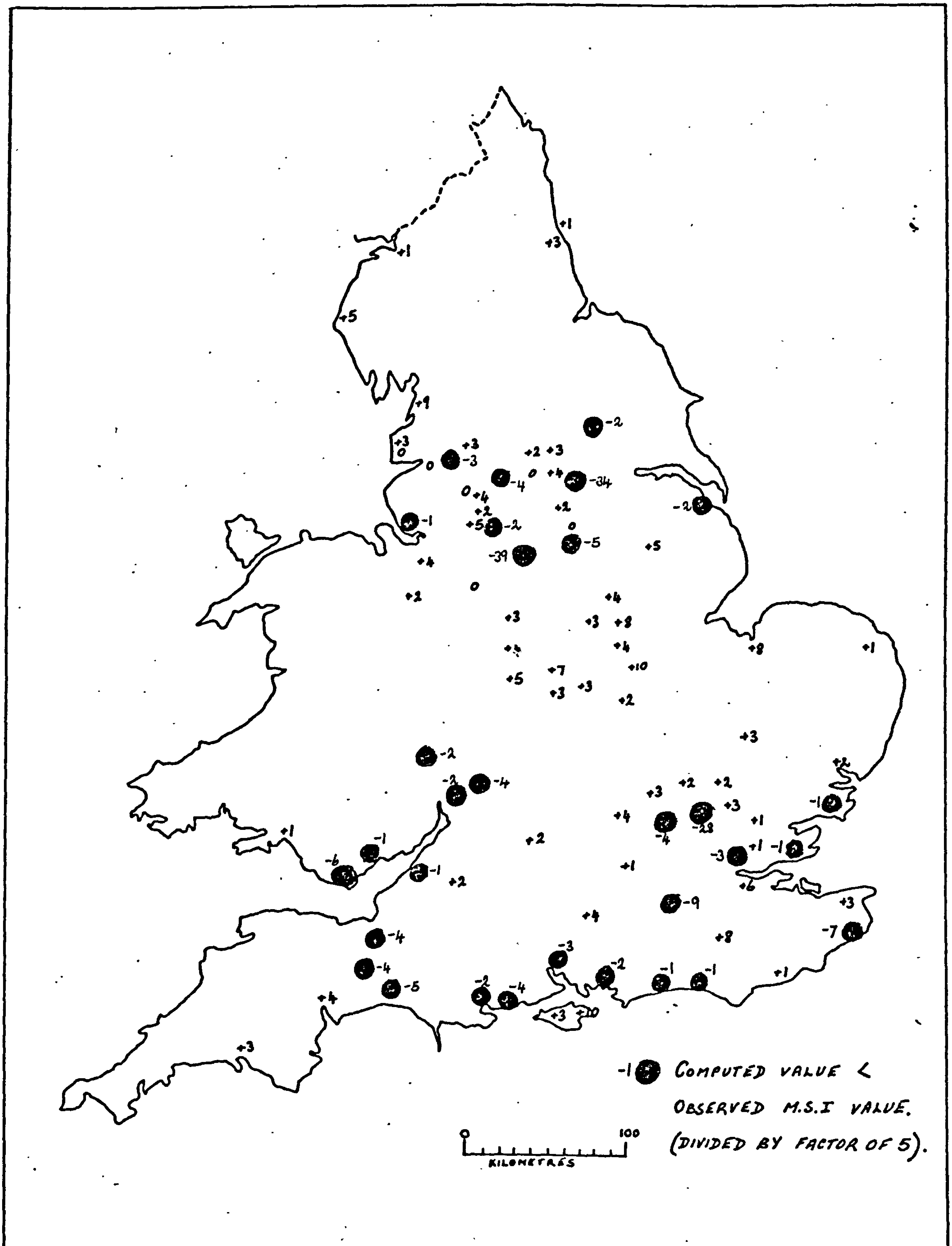
If this relationship with the transport networks is realistic, as would seem reasonable, a further interesting feature of the surface can be seen. This concerns the location of the minimum and maximum axes of the minimax solution. It can be seen that travel in a general N.W. - S.E. direction is along the maximum axis. That is the centroid of the surface is the maximum value for this axis whether one moves in a N.W. direction or S.E. direction. Alternatively the minimum axis would seem to run in a S.W. - N.E. or east direction, meaning that as one moves along this axis in either of these directions from the centroid area, the values of the MSI increase. Remembering that this is a surface for 1970, it would seem to be indicating the N.W. - S.E. bias in the major transport network systems applicable at that time. Similarly, the rising surface from both the east midlands and the urban conurbations of west Yorkshire, would also seem to indicate the lack of satisfactory transport facilities from these areas to the east coast ports of Hull and Grimsby which deal with considerable trade from these areas. These conclusions would seem therefore to further justify the completion of the Trans- and E. Pennine motorways as well as the extension of the M5 motorway into S.W. England.

It was pointed out previously that the low order surfaces are generally deemed to portray the regional trend of the phenomena under study therefore it would seem reasonable also to consider the "deviations" from these derived surfaces as representing the variability unexplained by the regional generalisations, due to local abnormalities. However, with this variable, because the explained variation is very limited (12.2%), the residuals could also indicate systematic variations which could have been accounted for by higher order surfaces. The deviations from the cubic surface for TC7Ø are therefore shown in Fig. (9.4.8). In this figure positive deviations are those where the values of the computed surface are greater than the actual value at that point. Therefore areas with positive deviations have been over-predicted by the trend surface. The converse is true for negative deviations, and so high negative values are areas where the regional component fails to explain the high incidence of traffic conflicts.

Observing this deviations map, there are various features which are immediately obvious and deserving of some further consideration, within the context of the regional trend. These features can be split into 5 groups:-

1) The most important feature of the deviations map is the heavy concentration of positive residuals in the Midlands region. Thus it seems that this region has been over-predicted by the regional trend, even though this area is on the edge of the centroid of the minimax surface. It could be that this over-prediction is a direct result of being the node of the national network system, since increased attention has obviously been given to the satisfactory movement of these quantities of long distance, and through traffic, around these urban areas. As a result, it would seem that these midland areas are able to cope with the road safety problem better than would be expected according to the national average expectation.

2) A second feature of the residuals map is the group of negative residuals which control entry into the S.W. Peninsula,



Fi . 9.4.8 DEVIATIONS FROM THE CUBIC TREND SURFACE FOR TC70 both along the north and south coasts as well as the central areas. Thus whilst the regional trend showed the need for a better network system in this area, the deviations show that the problem was under-estimated by this regional trend. It is interesting to note that all the sampled towns on the A38 south of Birmingham have negative values, the implications of which needs no further amplification. A further interesting point here however, is that once past this gateway, the sampled towns have positive deviations, perhaps implying that the high casualty levels are not only related to traffic volumes, but also the degree of congestion due to limited access into this area. However, since the number of sampled points is very limited at these points, no degree of certainty can be made upon such an assumption.

3) The third feature of the deviations map is the ring of positive deviations around London, within which there appears a smaller ring of negative values. This suggests that there are perhaps two patterns operating in this region. The inner region on the very outskirts of London seems to be characterised by high MSI values which is the opposite of the outer-ring, which incidentally includes the outer ring of new towns, and garden cities developed over the last thirty years. The result has been that the surface has tried to average these two groups, and thereby over-predicts the outer ring of urban areas and under-predicts the inner ring. The obvious reason for these two rings and this feature in the residual map is the disconcerting influence of London, which was not included in this research for that very reason.

4) The fourth feature which needs mentioning here is the consistent string of negative values associated with the coastal areas of south and south-east England. This is probably explained by the fact that this variable includes both driver and rider casualties plus pedestrian casualties, yet this latter group has already been shown to be influenced more by local factors than regional trends. Therefore where high pedestrian levels are expected, as in these holiday resort areas, it is only natural for the regional trend to underestimate the incidence of casualties for these areas. This is also further modified by the abnormal seasonal traffic variations associated with these towns purely because of their function. Thus the highest negative value for these areas is -7 for Dover, which is obviously related to its function as a cross-channel ferry port.

5) The final feature to be noted is the mixed residuals associated with the Lancashire and South Yorkshire conurbations, which fail to indicate even a marginal local trend, except perhaps for the fact that as one moves east, there is a relative increase in the number of negative values. Consequently one can only assume that these areas are more influenced by local anomalies than regional trends.

Although the 6th order-surface fails to reach a satisfactory confidence level, it is worth observing since the explanation level has been increased to 22.3%. This surface is shown in Fig. (9.4.7). As is only to be expected, this surface is much more complicated, but in many places it does agree with or satisfy some of the features noted above as regards the deviations from the cubic surface. The modifications and regions produced in this 6th order surface can best be described as follows:-

(1) A circular high centred on the Bristol Channel area which effectively blocks the approach to the S.W. by means of a ridge which extends across the area between Bristol and the Dorset coast. This area of high MSI values also extends to the industrial regions of S. Wales and the approaches to that area. Both of these modifications could have been predicted from the cubic surface as argued above.

(2) The N.W. - S.E. transport axis is maintained as a trough of low MSI values between the highs of the S.W. and the N.E..

(3) A region of high MSI values centred upon York and the Humber estuary, and covering all of the N.E. and the east midlands.

(4) An expected small high centred on the London region.

(5) An area of low MSI values centred near Kings Lynn, but increasing as one moves into the more easterly areas of East Anglia.

(6) A gently declining surface from both the Welsh and North border lands.

(7) An area of low values centred on the north Cornwall coast and covering the major part of the S.W. peninsula. This is an interesting modification since it was already noted when discussing the areas past the entry to the S.W. that the deviations tended to be positive and it was postulated then that the areas of conflict

might be restricted to a limited area, due to the congestion factor. Although this seems to be supported by this sixth-order surface it must still be treated with some caution because of the low confidence value associated with this surface.

In conclusion therefore, it would seem fair to state that although the amount of variation explained by these trend surfaces for TC7Ø is only moderate, there would seem to be some definite spatial (regional) trends which should be taken into account when studying the road accident problem, and which can perhaps be related to the national transport network systems and the function of different urban areas.

(2) TOTAL DRIVER AND RIDER CASUALTY TREND SURFACE ANALYSIS (1970) TDRC7Ø

It was suggested in the analysis of the previous variable TC7Ø that one of the reasons for the lack of explanation in terms of a regional trend, could be the influence of the pedestrian casualties included within the total number of casualties in any urban area. It was further argued that this was due to the local causation factors within the pedestrian casualty situation. Accordingly one would expect, if this were true, that the variable TDRC7Ø should exhibit a better fit and a surface more influenced by external factors when subjected to a trend surface analysis.

As regards the goodness of fit, by the time the 6th order surface was derived, 39.7% of the total variation remaining from the regression on TDRC7Ø had been explained, and the resultant surface was highly significant ($F > 95\%$). Similarly the cubic surface, which can be used to observe the initial trend, accounted for 21.7% of the variation and was also very highly significant ($F > 99\%$). These two surfaces are shown in Fig (9.4.9) and Fig (9.4.12) respectively and will be described in the following section.

The cubic surface obtained in this trend surface analysis is in many ways similar to that described for TC7Ø. However, as was hypothesised there does seem to be a much better fit to external influences, as well as a better overall fit. In fact the derived surface fits remarkedly well, onto the major national transport network system. Once again several distinct regions can be identified.

(i) A central region along the N.W. - S.E. axis with MSI values falling between 100.00 and 105.00, extending approximately from Luton to the very north of England.

(ii) An area including S.E. England within which MSI values decrease in a S.E. direction.

(iii) An area covering southern England and the beginning of the S.W. Peninsula, within which MSI values are increasing in a southerly direction.

(iv) A rising surface which covers almost all of eastern England, reaching a peak in E. Anglia.

(v) A steadily declining surface westwards from the Welsh border lands.

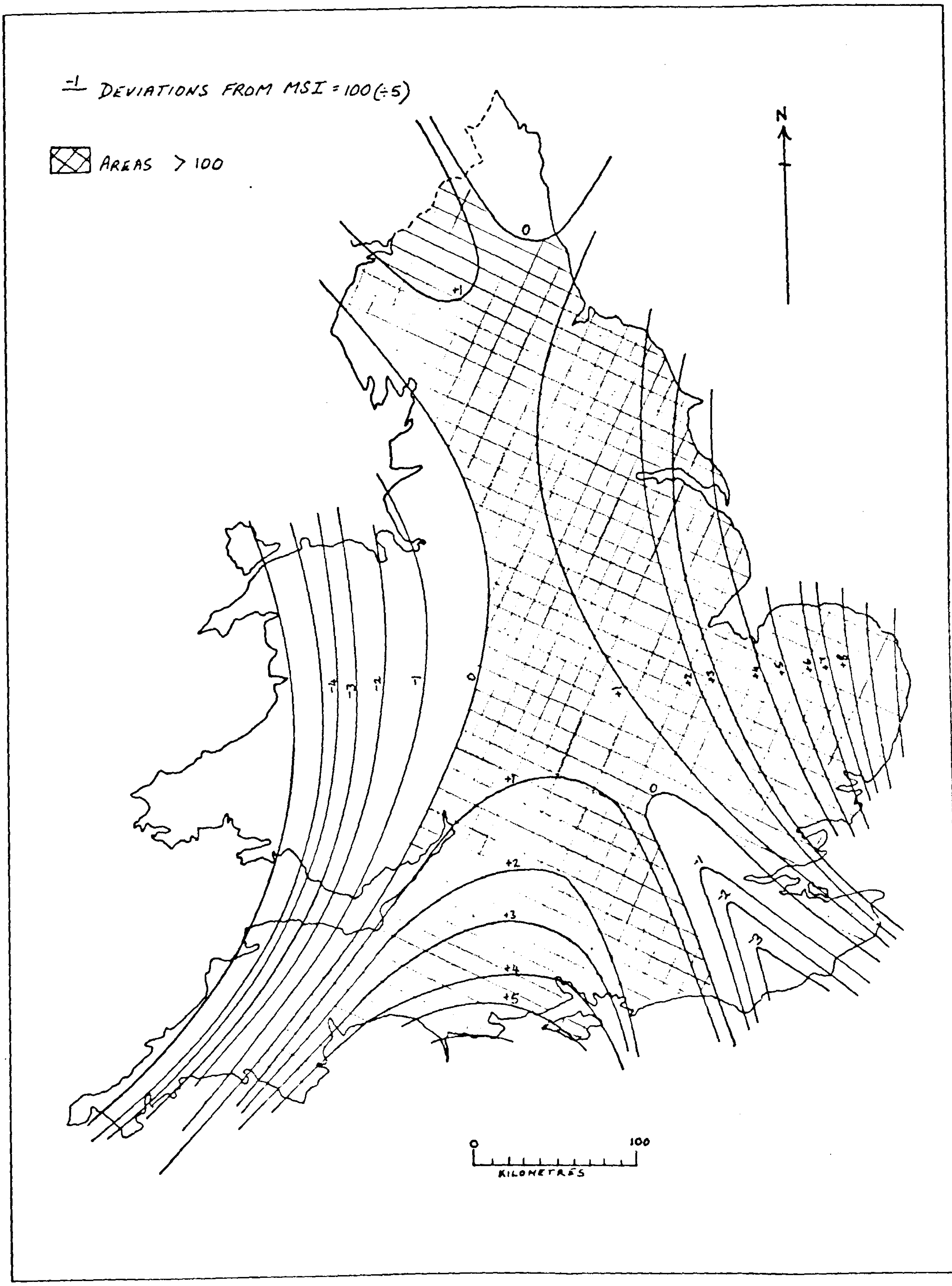


Fig. 9.4.9. CUBIC TREND SURFACE FOR TDRC70

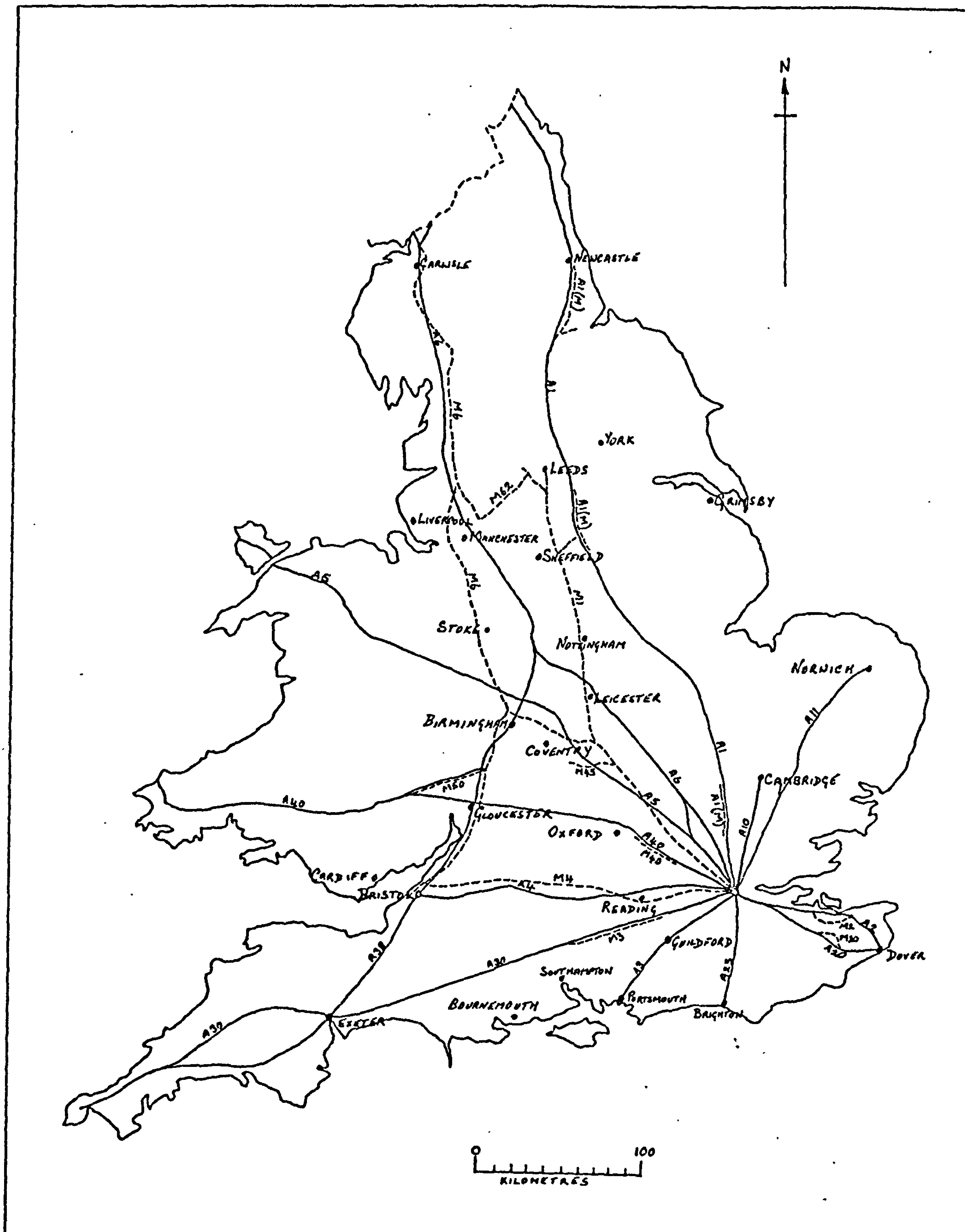


Fig. 9.4.10. THE ROAD NETWORK SYSTEM OF ENGLAND AND WALES.

Since most of these regions are comparable to those described for TC7Ø, it is unnecessary to put forward the same arguments here, except to say that once again the importance of the lack of major east-west communications and the problems of the S.W. Peninsula, are clearly elucidated. For means of comparison also, the major network highways are shown in Fig. (9.4.10).

The map of deviations from this cubic surface meanwhile are not too easy to identify (Fig. 9.4.11). However, certain features do appear to need to be qualified in higher order surfaces and these are listed below.

1) The cubic surface seems to be underestimating the MSI values especially in the S.E. area of England.

2) The areas around the Bristol Channel also seem to be underestimated and modifications to the surface are needed here.

3) Whilst there is some over-prediction in the midlands area, it is nowhere near as uniform as with variable TC7Ø.

4) The surface is definitely over-predicting in the N.W., and the Lancashire conurbations in particular.

5) Although the surface rises towards the east coast, it is obvious from the residuals that this rise should be much steeper.

6) One difference in this set of deviations and those obtained for TC7Ø which should be noted here, is that whereas for TC7Ø all the southern coast resorts had negative residuals, the deviations from this surface show a marked variation between the different

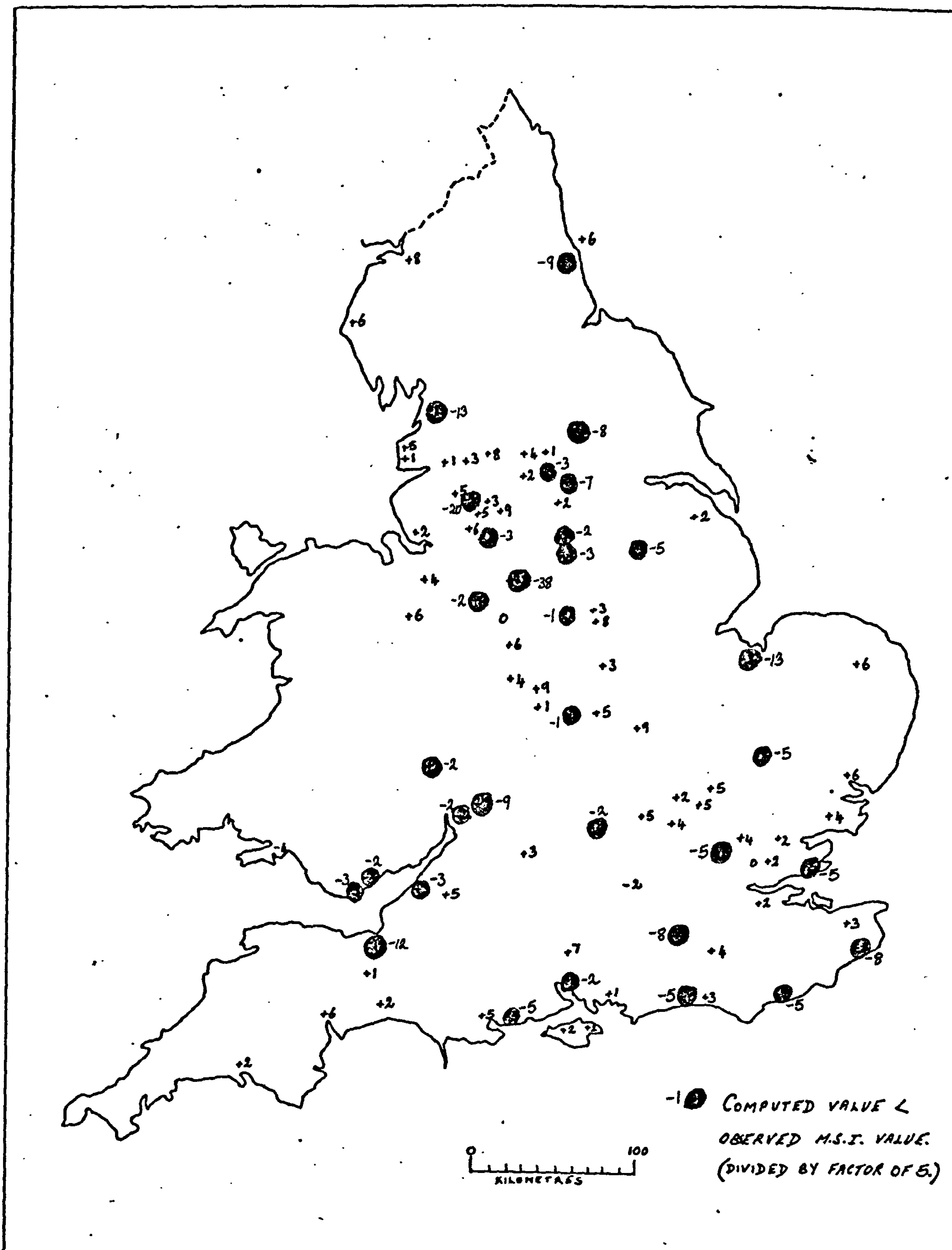


Fig. 9.4.11. DEVIATIONS FROM THE CUBIC TREND SURFACE FOR TDRC70

resorts. Therefore it would seem that the arguments put forward previously concerning the signifcance of pedestrian casualties in these areas are justified.

Since the 6th order surface accounts for 39% of the variation in TDRC7∅ it would be assumed that a lot of these deviations mentioned above would be removed within that surface. As can be seen from Fig. (9.4.12) this is very much the case with the central plateau being split into two groups, one being a 'low' and being centred over the southern midlands, and the other giving a gentle rise in the Lancashire conurbation area.

Similarly, the group of negative residuals around the Bristol Channel are removed by the introduction of a ridge into the north Devon area and extending into the west Hampshire basin. Another ridge along the eastern coastline also helps to remove the negative values in that area.

The general conclusions from this trend surface analysis therefore must be similar to that for TC7∅. That is, the residuals from the regression TDRC7∅ would seem to be strongly related to the national road network system, and the function of certain urban areas. However, it must be pointed out that whilst the southern coast resorts do still provide some anomalies, the removal of the pedestrian casualties has reduced the major abnormalities of these areas, thereby enabling these areas to be satisfactorily incorporated within the trend surfaces.

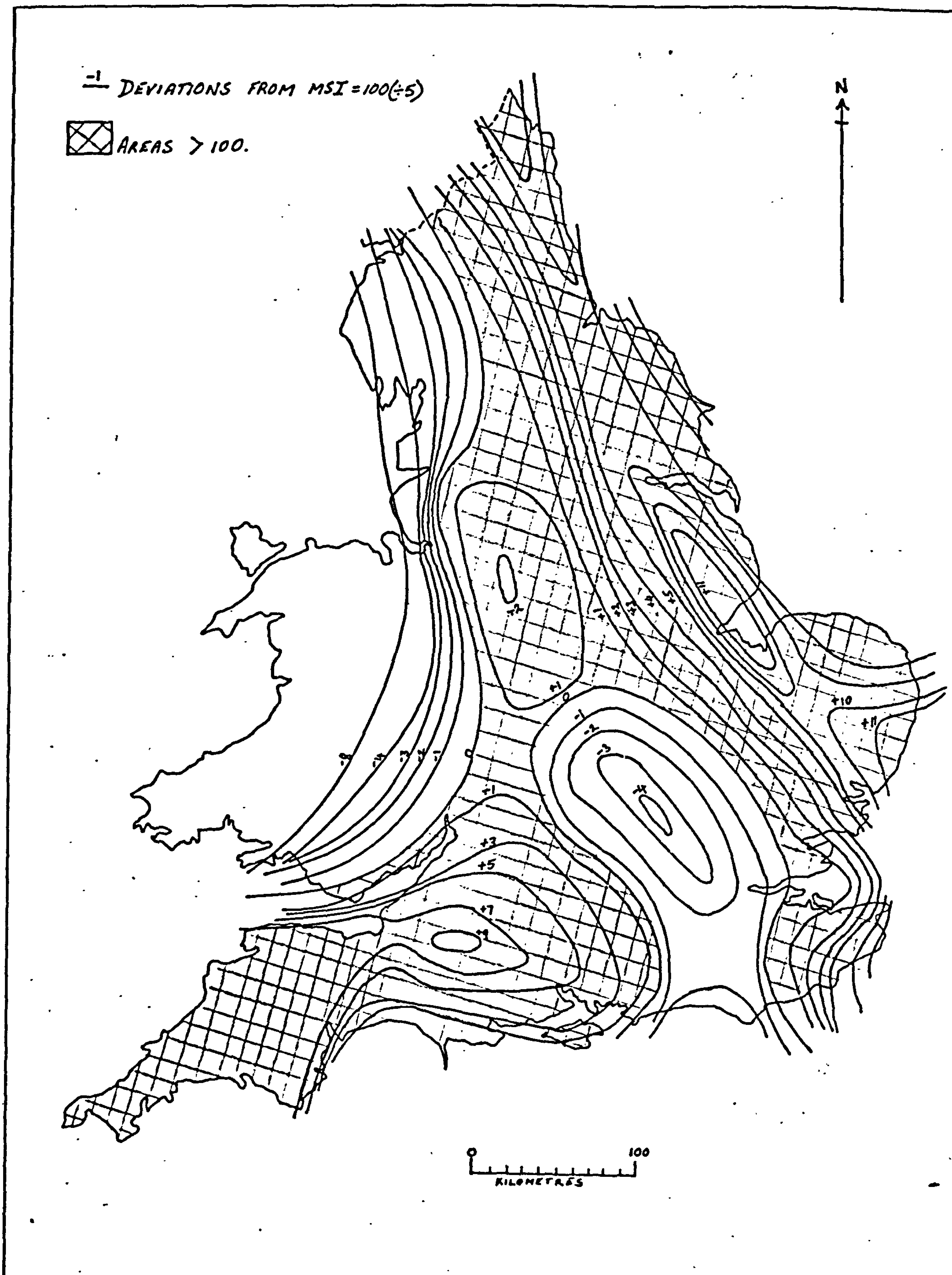


Fig. 9.4.12. SIXTH ORDER TREND SURFACE FOR TDRC70

(3) TOTAL ACCIDENTS PER 10^3 POPULATION TREND SURFACE ANALYSIS 1970 (TAPP7Ø)

If reference is made back to the original regression results for this variable, it will be seen that it was argued at that time that this variable measured the variation in road accidents once the size factor (population) had been removed. Similarly it was also shown that the structural variables accounted for a very low percentage of the total variation within the dependent variable (29%). However, since one of the major independent variables introduced into the regression equation was that of ETT7Ø it was postulated that perhaps a substantial part of the remaining variation could be accounted for by spatial and regional variations, especially those also related to ETT7Ø. This therefore is the main feature of interest in looking at this trend surface analysis.

From the results obtained for TAPP7Ø it can be seen that the trend surface analysis was significant at all levels and by the time the 6th order surface had been reached 43% of the variation in the variable had been accounted for, and the surface as a whole was highly significant. ($95 < F < 99\%$). Similarly, the cubic surface which accounted for 25% of the variation was also significant with the calculated value of F being significant at the 99% level. The cubic surface can therefore be observed as the more simplistic regional trend surface.

The cubic surface which is shown in Fig. (9:4.13) is so organised that the trends noted in the previous two trend surface analyses, are much more emphasised, within this variable. Thus the low trough

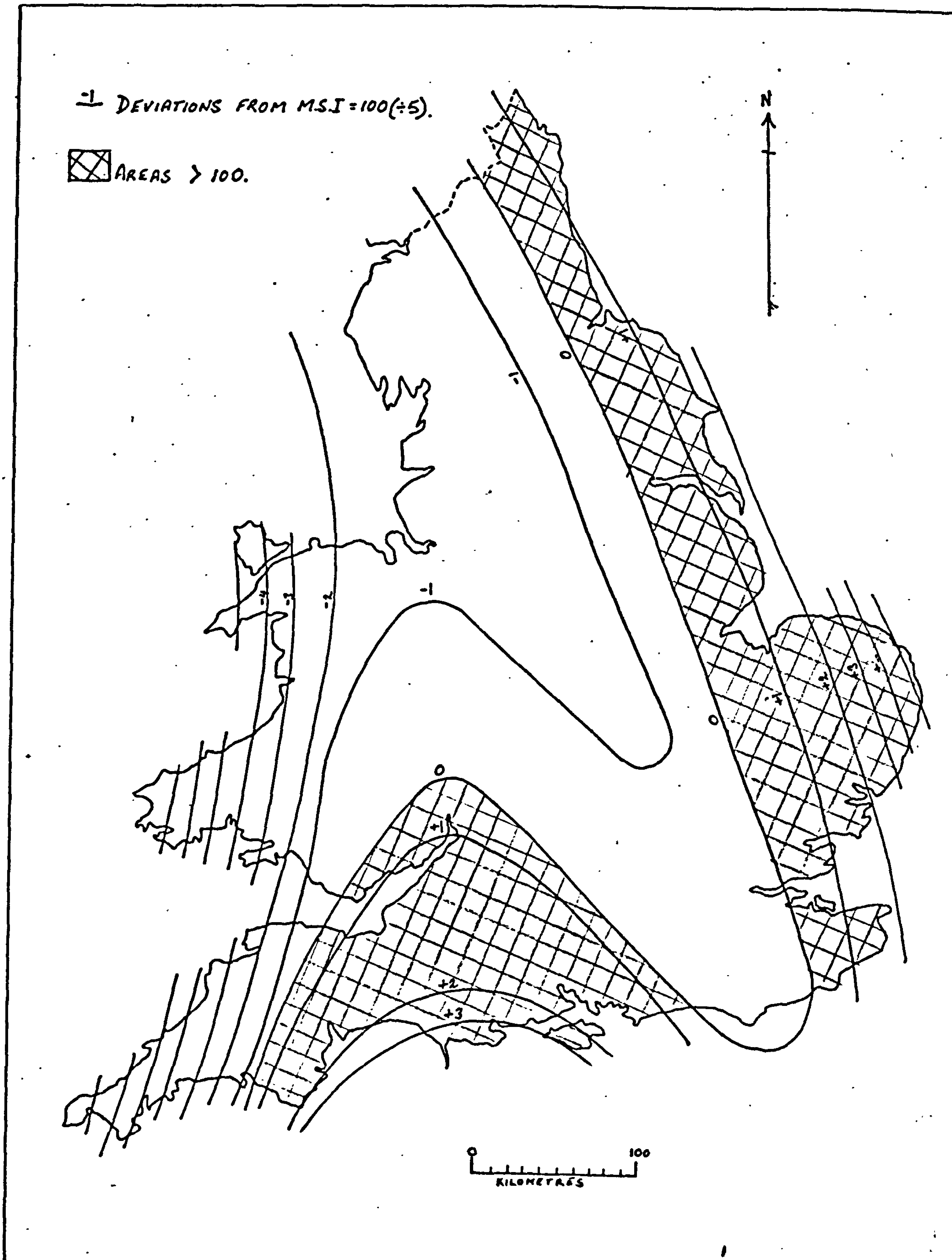


Fig. 9.4.13. CUBIC TREND SURFACE FOR TAPP70

which covers a major portion of England is again lined along the N.W. - S.E. axis. Towards the east of this area, the surface rises rapidly towards the east coast region, with considerable rises in East Anglia. There is also a similar rise in the surface from the south midlands into the S.W. peninsula and onto the south coast area. It should be noted, however, that once the surface has moved into the S.W. peninsula itself, the surface again begins to fall fairly rapidly. Thus once again the features of this surface which should be noted are those which can be related to the major transport network system, with the main emphasis being on the unsatisfactory surface rises in an east-west direction and also as one approaches the entrance to the S.W. peninsula.

Observation of the residuals from this surface meanwhile shows various features which must be accounted for by any higher order surfaces (Fig. 9.4.14).

(i) Under-prediction by the cubic surface for areas in the eastern half of England, including the south Yorkshire conurbations.

(ii) Under-prediction of most of the areas around London.

(iii) Under-prediction of the areas approaching the south-west and also the south midlands.

This predominance of underestimation by the surface is an indication of the difference between areas of high MSI values and those of low MSI values, which obviously cannot be accounted for by the cubic surface. Similarly it would also seem to be a direct

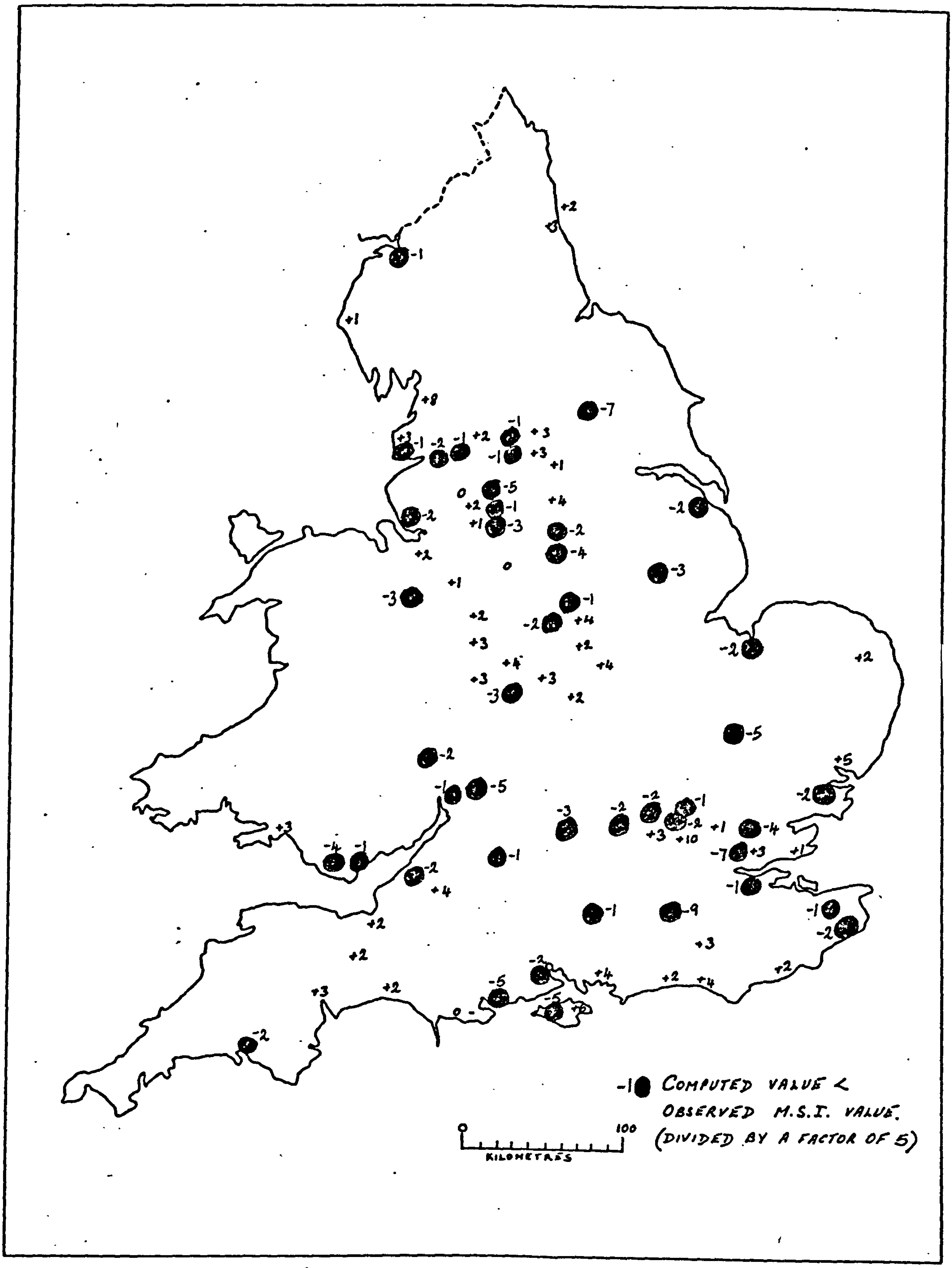


Fig. 9.4.14. DEVIATIONS FROM THE CUBIC TREND SURFACE FOR TAPP70

indication of the impact of any spatial variable such as road transport.

Accordingly, the 6th order surface which is shown in Fig. (9.4.15) is considerably more complex although the major trend, the N.W. - S.E. trough still remains as the central theme of the surface. The major modifications, as can be seen, are the introduction of new "highs" over central Wales to account for the high values for the industrial regions of South Wales, in particular Cardiff and Barry, and also over the eastern coastal region, in order to increase the high values noted for this region of England. Although this surface greatly increases the level of explanation for TAPP7Ø it still fails to account for considerable variation as can be seen from the map of residuals for this surface. However, it must be presumed that these remaining variations are due to local components within the regional trends, and as such cannot be explained within these trend surface analyses. An indication of these local components can clearly be seen in the residuals for both the Lancashire and Yorkshire conurbations, where there is neither a predominance of negative nor positive values which would permit further regional explanation for these areas.

The conclusion to be drawn from this analysis therefore is that the postulates made regarding the importance of through traffic in the discussion on the original regression equation, seem to be perfectly valid. However, it would seem also that

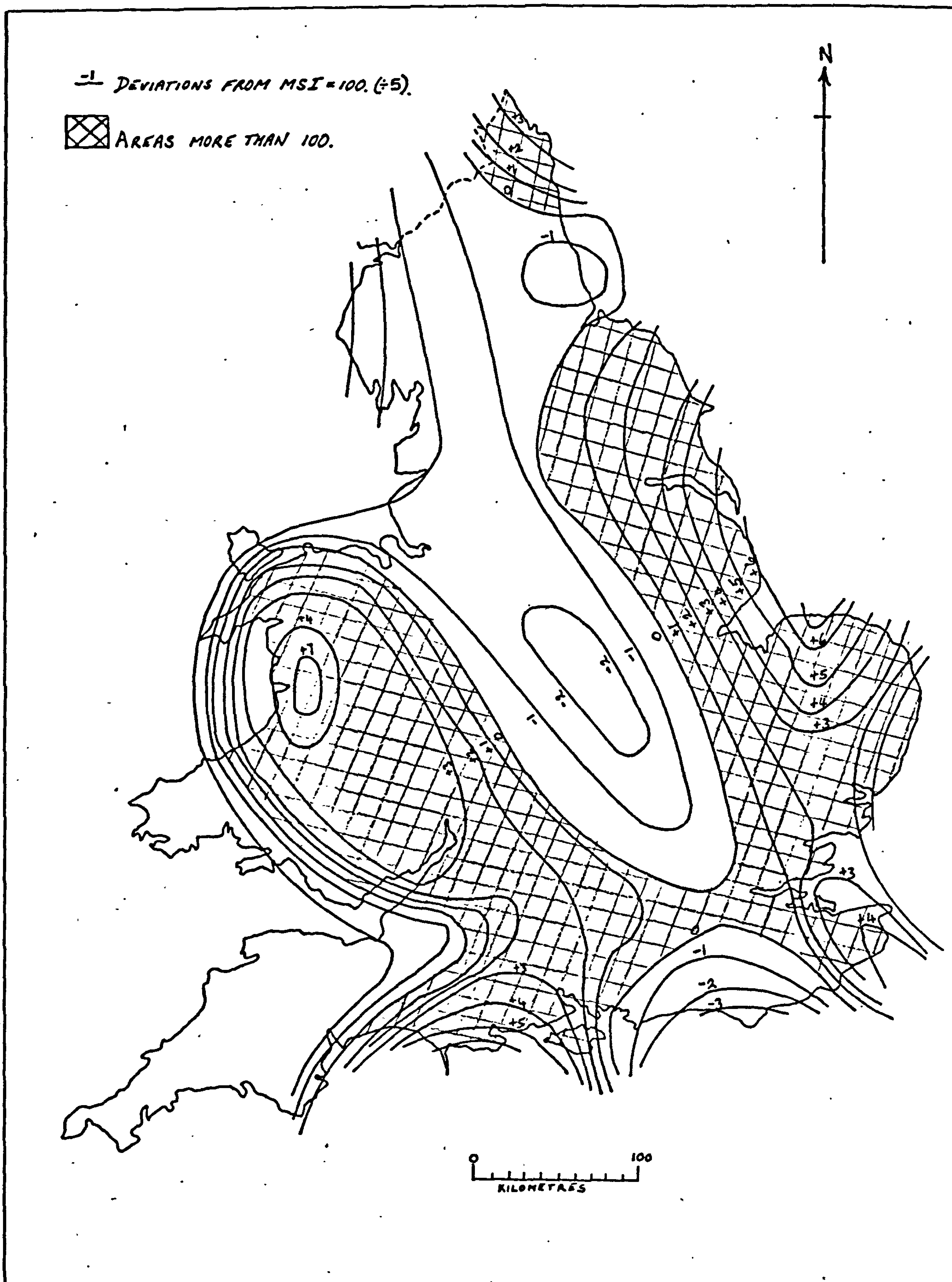


Fig. 9.4.15. SIXTH ORDER TREND SURFACE FOR TAPP70

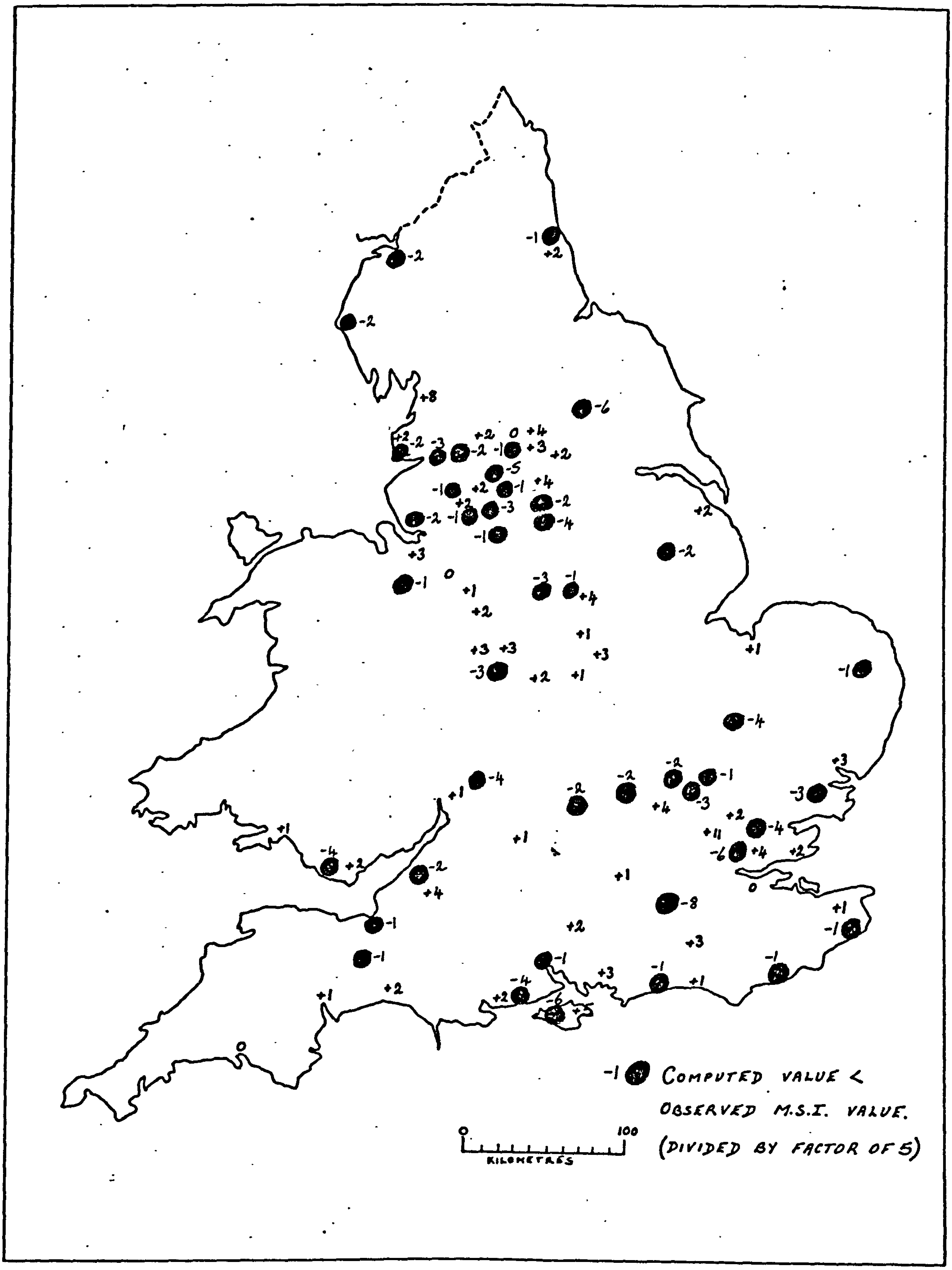


Fig. 9.4.16. DEVIATIONS FROM THE SIXTH ORDER TREND SURFACE FOR TAPP70

most of the remaining variation in TAPP7Ø can only be related to local factors, and as such cannot be accounted for within this analysis. Therefore once again the situation is reached where further understanding of the problem must arise from local studies of areas which can be placed generally within the context of this research.

(4) PEDESTRIAN CASUALTIES PER 10^2 WORK WALK TRIPS TREND ANALYSIS 1970 (TPEDWT7Ø)

It was stressed whilst looking at the results of the regression analysis that the volume of pedestrian casualties appeared to be dependent more upon local factors and arrangements, rather than regional or national factors and trends. The results of these trend surface analyses would seem to vindicate this line of argument since this variable is the only pedestrian variable which can produce a surface which is significant at any confidence level. From the relevant tables, it can be seen that at the 5th order surface, this surface accounts for approximately 34% of the total variation in TPEDWT7Ø and is significant at the 95% level. It must be remembered however, that this variable is somewhat experimental in nature due to the fact that the actual volume of pedestrian casualties was modified in order to take account of pedestrian exposure. The volume of casualties was related to the number of work walk trips as given by the journey to work tables in the 1966 Census. Therefore it would be naive to assume that this variable is without error, but then again it would seem reasonable to expect this variable to give some degree of indication in the relative risk of pedestrian casualties in different areas.

The 5th order surface for TPEDWT7Ø is shown in Fig. (9.4.17) and the most obvious feature of this figure is the low values derived for the midlands region of England, centred on Birmingham and Coventry. In some wasys such an area of low values is quite a surprise since with this region's high car ownership values one would have expected the relative risk to pedestrians to be somewhat higher in these areas. The explanation to this seeming anomaly would appear to be that because of the high levels of car ownership, it has become necessary in these areas to find ways of accommodating and integrating these extra cars into the urban system, whilst also protecting both the environment and the pedestrians. The result has been the creation and building of better transport network systems in these areas, and the simultaneous creation of pedestrian areas and facilities, in order to maintain the principle of single purpose thoroughways.

A second feature of this trend surface analysis is the way all the major coastal resort areas are characterised by rising surfaces. The only exception to this role would seem to be along the S.E. coast where the surface is falling. However, observation of the residuals map shows that this area includes negative values along the coastal area, indicating that these areas are being under-estimated. Whilst this is really of little surprise it does indicate that these resort areas should spend more time and money on the protection of the pedestrians in these areas, who after all, are the main source of revenue for these towns.

Finally, as regards this variable, the remaining point of interest is related to the deviations map for this 5th order surface

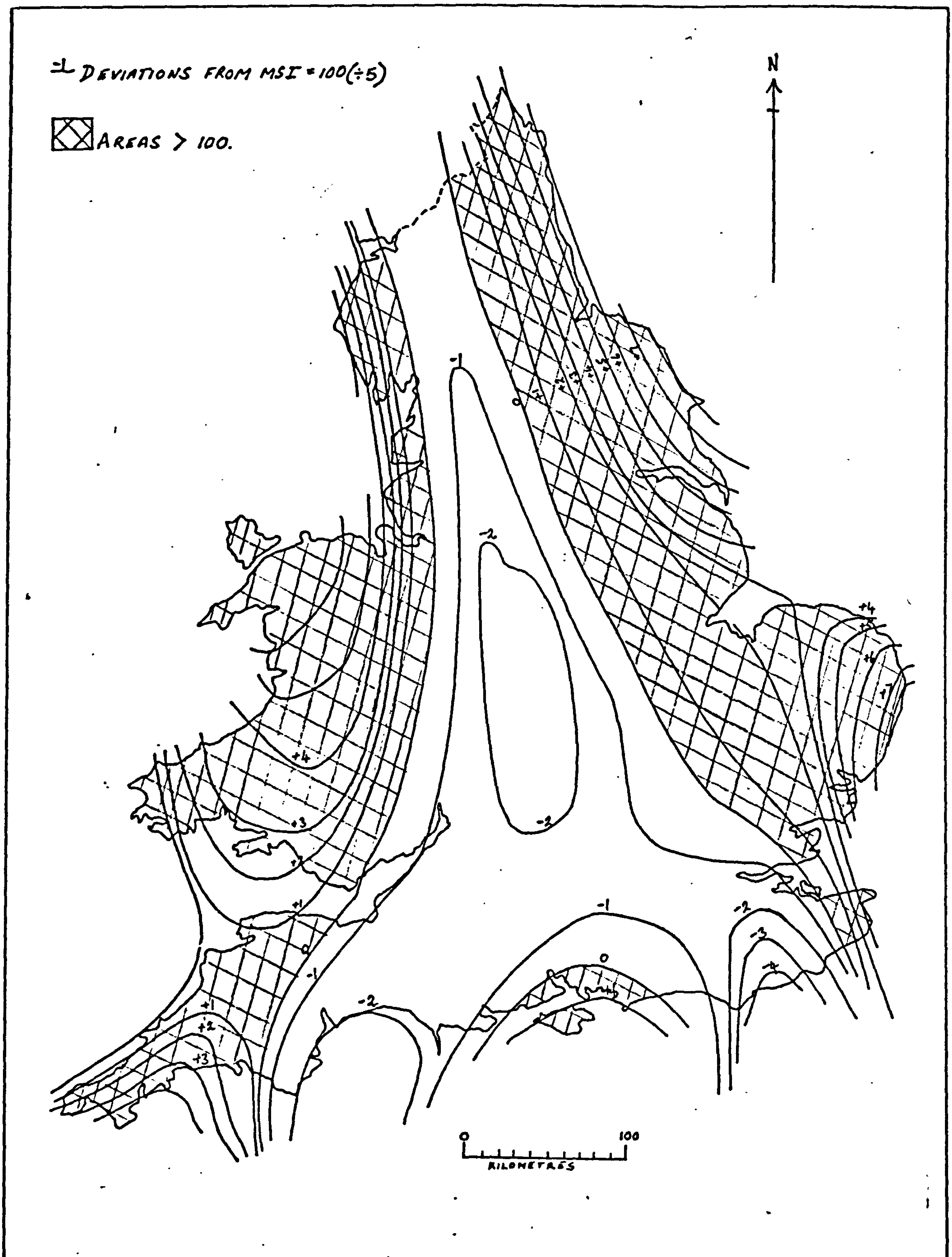


Fig. 9.4.17. FIFTH ORDER TREND SUFACE FOR PEDWT70

(Fig. 9.4.18). From this map one can see that there is a high negative group of residuals around the edge of London. Once again therefore it would seem that the derived regional trend is unable to account for the areal impact of the London conurbation, consequently in most instances, the regional trend will underestimate values in these areas.

Summarising therefore, although this variable must be treated with some caution, it would seem that three important facts can be derived from this surface:-

(i) The region which has seemed to cope best with regards to the traffic/pedestrian conflict, and where the relative risk to pedestrians is lowest, is the Midlands region. Therefore it would seem reasonable to advocate studies of these towns for a further understanding of this problem.

(ii) As regards relative pedestrian risk, it would seem that the coastal, and resort towns, have much that needs doing in order to reduce the level of pedestrian vulnerability.

(iii) The area surrounding London also has a high pedestrian risk value, and this includes the ring of new towns and garden city areas.

The trend surfaces which have been described individually in this report have pinpointed the major features of the spatial analysis. The remaining trend surfaces either merely emphasise the trends already indicated, or alternatively are insignificant according to the analysis of variance test. However, where the

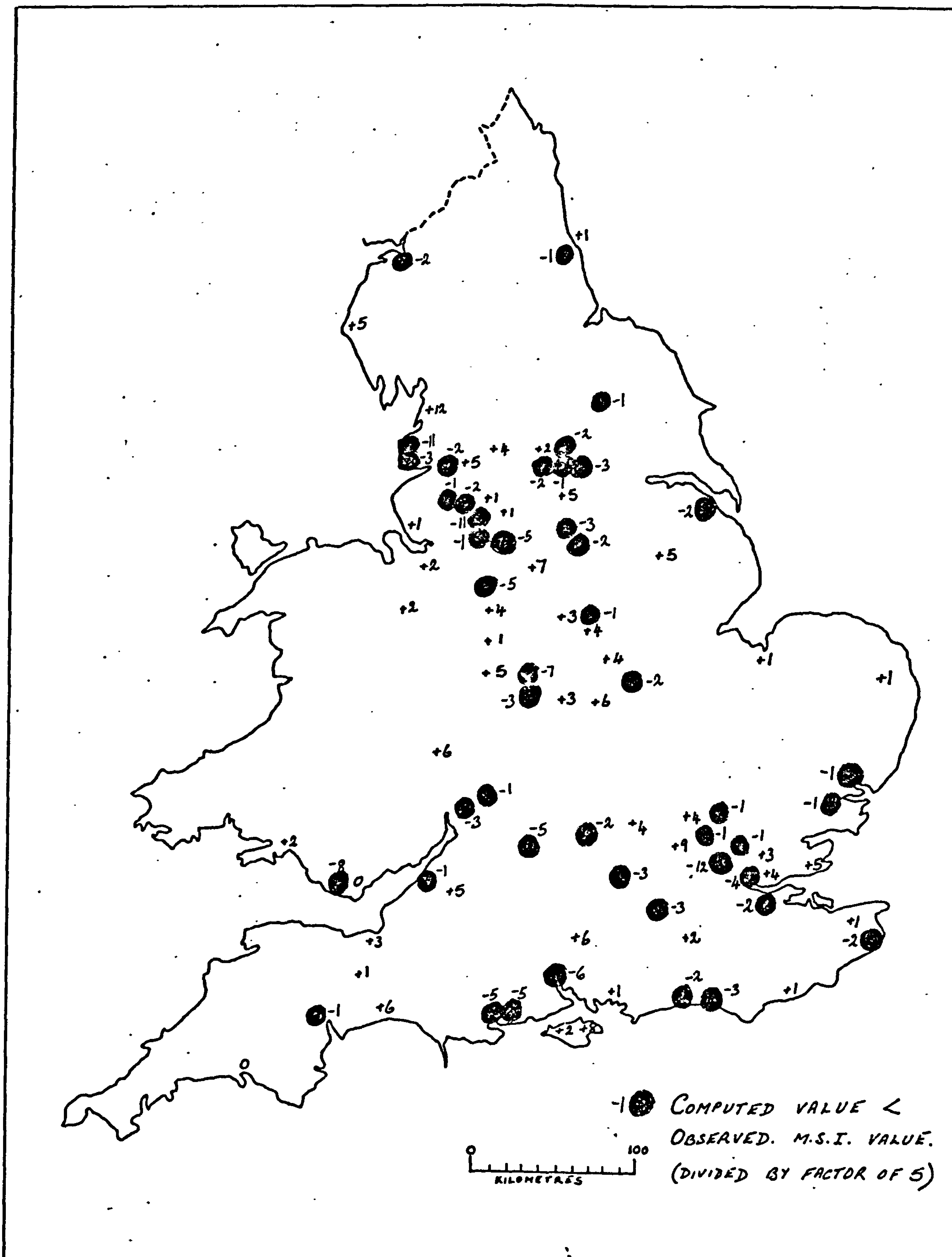


Fig. 9.4.18. DEVIATIONS FROM THE FIFTH ORDER TREND SURFACE FOR PEDWT70

trend surfaces have proved significant at any surface, the relevant maps are shown in Appendix 2, and for further information as regards these surfaces the reader is referred to that section of this report.

Before leaving this section of the study of the spatial variations it is pertinent at this point to look at the results of the calculation of the Mean Safety Index itself.

One of the major reasons for looking for, and at, these spatial variations was to look at regions and see if there were any specific problems relevant to that one region. However, at the same time it is also possible to group towns according to their response to any variable in a non-spatial sense. That is, towns can be grouped according to their actual MSI value. Once again the logic behind this approach is to see if certain groups of towns have similar problems which can be solved accordingly.

Two methods of grouping the 100 sampled areas were tried, one which simply used the value of the MSI, and thereby placed each area into a predetermined group, and the other which proved less reliable, and which made use of a Q-mode factor analysis.

According to the first mode of analysis, it would have been possible to derive a grouping system for every regression equation. However, this was considered excessive and accordingly a grouping system was attained for only five regressions, which display the absolute levels of conflict and also some degree of exposure. Thus the five regression variables which were grouped according to their MSI values were:-

1) Total Casualties 1970

2) Total Fatal and Serious Casualties 1970

3) Total Pedestrian Casualties 1970

4) Total Casualties for 10^6 Vehicle-Mile 1970

5) Total Pedestrian Casualties 10^2 Work Walk Trips 1970

An additional grouping which was also included was that of the calculated MSI values from the trend surface analysis on the variable TC7∅. The logic here was that using these MSI values would remove the external through traffic variable variation, thereby allowing a more realistic assessment of the group problems of each urban area. These groups are shown in Tables (9.4.19) to Table (9.4.23).

As can be seen from these tables, the distributions about 100.00 MSI value, appear almost normally distributed, which is only to be expected considering their derivation. Therefore the groups (and constituent parts) which are of most interest are those which occur at either end of the distribution, and are therefore showing extreme responses to each of the regression variables. Thus as regards this grouping procedure, the important groups were deemed to be those with MSI values greater than 120.00, and less than 80.00. In the following section therefore, a quick summary of these groups will be given according to the individual regression variable.

1) TOTAL CASUALTIES 1970 - (MSI GROUPINGS)

The urban areas grouped within the high MSI groups for this variable seem to be characterised by their very size. All three

50	50 - 60	60 - 70	70 - 80	80 - 90	90 - 100	100 - 110	110 - 120	120 - 130	130 - 140	140+
Horsham	Lancaster	Canterbury	Chester	Brighton	Aylesbury	Barnsley	Blackburn	Barry	Chesterfield	Buxton
Ryde		Corby	E.Kilbride	Burnley	Basildon	Bolton	Brentwood	Bournemouth	Cumbernauld	Cheshun
		Dartford	Stretford	Durham	Bath	Bristol	Cheltenham	Bridgewater	Yeovil	Guildfo
		Glenrothes	Sutton Cold.	Farnworth	Birmingham	Cambridge	Colchester	Grimsby		Inverne
		Hastings	W. Bridge-	K.Lynn	Blackpool	Cardiff	Dundee	Norwich		Pontefra
		Whitehaven	ford	Newport	Bradford	Crewe	Gloucester	Rochdale		
				Stafford	Carlisle	Dover	Hereford	Taunton		
				Swansea	Chelmsford	Huddersfield	Lincoln	Welwyn		
				Winchester	Coventry	Ipswich	Manchester			
					Derby	Luton				
				Wolverhampton	Dewsbury	Portsmouth	Poole			
				Worthing	Edinburgh	Reading	Sheffield			
				Wrexham	Exeter	Rugby	Southampton			
					Harlow	Southend	York			
					H.Hempstead	Stevenage				
					Kirkcaldy	Stirling				
					Leeds	Thornton Clev.				
					Leicester					
					Liverpool					
					Nottingham					
					Oldham					
					Oxford					
					Plymouth					
					Preston					
					Salford					
					Stoke					
					Sunderland					
					Swindon					

TABLE 9.4.19 MEAN SAFETY INDEX GROUPINGS FOR TC70

groups seem to be mainly composed of the lower sized urban area which also are towns which lie upon a major road traffic route. Thus in the top group one finds the towns Buxton and Guildford. A further characteristic of this top group (with perhaps the exception of Inverness) is that these towns are mainly dormitory towns around the major conurbation areas of London, Yorkshire and Lancashire.

The 120.00 - 130.00 group meanwhile has a more varied set of towns than the other groups, with the result that it is almost impossible to find any general characteristic for the whole group. Thus one finds that such towns as Bournemouth and Bridgewater are in the same group.

At the other end of the scale there is even less uniformity between the various groups, although once again it is obvious that only small to average sized towns are included within these values. Amongst the towns contained within these groups two towns seem to stand out as unexpected constituents. These are Lancaster and Chester, with the former having the lower value of the two. One tentative hypothesis as to these low values could be the ring road system which each has developed recently, removing unnecessary traffic from regions of the town centre. However, such hypotheses can only be proved after further research along these particular lines.

Before leaving this variable it should be pointed out that within these low value groups, there are two new town developments and one town (Corby) which has also been expanded under the aegis

L50	50 - 60	60 - 70	70 - 80	80 - 90	90 - 100	100 - 110	110 - 120	120 - 130	130 - 140	140+
Cumbernauld	Barnsley	Brighton	Bath	Basildon	Birmingham	Aylesbury	Bournemouth	Bradford	Cardiff	Barry
Dewsbury	Burnley	Cambridge	Blackburn	Blackpool	Canterbury	Bristol	Brentwood	Guildford	Chelmsford	Bridgewater
Durham	Leicester	Chester	Huddersfield	Bolton	Chesterfld.	Edinburgh	Colchester	Harlow	Corby	Buxton
	Oldham	Crewe	Stoke	Carlisle	Exeter	Hastings	Coventry	Oxford	Derby	Cheltenham
		E.Kilbride	Sutton Cold.	Cheshunt	Lancaster	Ipswich	Gloucester		Dundee	Dartford
		Pontefract	Swansea	H.Hempstead	Liverpool	Leeds	Lincoln		Inverness	Dover
		Stafford	Wolverhampton	Kirkcaldy	Luton	Preston	Norwich		Taunton	Farnworth
		Stretford		Rochdale	Manchester	Rugby	Nottingham			Glenrothes
				Sheffield	Newport	Southampton	Portsmouth			Hereford
				Stevenage	Plymouth	Southend	York			Horsham
				Swindon	Poole					K. Lynn
				Welwyn	Reading					Ryde
					Salford					Stirling
					Sunderland					Thornton Clev.
										W.Bridgeford
					Worthing					Whitehaven
										Winchester
										Wrexham
										Yeovil

TABLE 9.4.20. MEAN SAFETY INDEX GROUPINGS FOR TFS70

of a Development Corporation. This point however, will not be developed at this stage in the report, but is left until a later section dealing solely with the new town developments.

(2) TOTAL FATAL AND SERIOUS CASUALTIES 1970 (MSI GROUPINGS)

Once again the groupings noted within this variable prove very difficult to interpret although there are several features which are deserving of note, simply because of their appearance at any one location.

The first thing that should be said about these groupings is that the distribution is no longer normal, and the distribution of urban areas of differing size is much more random within the distribution. Thus Leicester and Barnsley are amongst the lowest values, whilst Cardiff and Derby are amongst the highest values. However, it should be noted that the highest group once again mainly includes only small and medium sized urban areas.

If individual areas are observed further interesting information can be derived. Thus if the location of Horsham and Ryde is noted in the groupings for TC7Ø and also for TFS7Ø, it will be seen that whilst both are in the lowest values for all casualties, they are both in the highest group for fatal and serious casualties. Thus these two areas can be said to exhibit low accident frequencies but high severity rates.

Perhaps the only groups which can be identified again within this variable is the highest group with values of more than 140.00. As with TC7Ø this group seems to be made up of towns which lie on

L50	50 - 60	60 - 70	70 - 80	80 - 90	90 - 100	100 - 110	110 - 120	120 - 130	130-140	140+
Lancaster	Glenrothes	Cumbernauld	Barnsley	Basildon	Bath	Bolton	Blackburn	Colchester	Taunton	Aylesbury
	Horsham	K. Lynn	Bridgewater	Cambridge	Birmingham	Bournemouth	Carlisle			Barry
		Ryde	Corby	Chester	Blackpool		Dover			Brentwood
		Stretford	Dartford	Coventry	Bristol	Bradford	Dundee			Buxton
			Dewsbury	Derby	Brighton	Burnley	Manchester			Canterbury
			E.Kilbride	Durham	Chelmsford	Cardiff	Newport			Cheshunt
			Farnworth	H.Hempstead	Cheltenham	Chesterfield	Norwich			Guildford
			Hastings	Kirkcaldy	Crewe	Edinburgh	Oldham			Inverness
			Southend	Lincoln	Gloucester	Exeter	Rochdale			Pontefract
			Whitehaven	Oxford	Harlow	Hereford	Sheffield			Sutton Cold.
				Portsmouth	Ipswich	Huddersfield				Thornton Clev
				Wolverhampton	Leeds	Grimsby	Sunderland			
					Luton	Leicester				Welwyn
				Worthing	Plymouth	Liverpool				W.Bridgeford
					Rugby	Nottingham				
					Stafford	Poole				Winchester
					Stevenage	Preston				Yeovil
					Stoke	Reading				
					Swansea	Southampton				
						Stirling				
						Swindon				
						York				

TABLE 9.4.21. MEAN SAFETY INDEX GROUPINGS FOR TPED70

major trunk roads and are themselves relatively small. Accordingly such towns would seem to be unable to cope with the excessive demands made upon them by the increased traffic flows. Similarly, since this through traffic also tends to be of the high speed nature, such accidents tend to cause casualties with higher degrees of severity.

(3) TOTAL CASUALTIES PER 10^6 VEHICLE MILES (MSI GROUPINGS)

When the exposure variable for total casualties is introduced, the groupings of urban areas appears totally haphazard. However, it should be recalled that this variable within the regression model showed a considerable degree of lack of fit, especially at the higher and lower values. Consequently it is of no surprise when one finds a considerable number of large towns within the highest value group. Thus although it is impossible to make any statements about these groupings, it is obviously significant when some large towns finish with low MSI values, and in such cases these towns are identified as being in need of special case studies. Towns of this kind in the lowest group of MSI values include Birmingham, Sheffield and Dundee.

(4) TOTAL PEDESTRIAN CASUALTIES (MSI GROUPINGS)

The groupings attained with this variable are very comparable to those obtained for TC7Ø and TFS7Ø with just modifications for individual towns. Thus Salford which has a high grouping for this variable, is below the 100.00 mark for all casualties, with the result that there must be a relatively high number of pedestrian casualties for that area.

The problem with this grouping, and the variable TPED7Ø in general is that without some measure of exposure, very misleading

L50	50 - 60	60 - 70	70 - 80	80 - 90	90 - 100	100-110	110 - 120	120 - 130	130 - 140	140+
Birmingham Swansea	Basildon Chesterfield Dundee Durham Lancaster Sheffield Stretford	E.Kilbride H.Hempstead Oldham Poole Preston Stafford Sutton Cold. Thornton Cleveleys	Burnley Bath Chester Dewsbury Inverness Rochdale Southend Swindon	Aylesbury Brighton Burnley Carlisle Chelmsford Crewe Harlow Hastings Huddersfield Luton Oxford Rugby Stevenage Whitehaven Yeovil	Bournemouth Brentwood Cambridge Cheltenham Exeter Hereford Horsham Grimsby Ipswich Kirkcaldy Lincoln Reading Stirling Stoke	Colchester Corby Farnworth Gloucester K. Lynn Ryde Winchester Worthing Wrexham York	Bolton Bridgewater Canterbury Dartford Norwich Taunton Welwyn	Blackpool Cumbernauld Guildford Newport Pontefract Salford	Bury Buxton Portsmouth W.Bridge-ford	Blackburn Bradford Bristol Cardiff Cheshunt Coventry Derby Dover Edinburgh Glenrothe Leeds Leicester Liverpool Mancheste Nottingha Plymouth Southampt Sunderlan Wolverhampton

TABLE 9.4.22. MEAN SAFETY INDEX GROUPINGS FOR TCVM70

conclusions can be made. For example where the environment has been made more conducive to pedestrian movement, more journeys may be made on foot. The result on pedestrian casualties could be to increase the absolute volume of casualties thereby implying that the safety level for pedestrians had been reduced. In actual fact, however, considering the increase in the volume of pedestrian movement, it could be that the true change in the safety level for pedestrians is upwards. Therefore, when considering urban groupings it is best to consider those for total pedestrian casualties in association with the pedestrian exposure variable given below.

(5) TOTAL PEDESTRIAN CASUALTIES PER 10^2 WORK WALK TRIPS (MSI GROUPINGS)

The major impact of introducing the exposure level is seen amongst some of the larger urban areas, with the majority of the smaller towns remaining in their original MSI group for the absolute variable. This movement within the larger towns works in both directions with some increasing and others decreasing. Examples of those areas lowering their MSI grouping are Leicester, Coventry, Bradford, Oldham, Blackburn etc., whilst those raising their MSI groupings include Norwich, York and Bournemouth. It must be remembered however, that the exposure introduced in this variable only includes "work trips" and most of the towns which move to higher MSI groups are those who exhibit high pedestrian movements for non-work trips. Thus these towns include the coastal resort towns and also the historic towns, which one would expect to have high pedestrian movements at all times of the day.

L50	50 - 60	60 - 70	70 - 80	80 - 90	90 - 100	100 - 110	110 - 120	120 - 130	130-140	140+
Lancaster	Buxton	Bath	Aylesbury	Basildon	Birmingham	Cardiff	Blackpool	Bournemouth	Dover	Barry
	H.Hempstead	Rugby	Barnsley	Bradford	Bolton	Carlisle	Brentwood	Cambridge	Grimsby	Cheshunt
		Ryde	Blackburn	Bridgewater	Brighton	Chesterfield	Canterbury	Colchester	Ipswich	Cumbernauld
		Wolverhampton	Burnley	Chelmsford	Bristol	Corby	Crewe	Norwich		
		Yeovil	Coventry	Derby	Cheltenham	Dartford	K. Lynn	Southampton		Durham
			E.Kilbride	Dundee	Chester	Glenrothes	Manchester	Sutton Cold.		Salford
			Hastings	Horsham	Dewsbury	Gloucester	Pontefract	W.Bridgeford		Sunderland
			Hereford	Lincoln	Edinburgh	Guildford	Poole	York		
			Inverness	Oldham	Exeter	Harlow	Swindon			Thornton
			Leicester	Rochdale	Farnworth	Leeds				Cleveley
			Luton	Southend	Huddersfld.	Liverpool				
			Stoke	Stafford	Kirkcaldy	Preston				
			Winchester		Newport	Reading				
					Nottingham	Sheffield				
					Oxford	Stirling				
					Plymouth	Swansea				
					Portsmouth	Welwyn				
					Stevenage	Wrexham				
					Stretford					
					Taunton					
					Whitehaven					
					Worthing					

TABLE 9.4.23 MEAN SAFETY INDEX GROUPINGS FOR TPEDWT70

The second grouping method used involved the application of a Q-mode factor analysis technique upon the 20 groups of MSI values calculated for the trend surface analyses. In the Q-mode factor analysis the data matrix is collapsed so that the rows are combined rather than the columns as in the normal R-mode factor analysis model. The result therefore in this context was to combine the urban areas according to their constituent MSI values. However, as was mentioned earlier the results from this approach proved less reliable for two basic reasons;-

1) The MSI values are such that several variables will create totally opposite MSI values due to the nature of the variable being used. Thus the composite factors which were derived would be almost average values of the MSI range.

2) Similarly little confidence can be placed in a factor analysis solution when there is such a divergence between the number of columns and rows in the data matrix when the rows are being combined. In such a situation the optimum kind of data matrix is obviously a square matrix.

Consequently the results obtained by this technique are considered to be of somewhat dubious value and relevance. For that reason no results are presented here and the method is left as a possible approach for future research.

The conclusions from these sets of urban groupings, therefore, are only very tentative, especially since it is only marginal to the main body of this research. However, the results do seem to suggest that there is some future in this line of research if, and when some more sophisticated grouping techniques are used. However, the greatest use at present simply lies in the identification of certain groups and constituent urban members. Further developments must be the more detailed study of individual groups and also individual urban areas within that group. With this thought in mind, therefore, it is worthwhile to leave this section of the analysis and look at the time series analysis.

9.5 THE TIME SERIES ANALYSIS

The methodology used in this analysis has been described already in Chapter 8 of this report. However, in order to simplify matters, a simple reiteration of the procedure will be applied here.

Three variables; TC, TCPP and TCVM, were studied in relation to their differing responses from 1966 to 1970. Comparison between years was performed in two ways, 1) Comparison of the overall regression explanation levels, 2) The varying importance of individual regression variables was studied by means of comparing the relevant coefficients. Since however, the results of the principal components analysis had differed for every year, in order to attain a reasonable comparison, it was necessary to return to the original data matrix. From this matrix variables were extracted which had been deemed significant by the step-wise regression procedure, and where necessary were subjected to a logarithmic transformation. The results of this temporal analysis will be presented below according to each of the three variables studied.

(i) TOTAL CASUALTIES TEMPORAL ANALYSIS

The five significant variables used in the analysis for this dependent variable are shown in Table (9.5.1), along with the derived beta coefficients (standardised b coefficients) for the years 1966 - 1970. Further results shown in Table (9.5.1) are for the regression equation containing the transformed variables, which were necessary because of the lack of symmetry in the original variable distributions.

As can be seen from this table, the overall explanation level remained approximately consistent over this time period with a steady increase in the absolute r^2 value from 0.954 in 1966 to 0.982 in 1970, for the non-transformed regression model. As regards the transformed regression meanwhile, the range for the r^2 value is much less, with the peak of 0.967 appearing in 1970 and the minimum value being 0.951 in 1968. Thus it can be seen that whichever model is used the actual variation in the explanation level is very small and seemingly insignificant. Consequently since the predominant variable in all of these regressions is the POP variable, one must presume that over this time period the major contributing factor in the variation in road casualties between different urban areas is the actual size of these urban areas, and the degree of interaction within these areas. The importance of this variable can be further qualified by looking at the actual value of the beta coefficients. Thus in 1970, POP7Ø had a beta coefficient of 0.980 as related to the second most important variable PPI7Ø, which had a value of 0.048.

In order to look at the temporal significance of the individual coefficients meanwhile, it is necessary to construct confidence

VARIABLES	BETA COEFFICIENTS				
	1970	1969	1968	1967	1966
POP70	0.980	0.953	0.969	0.963	0.958
SRI70	-0.013	-0.009	-0.032	-0.003	0.001
CI70	-0.041	-0.062	-0.053	-0.051	-0.042
DVI70	0.026	0.037	0.041	0.028	0.030
PFI70	0.047	0.082	0.038	0.045	0.049
X_o	682.486	601.024	1429.217	455.132	139.120
R^2	0.982	0.953	0.969	0.959	0.955
CI70	-0.017	-0.019	-0.035	-0.024	-0.004
LNPOP70	0.975	0.956	0.946	0.957	0.945
LNSRI70	-0.014	-0.004	-0.039	-0.003	-0.004
LNDVI70	0.027	0.050	0.095	0.065	0.063
LNPFI70	0.028	0.073	0.032	0.035	0.064
X_o	-2.350	-2.487	-1.672	-2.092	-2.272
R^2	0.967	0.962	0.951	0.965	0.962

TABLE 9.5.1. TIME SERIES ANALYSIS FOR SIMPLE AND TRANSFORMED VARIABLES IN RELATION TO TOTAL CASUALTIES.

regions around these coefficients. In order to do this, one can make use of the simple t test such that the confidence region is calculated by appying the formula,

$$c.f = b_i \pm S.E(b_i).\ t_{(n)}\ \alpha = 0.005\ (1\%) \quad \text{with}$$

$$t(n - (k + 1)\ \text{degrees of freedom.}$$

where,

b_i is the calculated value for the relevant b coefficient

$S.E(b_i)$ = the standard error of the b coefficient

n = the number of observations

k = the number of independent variables.

Thus for the total casualties variable $t_{(94)}\alpha = 0.005 = 2.62$, and $\alpha = 0.025 = 1.98$.

The confidence regions for all the independent variables in both the transformed and non-transformed regression models, are shown in Table (9.5.2) for 1970. The null hypothesis being tested in this table is that the values for the b coefficient for 1970 and 1966 are drawn from the same population, and consequently do not differ significantly. Thus it can only be stated that the values of these coefficients are significantly different if the 1966 value lies outside the confidence region for that same coefficient for 1970. At the same time it is also possible to test the individual significance of each variable within the regression by testing whether the confidence region includes the value of zero. If zero does fall within the confidence region then no confidence can be placed on the level or direction of the relationship indicated

VARIABLE		B. COEFFICIENT (1970)	STANDARD ERROR (b)	99% CONFID. LEVEL	95% CONFID. LEVEL	B.COEFFICIENT 1966
POP		0.006	0.00	0.00*	0.00	0.007
SRI		-252.253	303.14	794.23	600.22	13.703
CI		-600.088	209.48	547.58	414.77	-736.268
DVI		883.423	548.84	1437.96	1086.70	1477.861
PPI		7.881	2.35	6.15	4.65	9.611
TRANSFORMED	CI	-0.107	0.12	0.314	0.238	-0.029
	LNPOP	1.069	0.02	0.052	0.040	1.041
	LNSRI	-0.499	0.79	2.070	1.564	-0.150
	LNDVI	0.169	0.14	0.367	0.277*	0.465
	LNPPI	0.099	0.07	0.183	0.139	0.230

* Significant Difference at that Confidence Level.

TABLE 9.5.2. CONFIDENCE INTERVALS FOR THE b. COEFFICIENTS WITHIN THE TIME SERIES ANALYSIS OF TOTAL CASUALTIES

by the b coefficient, since the coefficient cannot be said to differ significantly from zero.

The results as indicated by this table reveal that for the two regressions TC7Ø and TC66, only one variable POP, differed significantly between these two years. Similarly only four variables appeared to be significantly different from zero, these being POP, CI, PPI[17] and LNPOP. The implications from these results are that whilst the overall explanation level has remained reasonably consistent over the period 1966 - 1970, the impact of the variable POP has changed significantly during this same period. The direction of this change can be ascertained by referring back to the original coefficient values. From these it can be seen that in 1966 the population variable influenced the dependent variable less than in 1970. Thus it can be stated that whilst this independent variable has remained the major determinant of the value for the total casualties variable over this time period, the relative importance has increased with time. Thus in the earlier periods it would seem that other variables played a more important role in explaining the observed variation. Since this was not true of any of the remaining variables in the regression, it becomes necessary to look at other variables which had not been incorporated due to the recommendations of the 1970 step-wise regression. Consequently, full

---

17 It is interesting to note that this variable PPI is significantly different from zero since the direction of this variable is positive. This however, is totally opposite to the results obtained in the major regression analysis, and one must assume that this is solely the result of the non-normality of this variable in this regression. This conclusion would seem to be supported by the lack of significance in transformed variable LNPPI.

YEAR	ESTIMATE OF THROUGH TRAFFIC				
	B.COEFF-ICIENT	STANDARD ERROR	99% CONFID. INTERVAL	95% CONFID INTERVAL	SIGNIFICANCE LEVEL
1970	2.247	1.359	3.561	2.690	190%
1969	5.589	2.358	6.178	4.669	95%
1968	8.854	1.889	4.949	3.740	95%
1967	6.942	2.310	6.052	4.574	99%
1966	5.538	2.871	7.522	5.684	90%

TABLE 9.5.3. VARIATION IN THE INFLUENCE OF THE INDEPENDENT VARIABLE E.T.T. IN RESPECT TO TOTAL CASUALTIES 1966-1970 - (SIGNIFICANT DIFFERENCE FROM ZERO).

regressions were computed for each of the years and the significance levels of new variables noted; the result being that only one variable appeared to contribute anything to the explanation level, that being the variable ETT. The variations in this variable are given in Table (9.5.3) along with the relative significance of the coefficient values. It can be seen from this that in 1967 the amount of through traffic was significant at the 99% level from whence it has declined until in 1970 it is not even significant at the 90% confidence level. The conclusions to be made here therefore are that, over the period 1966 - 1970 the level of explanation of the dependent variable has remained fairly constant, but the impact of two variables has varied with time. That is as time has progressed from 1966, the impact of the estimate of through traffic variable has decreased whilst the influence of the population variable has increased. Consequently, one must presume that over this period of time, the problem caused by the volume of through traffic in urban areas has been satisfactorily reduced until the impact of this variable is only marginal.

(ii) TOTAL CASUALTIES PER 10^3 POPULATION TEMPORAL ANALYSIS

The same procedure as used for the last variable was applied in this analysis except for the fact that only the years 1970, 1969 and 1966 were calculated because of the unreliability of the data for 1967 and 1968 due to considerable police re-organisation. The relevant data for this dependent variable are given in Table (9.5.4) and Table (9.5.5).

BETA COEFFICIENTS			
VARIABLES	1970	1969	1966
POP	0,079	-0.006	-0.020
DEN	-0.062	-0.035	-0.004
ETT	0.268	0.413	0.277
CT	-0.130	-0.173	-0.053
DOS	0.025	-0.134	-0.043
SH1	-0.348	-0.273	-0.304
DV1	0.079	0.156	0.230
COW	0.203	0.289	0.224
X_o	6.057	5.483	3.725
R^2	0.275	0.364	0.307
DEN	-0.147	-0.072	-0.010
ETT	0.081	0.254	0.168
CI	-0.168	-0.179	-0.047
COW	0.411	0.419	0.349
LNPOP	0.259	0.132	0.005
LNSHI	-0.573	-0.435	-0.465
LNDOS	0.060	-0.131	0.025
LNDVI	0.086	0.148	0.174
X_o	0.681	0.878	0.987
R^z	0.411	0.403	0.320

TABLE 9.5.4. TIME SERIOUS ANALYSIS FOR SIMPLE AND TRANSFORMED VARIABLES IN RELATION TO TOTAL CASUALTIES PER 10^3 POPULATION.

VARIABLE	B. COEFFICIENT (1970)	STANDARD ERROR (b)	99% CONFID. LEVEL	95% CONFID. LEVEL	B.COEFFICIENT 1966
POP	0.000	0.000	0.000	0.000	-0.000
DEN	-0.015	0.03	0.079	0.156	-0.001
ETT	0.018	0.01	0.026	0.020	0.026
CI	-2.094	1.51	3.956	2.990	-1.077
DOS	0.003	0.01	0.026	0.020	-0.007
SHI	-0.016	0.01	0.026	0.020	-0.017
DVI	2.952	3.45	9.039*	6.831*	13.343
COW	1.811	0.91	2.384	1.801	2.531
DEN	-0.003	0.003	0.008	0.006*	-0.000
ETT	0.000	0.0008	0.002	0.002	0.001
CI	-0.211	0.11	0.288	0.218	-0.069
COW	0.285	0.07	0.183	0.139*	0.284
LNPOP	0.057	0.03	0.079	0.059	0. 001
LNSHI	-0.113	0.02	0.052	0.040	-0.107
LNDOS	0.012	0.02	0.052	0.040	0.006
LNDVI	0.109	0.11	0.288	0.218	0.295

* SIGNIFICANT DIFFERENCE AT THAT CONFIDENCE LEVEL

TABLE 9.5.5. CONFIDENCE INTERVALS FOR THE b COEFFICIENTS WITHIN THE TIME SERIES ANALYSIS OF TOTAL CASUALTIES PER 10^3 POPULATION.

The second variable of interest in this analysis is the variable COW. Besides being the second most important variable in the regression series, COW, like LNSHI, would seem to have a varying impact upon the dependent variable with time. Thus from the beta coefficient values one can see that in 1970 COW had a value of 0.411 and a value of 0.349 in 1966. However once again when reference is made to the significance tables, it can be seen that whilst the variable is significantly different from Zero, there is no evidence that the variable has exerted any different influence at different time periods.

The variable ETT is of interest once again because of the increase in the beta value as one moves backwards in time towards 1966. Thus the beta value for 1970 of 0.081 is increased to 0.168 in 1966. However as with the previous two variables, when ETT is noted within the significance tables it can be seen that statistically there is no difference between the values for 1966 and 1970. Consequently it would be incorrect to assume any change in this variable with time. In fact further study of the significance table goes as far as to state that the variable ETT is not even significant within the 1970 regression model even at the 95% confidence level.

Summarising therefore one can say that for the dependent variable TCPP there would seem to have been no statistically significant change in the variables and relationships achieved, during the time period 1966 - 1970, since any change in the coefficient values for any of the independent variables could have occured simply by chance at the 95% confidence level. If it is therefore hypothesised that this dependent variable gives the casualty variation for urban areas once the size component has been removed, it can also be stated that the impact of the structural variables has remained constant over this time period, according to these results.

(iii) TOTAL CASUALTIES PER 10^6 VEHICLE MILES TEMPORAL ANALYSIS

The previous two variables included in this temporal analysis have both belonged to the "static" family of road accident indicators. Therefore one of the further reasons for including this dependent variable in the temporal analysis was to study the impact of time upon an "exposure" variable.

It was pointed out in the section describing the results of the multiple regression that the best results for this variable were obtained when the dependent variable was subjected to a logarithmic transformation. The results in this analysis add further proof to this postulated logarithmic relationship, in terms of the level of explanation. Thus when the non-transformed variables are used the r^2 value increases from 0.462 in 1966 to 0.494 = r^2 in 1970. When both the dependent and independent variables are transformed however, the range of r^2 values is from 0.910 in 1966 to 0.930 in 1970. Consequently attention for this variable is restricted to the transformed regression model.

The consistent increase in the level of explanation from 1966 to 1970 would seem at first to indicate some increase in the explanatory nature of the independent variables used in this regression. However, the results given in the significance tables (9.5.6) and (9.5.7) give no indication of this variation with time. Instead

VARIABLES	BETA COEFFICIENTS		
	1970	1969	1966
POP	-0.503	-0.408	-0.471
DEN	0.094	0.042	-0.014
CI	-0.220	-0.181	-0.149
CHSI	0.126	0.130	0.009
RI	0.063	0.195	0.123
RDI	-0.389	-0.432	-0.349
X_o	58.301	46.741	64.807
R^2	0.494	0.480	0.462
DEN	0.249	0.167	0.285
CI	-0.082	-0.083	-0.054
RI	-0.066	-0.007	-0.035
RDI	-0.274	-0.254	-0.272
LNPOP	-0.972	-0.925	-1.008
LNCHSI	0.077	0.066	0.030
X_o	5.423	5.237	5.606
R^2	0.930	0.912	0.910

TABLE 9.5.6. TIME SERIES ANALYSIS FOR SIMPLE AND TRANSFORMED VARIABLES IN RELATION TO TOTAL CASUALTIES PER 10^6 VEHICLE MILES.

VARIABLE	B. CCEFF ICIENT (1970)	STANDARD ERROR(b)	99% CONFID ENCE LEVEL	95% CONFID ENCE LEVEL	B CCEFF ICIENT 1966
PCP	-0.000	0.00	0.000	0.000	0.000
DEN	0.270	0.49	1.284	0.970	-0.043
CI	-42.361	14.46	37.885	28.630	-34.036
CHSI	0.968	0.61	1.598	1.208	0.073
RI	0.124	0.16	0.419	0.317	0.286
RDI	-0.785	0.32	0.838	0.634	-0.835
DEN	0.019	0.005	0.013	0.010	0.020
CI	-0.413	0.14	0.367	0.277	-0.267
RI	-0.003	0.002	6.005	0.004	-0.002
RDI	-0.014	0.003	0.008	0.006	-0.014
LNPOP	-0.859	0.03	0.079	0.059	-0.862
LNCHSI	0.460	0.13	0.472	0.356	0.173

TABLE 9.5.7. CONFIDENCE INTERVALS FOR THE B, COEFFICIENTS WITHIN THE TIME SERIES ANALYSIS OF TOTAL CASUALTIES PER 10^6 VEHICLE MILES.

these results suggest that both the increase in the r^2 values, and also the general relationships postulated within this regression, remain constant over time, and any changes which may appear are merely the result of chance random variation. This can be stated at both the 99% and 95% confidence level, since all the variables within these regression models fall within the relevent confidence intervals. Although there would seem to be no statistically significant difference between the b coefficients in a temporal sense, the significance tables do show that four of the independent variables are significantly different from zero at the 99% confidence level. Of these variables the major one would appear to be LNPOP which accounts for approximately 50% of the overall variation in the dependent variable, and this compares favourably with the results obtained in the major regression analysis. This is also true for the other significant variables, namely DIN, CI, FDI and LNCHSI (at the 95% level).

The major conclusions to be drawn from these temporal analyses can therefore be summarised as below:-[18]

(i) The relationships between the dependent variables and independent variables noted in the original regression analysis appears to be further justified by the results of the temporal analysis, with only one variable ETT, being inadequately represented in the original TC regression model.

---

18 Similar temporal analyses were carried out upon the data, when disaggregation had been completed for County Borough areas and Non-County Borough areas. Although in several instances there proved to be absolute differences between the coefficient values for variables, there was no statistical significant changes at the 99% confidence level, over the period 1966 to 1970. Therefore the conclusions drawn from the results for the full 100 sampled areas are also

/contd..

(ii) The relationships noted in the major regression analyses appear to remain constant with time and do not significantly differ in 1970 from 1966.

(iii) In terms of the total casualties dependent variable, it would seem that there has been an increase in the importance of the size variable (POP) as time has progressed from 1966 to 1970. This could perhaps indicate an increase in the relative importance of the human sub-system within the overall road accident system.

(iv) Also in terms of the total casualties dependent variable, there would appear to have been a decrease in the relative importance of the variable related to through traffic (ETT) as time has progressed from 1966 to 1970. This would seem to indicate a reduction in the level of problems caused by through traffic volumes during this time period, probably as a result of increased facilities within urban areas for through traffic, and also because of improved external, national transport facilities.

(v) The static nature of the relationships noted in these temporal analyses, would seem to be an indictment of the present

---

18 contd.. applicable to these two disaggregated groups once again stressing the constancy of the road accident problem.

Further analysis was also conducted into the relative response of the two urban area groups within the same time period. Once again the results proved insignificant and it could be concluded that there is no reason to suppose that these two groups were not drawn from the same population. Accordingly there would also appear to be no statistical reason to support the hypothesised disaggregated dichotomous response for these two groups of urban areas.

approach to road safety, and road safety analysis. Thus although the Ministry of Transport notes that, "any hopes that travel on the roads is getting safer are quite mistaken," (HMSO 1967) in their recent advisory publication, "Road Safety - A Fresh Approach," the point that has to be made is, "If travel on the road is not getting safer in spite of the enormous efforts being made by the Ministry and its associated bodies, something is seriously wrong somewhere." (Cohen 1968, p.122).

9.6 THE STUDY OF THE GROWTH OF ROAD ACCIDENT CASUALTIES 1966 - 1970

Although the temporal analysis has enabled several conclusions to be made as regards the constancy of the derived relationships with time, it was also considered desirable to look at the actual rates of growth for these urban areas, in order to assess if any further relationships could be uncovered.

In this section of the research, several hypotheses were put forward and consequently tested. The most obvious of these was, to what degree did the level of road accident casualties in time t + 1 depend upon the level of road accidents at time t. As was mentioned in Chapter 8 of this report, such a hypothesis can be rephrased so that one is testing whether towns of different size and therefore, different road accident levels, have accident growth levels which are proportionally equivalent. This can be simply tested by using the regression model,

$$X_{t+1} = a + bX_t + e^{19}$$

The results for this model are shown in Table (9.5.8) for all

19 Refer to Chapter 8

	TIME PERIOD	INTERCEPT (a)	SLOPE (b)	R^2	S.E.E.	$\bar{Y}$	See/Y %
TOTAL CASUALTIES.	1966/1967	19.1968	0.9363	0.990	115.1	843.5	13.6
	1967/1968	11.6733	0.9011	0.992	88.2	771.8	11.4
	1968/1969	16.9980	1.0422	0.987	120.3	787.4	15.3
	1969/1970	58.7652	0.9470	0.983	133.7	804.4	16.6
	1966/1970	66.0708	0.8387	0.967	186.5	804.4	23.2
LOG. TOT. CASUALTIES	1966/1967	0.0312	0.9825	0.978	0.065	2.694	2.4
	1967/1968	-0.0785	1.0132	0.963	0.086	2.651	3.3
	1968/1969	0.1419	0.9515	0.972	0.073	2.665	2.7
	1969/1970	0.0912	0.9731	0.979	0.062	2.684	2.3
	1966/1970	0.1077	0.9506	0.953	0.093	2.684	3.5
LOG. POPULATION.	1966/1967	0.0256	0.9974	0.997	0.022	4.923	0.4
	1967/1968	0.0578	0.9894	0.997	0.021	4.928	0.4
	1968/1969	0.0330	0.9938	0.999	0.008	4.931	0.1
	1969/1970	0.0316	0.9941	0.999	0.003	4.933	0.05
	1966/1970	0.1571	0.9727	0.989	0.041	4.933	0.8

TABLE 9.5.8. RESULTS FOR THE ANALYSIS WHEN THE LEVEL OF ROAD ACCIDENTS IN ONE YEAR IS REGRESSED AGAINST THE LEVEL OF ROAD ACCIDENTS IN SOME PROCEEDING YEAR.

adjacent years for 1966 to 1970, and also for the period 1966 to 1970 itself. Similarly the results are shown for when total casualties, log. total casualties, total population, and log. total population, are used as variables within this regression model. From this table it can be seen that for all time periods the slope of the regression line as given by the b coefficient is virtually 1.0 and the level of explanation is very high. Now in the exceptional case of b = 1.0, it would be acceptable to state that the absolute growth in road accident casualties is proportionate and independent of the original level of casualties. In this instance therefore one can conclude that whilst the value of the b coefficient indicates a virtual stochastic growth process, the fact that most of the b values are less than one would seem to indicate that smaller urban areas have growth rates which are proportionately greater than those for larger urban areas. In other words, there would seem to be a weak negative relationship between the growth rate and the previous level of road accident casualties. It is necessary however, to qualify this above conclusion by pointing out that the actual relationship is also affected by the variance about the regression line indicated by the error term (e_1) within the regression equation. An indication of this variance is given by the Standard Error of Estimate which should be at a minimum level if the weak negative relationship noted above is not to be invalidated. Thus looking at the tabulated results once again it becomes obvious that the smallest variances are associated with the logarithmically transformed variables. This is of no surprise since the untransformed variables both exhibit a strongly skewed distribution which affect the results and relationships accordingly. Since log. total casualties produces the minimum degree of error, with no real reduction in the overall level of explanation, it is this variable which should be analysed in some further detail. The two features of interest are the values of the intercept and the b coefficient.

The intercept can be regarded as indicating the overall change produced by the growth of road accident casualties within the total population of urban areas, and which would measure the vertical displacement of the curve, whose slope is defined by the b coefficient. Observing the values of the intercept given in Table (9.5.9) it can be seen that all the values are positive except for the period 1967 - 1968 which has a value of -0.079. In fact the whole sequence of the intercept values seems to indicate a quite convincing pattern. In the period 1966 - 1967 there would seem to have been a steady increase in road casualties which preceded the decline noted for the period 1967 - 1968. However, following this decline there then follows a two year period where casualties are increasing but at a decreasing rate. In other words, it would seem that the decline during the 1967 - 1968 period was being accounted for, whilst also accommodating the overall increase for these years noted by the 1966 - 1970 figure. Although any hypothesis put forward for this observed pattern must by nature be very tenuous, it is suggested here that perhaps there is more than a chance coincidence between the decline in casualties and the introduction of the 1967 Road Safety Act. Although this line of thought cannot be followed through in this research, it is also suggested that before any definite conclusions can be made in this direction other variables, such as traffic volumes must be taken into consideration. However, since the results of the original regression analysis showed the impact of police action to be beneficial it would tentatively seem to be a reasonable postulate, and if so, a postulate which would benefit by temporal analysis to consider the lasting effect of such innovations.

Turning to the b coefficient, there would seem to be no temporal pattern as above, although the b coefficient for 1967 - 1968 is the only value which is greater than 1.0. Even so from the results available, it would seem possible to arrive at two conclusions:-

(i) Since the error term is very small it would seem reasonable to accept the negative relationship between the growth in log. casualties and the original level of log. casualties, as being truely representative of the real world situation.

(ii) There would seem to be a gradual increase in this negative relationship with time, even though the increase is not consistent.

The results where the original opening variable was the population variable, produce very similar conclusions, which although varying in minor detail in some places, correspond with those described above for the log. total casualties variable, and can be seen in Table (9.5.8).

A second way in which the growth of road casualties[20] can be compared to the original level of road casualties for each urban area, is by the use of the simple regression equation,

$$GA_{(t.t+1)} = a + bTC_t + e^{21}$$

If the results from the previous regressions are correct then one would expect the results from this second equation to contain very low and very insignificant values. As can be seen from Table (9.5.9) this is exactly the case with very low levels of explanation, and b coefficient values which closely approximate zero. Consequently two conclusions would once again seem to be valid in this situation:-

(i) Because of the horizontal nature of the derived regression it can be concluded that there is no direct significant relationship between the original road casualty level, and the growth rate of casualties.

---

20 The growth rate was calculated by application of the formula;

$$GA_{(t.t+1)} = (TC_{t+1} - TC_t) \cdot \left(\frac{TC_{t+1} + TC_t}{2}\right)^{-1}$$

where, $GA(t.t+1)$ = the growth rate in casualties from time t to t + 1.
TC_t = the total number of casualties at time t.

This compound growth rate was used in order to counter the tendency for the results from a simple growth rate computation to be positively skewed, when normality is really desired.

21 A modification to this equation introduced the logarithmically transformed variable log TC_t in order to reduce the asymmetry of this distribution.

DEPENDENT VARIABLE (GROWTH RATES)	INDEPENDENT VARIABLE (TOTAL CASUALTIES)	INTERCEPT X_o	B. COEFFICIENT (b_1)	F. VALUE	R^2	S.E.E.
GR6766	TC66	-0.031	-0.000	-0.382	0.004	0.148
GR6766	LNTC66	0.071	-0.040	1.399	0.014	0.147
GR6867	TC67	-0.096	0.000	0.002	0.000	0.181
GR6867	LNTC67	-0.156	0.022	0.286	0.003	0.180
GR6968	TC68	0.039	-0.000	0.513	0.005	0.165
GR6968	LNTC68	0.303	-0.103	8.332	0.078	0.159
GR7069	TC69	0.061	-0.000	2.413	0.024	0.142
GR7069	LNTC69	0.207	-0.061	3.433	0.034	0.142
GR7066	TC66	-0.035	-0.000	2.567	0.026	0.212
GR7066	LNTC66	0.241	-0.111	5.336	0.052	0.209

TABLE 9.5.9. RESULTS FOR THE REGRESSION MODEL WHEN THE GROWTH IN ROAD CASUALTIES IS REGRESSED AGAINST THE TOTAL VOLUME OF ROAD CASUALTIES AT THAT TIME.

(ii) Although the magnitude of the b coefficents are very small, since all the values are negative (except for the regression of $GA_{1967-1968}$ on total casualties 1967), it would seem to be reasonable to assume that there is a weak inverse relationship between the level of road casualties and the following growth rates.

The lack of any definite relationships in this analysis therefore, suggests that further insight should be gained as to the actual growth rates themselves. This can be done in two ways,

(i) By studying the growth rates for different years in terms of their basic statistics, and their graphic representations.

(ii) Through disaggregation in order to see if there is any significant difference in the growth rates for different sized urban areas.

The basic statistics for the computed growth rates are shown in Table (9.5.10) from which it can be seen that there is some variation between the values for different years. Thus for the first two sets of growth rates the mean value is negative with $GA_{1967-1968}$ having both the higher mean value and also lower coefficient of variation. The last two sets of growth rates meanwhile both have positive mean values, once again with the values increasing with time. The overall growth rate ($GA_{1966-1970}$) however, has a negative value, but also a high degree of variation

YEARS OF GROWTH RATE	MEAN	STANDARD DEVIATION	COEFFICIENT OF VARIATION
1966–67	-0.038	0.147	393.1
1967–68	-0.096	0.180	188.3
1968–69	+0.030	0.165	558.3
1969–70	+0.044	0.144	323.2
1966–70	-0.061	0.214	351.7

TABLE 9.5.10. CALCULATED BASIC STATISTICS FOR THE ANNUAL GROWTH RATES IN ROAD CASUALTIES FOR THE 100 URBAN AREAS.

(CV = 351.7%). Thus it can be seen that once again there would appear to be a significant break and change in the direction of growth during the period 1967 - 1968, during which time the 1967 Road Safety Act was introduced, and enforced. Before any conclusive statements can be made, however, it was necessary to see if there was any statistically significant difference between these growth rate means. A simple "difference between means" test was employed making use of the t statistic.[22] The results of this analysis are shown in Table (9.5.11) below and indicate that there is a significant difference between the mean values for all years, except for those between 1969 and 1970. Thus it can be stated that the variations in these mean values are not chance variations, and therefore the pattern they imply would seem to be a truely significant pattern. Similarly it would appear that the fluctuations in the growth rates between 1966 and 1969 are beginning to even out, and perhaps return to some steady increase or decrease. It is impossible to state which direction future growth rates will follow because of the very fluctuations, even though the 1966 - 1970 growth rate has a negative value, and such information can only be obtained by further future observations.

---

22 The t test used can be defined by:

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\hat{\sigma}_{\bar{X}_1 - \bar{X}_2}} \quad \text{where:}$$

$$\hat{\sigma}_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{S_1^2}{N_1 - 1} + \frac{S_2^2}{N_2 - 1}}$$

with $(N_1 + N_2 - 2)$ degrees of freedom

YEARS

YEARS	1966 - 1967	1967 - 1968	1968 - 1969	1969 - 1970
1966 - 1967	-	2.48 *	3.02 **	3.96 ***
1967 - 1968	2.48 *	-	5.10 ***	6.05 ***
1968 - 1969	3.02 **	5.10 ***	-	0.68 NS
1969 - 1970	3.96***	6.05 ***	0.68 NS	-

* = 't' is significant at 0.02 level

** = 't' is significant at 0.01 level

*** = 't' is significant at 0.001 level

NS = Not significant.

Table 9.5.11 Results of the 't' test used to test the significance of the difference between the mean growth rate values for the years 1966 - 1970.

As a prelude to studying the growth rates in a disaggregated manner the 1966 - 1970 growth rate was plotted against the total population level for each urban area. This is shown in Fig (9.5.12) from which it can be seen that whilst there seems to be a fair variation about the mean value of -0.06, it also seems as though the variance is reduced at higher population levels. Accordingly it was decided that the growth rates should be disaggregated into

Fig. 9.5.12. PLOT OF ROAD CASUALTY GROWTH RATES (1966 - 1970) AGAINST 1966 POPULATION

four groups based on the poplulation levels of 50,000, 100,000, 250,000, and 250,000+. Similarly since it was hoped to be able to compare these groups, the growth rates were standardised such that;

$$G_{Z_i} = \frac{G_{in} - \bar{G}_n}{\sigma_n}$$

where,

G_{Z_i} = the standardised growth rate of the ith urban area within group n.

$\bar{G}_n$ = the mean value of group n

σ_n = the standard deviation of group n.

As normal these standardised growth rates will have a mean value of zero and a standard deviation of 1. (Table 9.5.13) The importance of these standardised growth rates lies in the fact that they place areas within a group, only respective to the other areas in that same group. Thus two areas from different groups may have totally different absolute growth rates, but may be relatively placed within their respective groups. If this is the case, then both areas will have similar standardised scores, due to their relative locations within their groups. Obviously therefore the greatest value of these scores must lie in a study of their spatial variations.

However, before going on to make use of these standardised growth rates, the original disaggregated absolute growth rates have to be analysed. These original values and basic statistics are shown for each of the four groups in Table (9.5.13) and a plot of these

GROUP 1

0 - 50,000 POPULATION

		ORIGINAL GROWTH RATES	STANDARDISED GROWTH RATES
1.	Aylesbury	-8.74	-0.207
2.	Bury	-1.76	0.065
3.	Bridgewater	-9.00	-0.217
4.	Buxton	3.70	0.278
5.	Canterbury	-10.12	-0.261
6.	Cheshunt	-3.26	0.007
7.	Corby	-22.91	-0.759
8.	Cumbernauld	73.77	3.011*
9.	Dartford	-10.05	-0.258
10.	Dover	27.03	1.188*
11.	Durham	-20.64	-0.671
12.	Farnworth	-19.46	0.625
13.	Glenrothes	36.56	1.560*
14.	Hereford	-12.41	-0.350
15.	Horsham	-4.15	-0.028
16.	Inverness	52.53	2.183*
17.	K. Lynn	6.12	0.373
18.	Lancaster	-42.86	- 1.534*
19.	Newport	3.85	0.284
20.	Pontefract	-16.77	- 0.520
21.	Ryde	-22.41	- 0.740
22.	Stirling	14.33	0.693
23.	Taunton	0.00	0.134
24.	Thornton Cleveleys	-38.78	- 1.378*
25.	Welwyn	-14.32	- 0.424
26.	W. Bridgeford	-31.73	- 1.103*
27.	Whitehaven	-22.71	-0.752
28.	Winchester	-22.83	- 0.756
29.	Wrexham	-12.21	- 0.342
30.	Yeovil	26.15	1.154*

$\bar{X}$ = 3.436
6 = 25.642
N = 30

* GREATER THAN = 1 STANDARD DEVIATION

TABLE 9.5.13. THE GROUPED URBAN AREAS WITH THEIR ORIGINAL AND STANDARDISED GROWTH RATES.

GROUP 2

50 - 100,000 POPULATION

	ORIGINAL GROWTH RATES	STANDARDISED GROWTH RATES
1. Barnsley	-44.08	-2.033*
2. Bath	-15.35	-0.421
3. Brentwood	-18.00	-0.570
4. Burnley	-25.41	-0.985
5. Carlisle	2.79	0.284
6. Chelmsford	-21.32	-0.756
7. Cheltenham	-1.94	0.332
8. Chester	-24.82	-0.952
9. Chesterfield	-11.31	-0.194
10. Colchester	-14.90	-0.396
11. Crewe	-29.44	-0.819
12. Durham	0.82	0.395
13. E. Kilbride	25.00	1.844*
14. Exeter	-14.34	-0.364
15. Gloucester	9.79	0.990
16. Guildford	30.39	2.146*
17. Harlow	17.99	1.453*
18. Hastings	-6.42	0.080
19. H. Hempstead	-3.54	0.242
20. Grimsby	-2.50	0.301
21. Kirkcaldy	-11.81	-0.222
22. Lincoln	3.80	0.654
23. Rochdale	3.85	0.657
24. Rugby	-19.61	-0.660
25. Stafford	-20.55	-0.713
26. Stevenage	-6.92	0.052
27. Stretford	-19.67	-0.663
28. Sutton Cold.	-0.78	-.397
29. Swindon	-4.03	0.215
30. Worthing	-13.30	-0.306

$\bar{X}$ = 7.8537
6 = 17.815
N = 30

* GREATER THAN = 1 STANDARD DEVIATION

TABLE 9.5.13

GROUP 3

100,000 - 250,000 POPULATION

		ORIGINAL GROWTH RATES	STANDARDISED GROWTH RATES
1.	Basildon	-9.37	-0.095
2.	Blackburn	-0.54	0.355
3.	Blackpool	-30.05	-1.146*
4.	Bolton	-0.03	-0.077
5.	Bournemouth	10.83	0.933
6.	Brighton	-13.70	-0.315
7.	Cambridge	-14.54	-0.358
8.	Derby	8,53	0.816
9.	Dundee	16.82	1.238*
10.	Huddersfield	-6.42	0.055
11.	Ipswich	-5.31	0.112
12.	Luton	-4.67	0.144
13.	Norwich	-4.18	0.169
14.	Oldham	-39.92	-1.649*
15.	Oxford	-34.39	-1.367*
16.	Poole	-6.40	0.056
17.	Portsmouth	-21.15	-0.694
18.	Preston	-20.18	-0.644
19.	Reading	-20.09	-0.640
20.	Salford	-4.30	0.163
21.	Southampton	-14.27	-0.344
22.	Southend	-17.87	-0.527
23.	Sunderland	20.31	1.415*
24.	Swansea	46.70	2.757*
25.	York	-14.55	-0.358

$\bar{X}$ = 7.5096
6 = 19.660
N = 25

* GREATER THAN = 1 STANDARD DEVIATION.

TABLE 9. 5. 13

GROUP 4

250,000 + POPULATION

	ORIGINAL GROWTH RATES	STANDARDISED GROWTH RATES
1. Birmingham	-13.45	-0.183
2. Bradford	-10.74	-0.030
3. Bristol	4.52	0.833
4. Cardiff	4.21	0.815
5. Coventry	13.87	1.361*
6. Edinburgh	-4.25	0.337
7. Leeds	12.69	1.295*
8. Leicester	-30.77	-1.162
9. Liverpool	-32.28	-1.247*
10. Manchester	-23.23	-0.736
11. Nottingham	-24.30	-0.796
12. Plymouth	- 9.98	0.013
13. Sheffield	- 7.99	0.126
14. Stoke	-18.03	-0.442
15. Wolves	-13.52	-0.187

$\bar{X}$ = 10.217

6 = 17.692

N = 15

* GREATER THAN = 1 STANDARD DEVIATION

TABLE 9.5.13

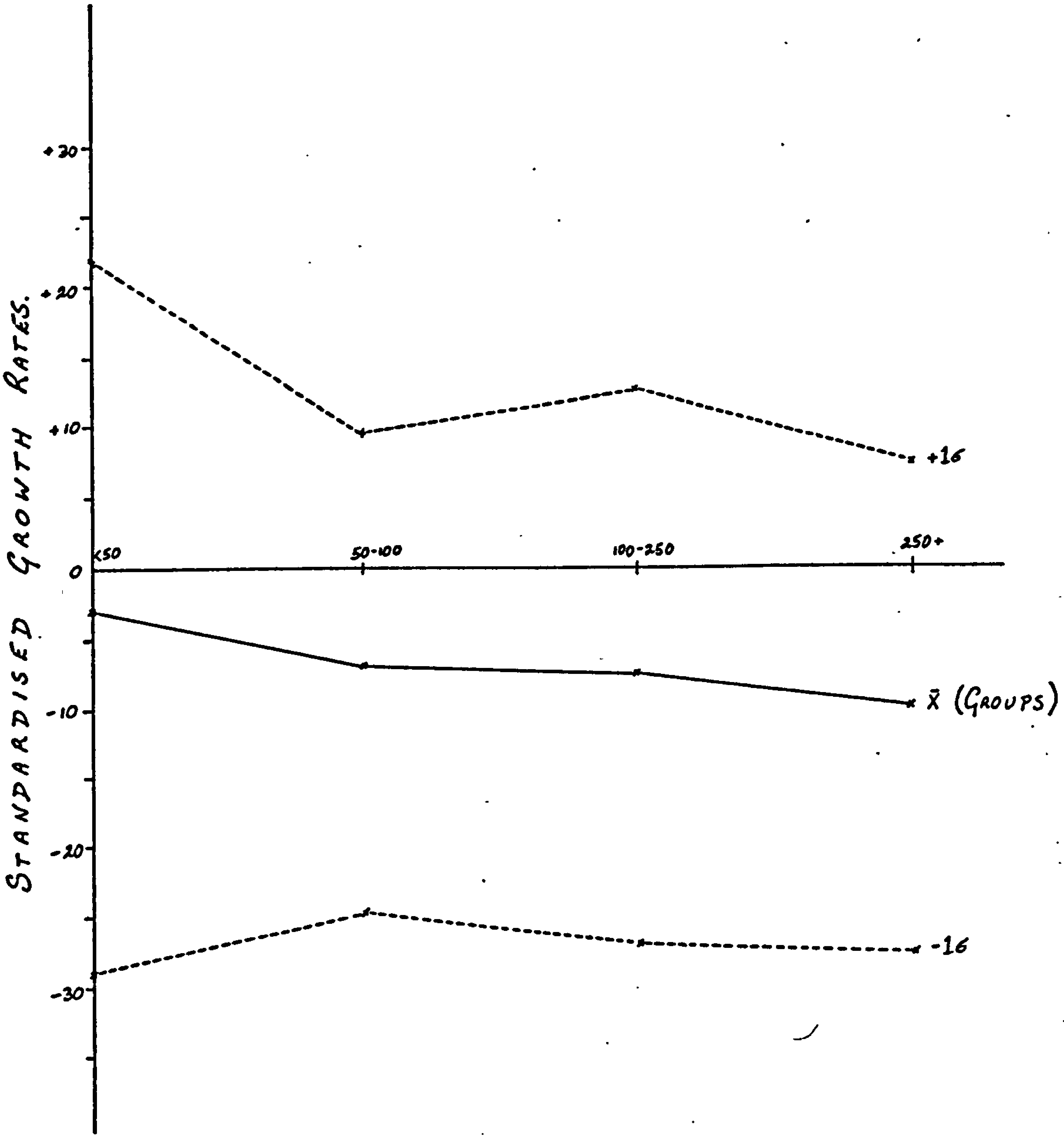


Fig. 9.5.14. PLOT OF STANDARDISED GROUP ROAD CASUALTY GROWTH RATES, MEANS AND STANDARD DEVIATIONS.

statistics for each group is shown in Fig. (9.5.14). It should be remembered that the reason for this disaggregation was to see if there actually was a different growth rate response for smaller urban areas, and also to see if the variance about the mean value similarly varied according to the size of the urban area. This last point can be simply tested by looking at both the graph plot in Fig. (9.5.14) and also the coefficient of variation for each of the four groups. For both of these approaches it can be seen that the variance is by far the largest in the smallest population group. However, the other three groups do not seem to alter too much although the smallest variance is associated with the largest population group. Thus it can be stated that whilst the smaller urban areas have a much greater variation in their growth rates, the degree of variation in the other groups cannot be statistically differentiated even though there is a decreasing trend as population size increases.

In order to test the hypothesis that each of the population groups has a different set of growth rate values, the simple difference of means test was used again in order to see if the difference between the mean growth rates for each of the population groups was at all significant. The results for these tests are presented below in Table (9.515) which shows that none of the group mean values are statistically significantly different from each other at any confidence level.

GROUPS

GROUPS	1	2	3	4
1	-	0.775 *	0.654 *	1.011 *
2	0.775 *	-	0.066 *	0.410 *
3	0.654 *	0.066 *	-	0.436 *
4	1.011 *	0.410 *	0.436 *	-

$$DIF. = N_1 + N_{2/2}$$

* = INSIGNIFICANT 't' value

Table 9.5.15 Results of the student's 't' test used in order to test for any significant difference between the growth rate means for different urban groups. (1966 - 1970)

The two conclusions to be derived from these disaggregated approaches are therefore,

1) Although small urban areas have a much larger variation in growth rates, there is no real discernible significant decrease as population levels increase for the remaining groups.

2) There is no statistical reason to presume that smaller urban areas have different growth rate characteristics to larger areas. The variations noted in their mean values could simply be the result of chance variations.

Although these results show that the disaggregation of the urban areas does not improve the understanding of these growth rates, the disaggregation was maintained in respect of the standardised growth rates. As has already been indicated the greatest use of these scores is in a spatial analysis. However, it should always be remembered that a town with a high score does not mean that its percentage rate of growth was high compared to the national average, but that it was high according to the variation within its respective group. The spatial variation of these standardised growth rates are shown in Fig. (9.5.16) with urban areas plotted according to their deviation from the mean value. Four groups of deviations are marked on either side of the mean value and these can be interpreted as:

1) $\pm 0.5\sigma$ = average growth rates.
2) $0.5 < G < 0.75\sigma$ = moderate growth rates.
3) $0.75 < G < 1.00\sigma$ = high growth rates.
4) $\geqslant 1.00\sigma$ = exceptional growth rates.

From this plotted spatial variation several features become immediately obvious and therefore deserving of further attention. These can be summarised as below:-

1) Although there are only eight sampled points in Scotland, all of these points, except for Edinburgh, have values above the mean and five have values which exceed plus one standard deviation, which perhaps indicates the growing road accident problem in this region. However, it must also be pointed out that these eight points include the three Scottish new towns of Cumbernauld, Glenrothes and E. Kilbride which have been continuing to grow during this period

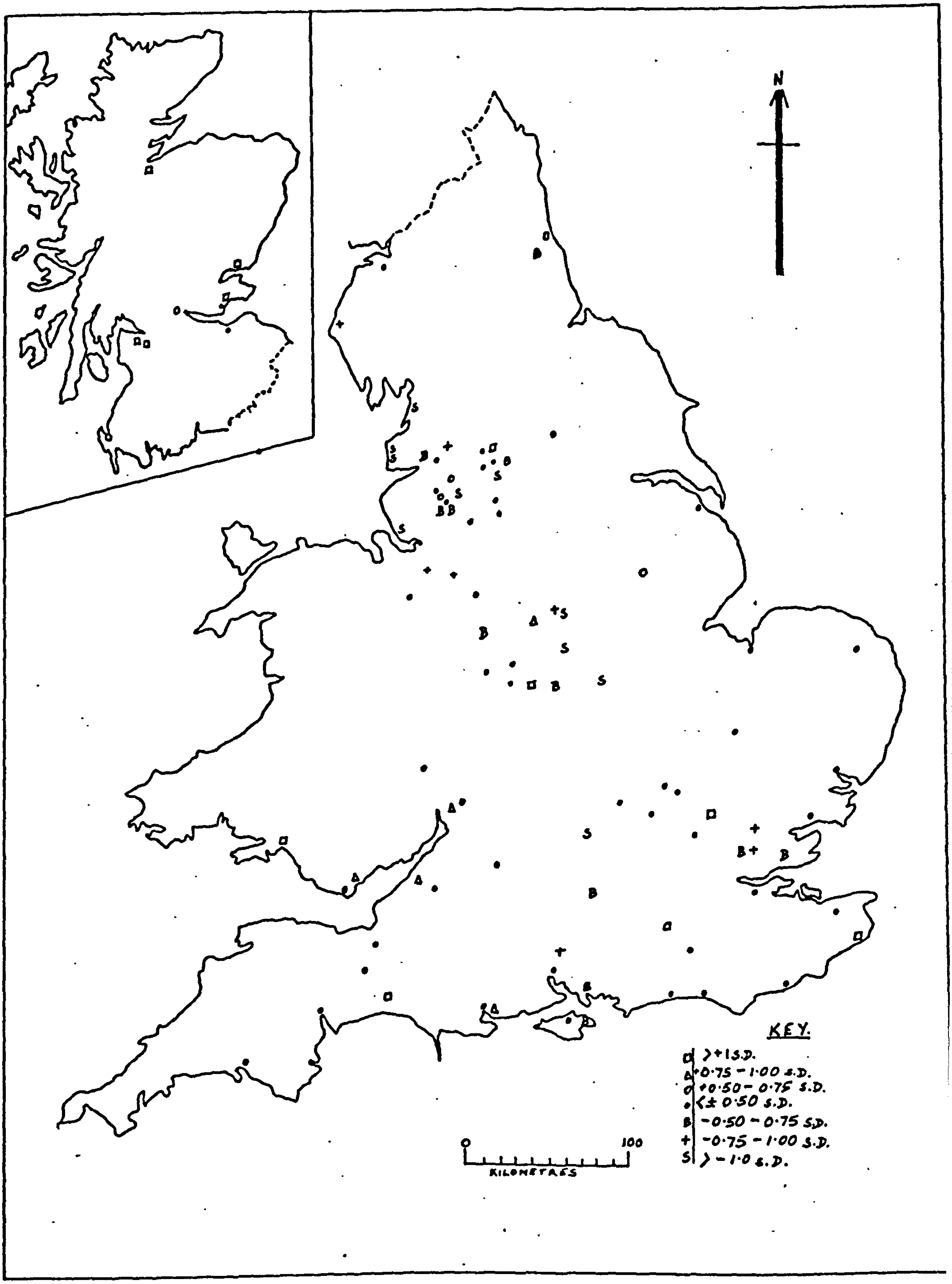


Fig. 9.5.16. MAP OF THE SPATIAL DISTRIBUTION OF GROUP STANDARDISED ROAD CASUALTY GROWTH RATES.

and therefore perhaps should be dealt with as exceptional cases.

2) It would seem that most of the sampled areas in the North-west of England have less than average rates of growth, with only Farnworth and Rochdale having above average values. Since one of the factors which has continually been reappearing is the volume of through traffic, these low values could be due to the improving road and motorway systems in that area.

3) It was noted in the spatial analysis using trend surface techniques, that the area around the gateway to the S.W. peninsula always had underestimated values. A similar situation is also revealed in this analysis, with urban areas in this region and South Wales having either average growth rates or above average growth values. Consequently the arguments put forward in the previous spatial analysis would seem to be verified by this present analysis.

4) Although it was noted earlier that the urban areas around London were usually underestimated in respect of their levels of accident casualties etc., the results of this analysis seem to be less critical of that region. Only two urban areas in the proximity to London have growth rates which are above average, these being Harlow and Guildford. The remainder have either average growth rates or below average rates. With reference to previous results, therefore it must be concluded that although these areas tend to have higher accident casualty levels than the national 'norm,' the rate of growth within these areas is no higher and often lower than in other areas of the country.

5) Finally this analysis pinpoints certain anomaly urban areas with high rates of growth which are out of character with the other urban areas in that region, and which are therefore in need of some special attention. Besides the two areas previously mentioned, Guildford and Harlow, similar urban areas include Bournemouth, Dover, Leeds, Sunderland and Swansea.

This then concludes the presentation of the results from this research. However, before drawing all the conclusions together, since the research was initiated to look at the influence of urban planning proposals, it will be useful to look at the results with particular reference to those areas specifically designed with road safety in mind. That is the new town areas included within the 100 sampled areas.

Chapter 10 THE NEW TOWNS AREAS - A CASE ASSESSMENT

10.1 INTRODUCTION

One of the major assumptions in new town developments is that through urban planning proposals and comprehensive planning in particular, it is possible to create an urban environment which is conducive to road safety both from the engineering point of view as well as the psychological point of view. Implicit in this statement therefore is the argument that the type and design of the urban structure of an area can influence the level of road accident casualties in that area. However, to date, this assumption has never been tested quantitatively, except as in the simplistic terms used by the Cumbernauld Development Corporation. Unless this assumption is true meanwhile, considerable amounts of money could have been expended, for which the return is unacceptable. Furthermore if such an approach is found to be unsatisfactory, then it could be argued that a more satisfactory approach to road safety has been ignored, or remained uninvestigated, purely as a result of a lack of scientific rigour and investigation. Consequently this section of this report attempts to look at the response of the six new town areas included within the 100 sampled areas, by means of the results obtained through the proposed methodological approach advocated in this research. The six areas studied here are; Harlow N.T., Hemel Hempstead N.T., Stevenage N.T., Cumbernauld, East Kilbride and Glenrothes which combine new towns of both the first and second generations. The assessment of these areas will be facilitated by making use of the mean safety indices for these areas, and also by observing the significant road accident casualty growth rates for the time period 1966 - 1970.

10.2 ASSESSMENT OF THE MEAN SAFETY INDICES

The simplest way of looking at these mean safety indices is by observing the response of these new towns in relation to some of the more important dependent variables used in the regression analyses. These MSI values are shown in Table (10.2.1) and will be continuously referred to in this section of the report.

(i) TOTAL CASUALTIES 1970

It can be seen from the results in the table that whilst five of the new town areas have MSI values of less than 100.0 in most the actual deviation from this expected value is only marginal and in the case of Cumbernauld N.T., the MSI value is well over 100.0 (MSI = 139.33). It is also interesting to note that the three older new towns[1] are all within eight MSI units of the expected value, whilst the remaining newer towns[2] have much higher deviation MSI values from the expected. When these MSI values are compared with those derived for other areas it becomes obvious that the new towns response is not as good as would be expected. In several instances towns of a similar size, which have not been planned, exhibit a better response. For example, Hastings, Horsham and Lancaster all have MSI values around the 50.0 mark.[3]

---

1 Harlow, Hemel Hempstead and Stevenage.

2 Excluding Cumbernauld.

3 For further examples of non-planned urban areas see Tables (9.4.19 to 9.4.23).

NEW TOWN	MEAN SAFETY INDICES								
AREA	T C	T A	T F S	TPED	TCPED	TAPP	TCPP	TCVM	TAVM
Cumbernauld	139.33	127.62	34.99	67.59	300.00	103.11	108.54	123.92	144.66
East Kilbride	72.77	69.12	66.33	75.78	110.40	73.82	98.14	67.81	66.27
Glenrothes	60.43	63.20	184.63	58.39	66.56	88.65	106.24	160.78	144.40
Harlow	92.25	85.32	127.13	93.10	115.29	91.38	105.63	86.82	82.56
Hemel Hempstead	92.18	88.98	84.13	88.75	74.46	81.42	78.19	66.03	65.97
Stevenage	100.06	105.47	86.49	99.93	85.42	101.02	84.11	88.15	76.50

TABLE 10.2.1. THE MEAN SAFETY INDICES FOR 9 DEPENDENT VARIABLES RELATING TO THE NEW TOWN AREAS (1970)

Another dependent variable which adds further information to TC7Ø is the total casualties variable for 1966. The important fact to be gained from this variable is that whilst the MSI values for the new town areas remain in the same relative position with each other, in terms of absolute value, they are substantially less.[4] Thus the two MSI values for Harlow in 1966 and 1970 are 70.08 and 92.25 respectively, and for Glenrothes 40.02 and 60.43. This fact should be compared with the fact that the three comparable towns mentioned previously either remain virtually constant or decrease from 1966 - 1970[5], a general decrease which has already been discussed in terms of the growth rates 1966 - 1970. Consequently the tempting conclusion to reach in accord with this data is that whilst these new towns may exhibit favourable accident response rates in their early years of development, they seem to tend towards the national "norm" irrespective of the average national growth rates during that same period of time. If this is true then it would seem to suggest that the new towns improved environment is only effective within the early stages of development.

Before leaving this variable some comment must be made about the apparent uncomplimentary results obtained for Cumbernauld. It was pointed out in the discussion on the areal units used in this research that where the new towns were concerned the unit used was the administrative unit. In most of these cases the administrative unit and the area of the New Town Corporation are almost identical.

---

4 Once again however, Cumbernauld is the exception to the rule.

5 The two values for Lancaster are 74.36(1966) and 51.37(1970) and for Hastings, 76.00(1966) and 65.98(1970).

However, in the case of Cumbernauld, this is not so since the administrative unit covers a much greater area than the new town. In order to see how large an influence this had on the ultimate MSI value for Cumbernauld, data was fed into the derived regression model for TC7Ø, which was based on information provided by Cumbernauld Development Corporation and therefore related solely to the new town area. The resultant MSI value for this new town area was 85.80. Consequently although the results for Cumbernauld will continue to be presented, it must be remembered that the MSI values relate to the administrative unit of Cumbernauld D.C. and not solely the new town area. For this reason comments on the Cumbernauld results will be very limited.

The results for the dependent variable TA7Ø are almost identical to those for TC7Ø, and therefore will not be discussed in any detail. The only significant point to be made is that except for Stevenage and Glenrothes, the MSI values relating to this variable are lower than those for TC7Ø, thus suggesting that these new town areas have more casualties per accident than is the national norm.

(ii) <u>TOTAL FATAL AND SERIOUS CASUALTIES 1970</u>

The implications of this variable are much more difficult to discern than the previous variables because the response varies from one to another new town. Glenrothes and Harlow exhibit high MSI values indicating that the accidents in these areas are more serious than would normally be expected. However, since both of these areas exhibit below average MSI values for TC7Ø it would seem that whilst the overall level of road casualties is reduced the remaining accidents (and therefore casualties) are much more serious. This is an interesting observation from a cost-benefit point of view since it must be doubtful whether the advantages gained from a reduction in total casualties outweighs the disadvantages derived from the high number of serious casualties due to the high financial (as well as human) costs associated with such accidents.

The remaining new towns meanwhile appear to exhibit low MSI values for this variable, and therefore it can be said that these areas have been more successful in reducing the number of serious accidents rather than total accidents. Without further detailed study the reason for this reduction cannot be assertained but it must be presumed that the cause is two-fold. In the first instance the removal of through traffic and high speed traffic to the peripheral areas must have affected this casualty level. Secondly, it must also be presumed that the level of fatal and serious casualties has been affected by the road system within the urban area itself.

(iii) TOTAL PEDESTRIAN CASUALTIES 1970

If the urban environment of the new towns is to achieve its desired aim of a safe environment for those people living within that urban area then one of the ways in which this safety level can be expressed is through the number of pedestrian casualties within these areas. In respect of this variable the response of the new town areas is very good with all six exhibiting MSI values of less than 100.0. However, what is perhaps of greater interest is the fact that the three newer new towns, have much lower MSI values then the other three older new towns. In fact the response of these older areas is not as good as would have been hoped with values well into the 90.0 level. Two conclusions can be made here, either of which could prove to be valid. In the first instance, it could be concluded that the second generation new towns have developed town plans which offer much higher levels of safety to their populations than do the earlier new towns. Alternatively it could also be argued that the higher levels in the earlier new towns is merely a function of time. That is, like the observation made in respect of TC7Ø, it could be that as these new town developments age, their safety responses tend towards the national norm. The only way to solve this dilemma is through more temporal analysis of this variable.

(iv) <u>TOTAL CHILD PEDESTRIAN CASUALTIES 1970</u>

In the context of safety for the inhabitants of an urban area, this variable is just as important as the previous variable for total pedestrian casualties. However, whilst the response for TPED 7Ø was good, the response to this variable varied from area to area. Three of the new towns had values exceeding 100.0 (Cumbernauld, E. Kilbride and Harlow) whilst the remainder had very good response rates, with Stevenage having the highest MSI value of 85.42. Of the three areas with poor response rates Cumbernauld can be dismissed for the reasons forwarded earlier, whilst there would seem no logical reason for the other two areas. For whilst the MSI value for TPED7Ø for Harlow was near 100.0, it did remain below the expected value. Thus it can be stated that in both these areas the facilities for child safety and perhaps the urban structure as a whole is unsatisfactory from the point of this type of casualty. Obviously therefore there is some need to investigate this problem in these areas and perhaps there is some added need for child road safety education in order to try and remedy the situation.

(v) TOTAL ACCIDENTS PER 10^3 POPULATION 1970

It was argued during the analysis of the major regression models that if the size and interaction variables accounted for such a high proportion of the variation in the absolute dependent variables, then this variable which holds the size component constant is a true indicator of the possible effect of urban structure. In these new town areas one would expect therefore that if these areas influence the level of road accident casualties at all, then there would be some low MSI values associated with this variable. From the table it can be seen however, that whilst the majority of MSI values are below the expected level, the overall deviations are only small, with the lowest value being 73.82 for East Kilbride. Similarly when these results are compared to those for the related variable TCPP7Ø, it can be seen that these results are even worse, although it must be pointed out that the above average values for Glenrothes and Harlow, are consistent with the high casualty/ accident rates noted previously for these areas. However, the conclusion must remain that the response of these new town areas to these two variables is nowhere near as good as would have been expected or desired.

(vi) TOTAL CASUALTIES PER 10^6 VEHICLE MILES 1970

The final variables to be examined in relation to the new town areas are the two exposure variables TCVM7Ø and TAVM7Ø. However, since the results for these two variables are so similar, comment will be restricted to the former. As can be seen from the tabled results the response to these variables is good except in the case of Glenrothes, with values as low as 66.03 for Hemel Hempstead and 67.81 for East Kilbride. Whilst these results could be indicating, therefore, that the poor results for the majority of dependent variables in these new town areas are due to the lack of exposure in the variables themselves, it must once again be noted that other areas continue to produce response rates which are as good, if not better than those obtained for the new towns. Thus Dundee, Chesterfield and Stretford have values of 56.91, 57.41 and 56.07 respectively. Consequently it would seem to be rather naive to try and explain the relatively poor results for the new town areas on this fact alone. Similarly since these poor results are fairly constant for whichever variable is chosen, it would seem to be reasonable to assume that these results are a true reflection of the actual performance of these areas.

10.3 ASSESSMENT OF THE GROWTH RATES FOR THE NEW TOWN AREAS 1966 - 1970

Although it is difficult to arrive at any specific conclusions concerning the functioning of these new towns through these growth rates, it has been mentioned in the previous section that there would seem to be a movement towards the national average in these areas according to time from inception, and with little regard to the average national growth rates. For this reason it was deemed useful to look at these growth rates and compare them with other areas. However, it must be remembered that in the case of the more recent new town areas this can only be a qualitative assessment because of the fact that they are still growing and are therefore not a completed unit.

The absolute growth rates and the standardised growth rates for these areas are shown in Table (10.3.1). From this it can be seen that four of the six areas have high positive absolute growth rates, with the remaining two having only small negative absolute growth rates, which incidentally are both less than their respective group mean value growth rate. Consequently when one observes the standardised growth rates it is also seen that all the areas have a positive deviation value, four of which are greater than plus one standard deviation.

The implications of these results are quite important in terms of assessing the success of the new town areas since several of the hypotheses put forward in the previous section appear to be totally vindicated. In the first instance it is obvious that during a period when the overall national situation had average negative growth rates, these new town areas continued, on the whole, to have positive growth rates. If the new town areas were to be as successful as was hoped, it is doubtful if this would be the case.

Secondly, and perhaps more important these above average growth rates would seem to verify the postulate put forward that these new

NEW TOWN AREA	GROWTH RATE 1966–1970 (10^2)	MEAN VALUE FOR ASSOCIATED GROUP (SIZE)	STANDARDISED GROWTH RATE
Cumbernauld	+73.77	– 3.436	3.011
East Kilbride	+25.00	– 7.854	1.844
Glenrothes	+36.56	– 3.436	1.560
Harlow	+17.99	– 7.854	1.453
Hemel Hemp-stead	-3.54	-7.854	0.242
Stevenage	-6.92	-7.854	0.052

TABLE 10.3.1 GROWTH RATES FOR THE NEW TOWN AREAS (1966–1970)

town areas, after an early successful period as regards road safety, appear to be gradually moving towards the national norm and losing the advantages that they originally enjoyed. Even taking into account the fact that these areas are still growing in some instances, this must still remain true due to the size of the deviations from the mean growth rates, which obviously take this fact into account.

10.4 SUMMARY OF THE NEW TOWN RESPONSE RATES

It was argued in the introduction to this chapter, that if it was possible to influence the level of road accident casualties, this must be indicated in the new town areas. However, the results put forward in this chapter do not seem to support this point of view, and in several instances it has been shown that the response of the new towns is sometimes worse than the national norm. As a whole these results seem to lead to five conclusions as regards the new town areas, and these are summarised below.

1) Allowing for a certain degree of variation, the new towns seem in general to exhibit response rates which are below the expected level. However, the size of this deviation in respect to the national norm, and also other similar sized urban areas, is not as great as would be expected for these areas, considering the reasons for their inception.

2) The new town areas do seem to have created an environment which results in considerable savings in the number of pedestrian casualties, and therefore from the point of view of the inhabitants of these areas they are justified developments.

3) The new town areas appear to have their best response rates during the early years of their development. However, after a certain period of time, it would seem that these advantages are eroded and as a result of high growth rates, tend towards the national "norm."

4) The level of these growth rates is both independent of the national average growth rate as well as the actual time period.

5) The response rates of these new town areas do not support the argument that the level of road accident casualties can be controlled by comprehensive planning and modifications of the urban structure, but instead seem to indicate that the road accident system is a homaeostatic system which maintains a constant operating environment in the face of external fluctuations. That is it is a system which resists any alteration in environmental conditions and exhibits a gradual return to equilibrium or steady-state behaviour after such an alteration. Consequently, a certain amount of doubt must be placed upon the justification of new town development solely from a road safety point of view.

Chapter 11. CONCLUSIONS AND POSSIBILITIES FOR FUTURE RESEARCH

11.1 CONCLUSIONS

It has been continually emphasised within this report that the major need within road accident analysis is the creation and application of a more quantitative, scientific approach in order to assess the achievements and progress of road safety proposals. Accordingly, therefore this research has tried to put forward such a methodological approach which enables the researcher to study variations in the volume of road accident casualties, both in a spatial and temporal dimension.

Whilst the development and application of such an approach has been the major purpose of this research, analysis of the data, and the urban structure sub-system, has also enabled certain conclusions to be drawn pertaining to these sections of the road accident problem. That is, it was possible to study the level of influence these structural variables exerted upon the observed volume of road accidents in any one urban area. In most instances these conclusions have been discussed in the relevant sections of the text of this report, but in order to achieve some overall assessment, it becomes necessary to draw all these conclusions together, and thence comment upon them as a whole.

Perhaps the major conclusion which can be drawn from the results of this research is that in most instances, it would seem that the major determinants of the volume of road accidents observed within any urban area, are those factors which are related to the human and psychological sub-systems. Put the opposite way around, one can therefore state that the impact of the structural sub-system upon the level of road accidents would seem to be marginal in the short-term and almost negligible in the long term. This sub-system would appear however, to be instrumental in determining the way in which road accident casualties are allocated between the various types of accidents, and persons (users) injured.

The implications of this conclusion are quite important as regards the total approach to road safety. In many ways it also seems to vindicate the work and suggestions of Cohen (1968), by implying that the major road accident problems are psychological, rather than engineering, in nature. Similarly it would seem that structural changes are only successful whilst they remain "unfamiliar" with, as the saying goes, familiarity breeding contempt. That is, whilst a new road or junction, or urban structure may in the short term lead to a reduction in road accidents, in the long term with an increased understanding of the new structure, the level of accidents may increase once more, simply as a human response to the improved, safer, environment. Therefore in the phraseology of Cohen, it would seem that in order to understand the road accident problem any further, one should be trying to understand the psychological factors inherent within the "bio-robot" model with perhaps slightly less emphasis on the various engineering factors.

The results of the various regression analyses besides emphasising this human dilemma, also pin-pointed two other major factors, which have a significant impact upon the level of road accident casualties. Perhaps the more important of these two factors is the importance of the police activity variables. The results obtained for these variables positively show the significance of police action in maintaining or reducing road accident levels. Although it is difficult to indicate any single causal chain, it must be postulated that the importance of police action is most useful as a simple deterrent to conscious disobedience of the various road safety (traffic) laws, and as such, should be widely encouraged by all sectors of the community.

The second major conclusion to be drawn from these regression models, which is strongly supported by the trend surface analyses, is the unquestionable importance of the degree of through traffic variable. In fact this latter analysis shows how in most instances, the importance of this variable is under-emphasised by the regression models. This can be shown by the fact that all the trend surfaces derived from the relevant regression residuals follow quite remarkably the major road network systems of this country, whilst also indicating the further justification for better east-west transport communications. Similarly with reference to the time series analysis, it would also seem that the importance of these improvements will be most noticeable in relation to the smaller "en route" towns which have not yet been provided with adequate facilities to deal with the present and future volumes of external traffic. Until such improvements are provided in these areas, these smaller towns would seem to be about to increase their road accident problem at a proportionally faster rate than most of the larger urban areas, where it is possibly reasonable to postulate that some equilibrium, saturation level, is slowly being obtained.

A further conclusion to be drawn from this time series analysis, especially with reference to the new town areas, is that from the data and results obtained in this research, it would seem reasonable to postulate that the road accident system, as regards the structural variables, is a homaeostatic system which is always trying to revert back to its equilibrium or steady state behaviour, following some structural fluctuations. If this is true then it would seem that some major change is necessary, probably in the psychological attitude of individuals, before any permanent change in the road accident pattern can be achieved. Consequently it would seem that where any piece-meal improvements are operative, the benefits which have been achieved are, or will be of very marginal value. In respect of the new towns therefore, one should not look for large improvements in the level of road accidents, and road safety in general. Similarly although road safety is only one aspect of the aims of new town developments, from this point of view one must also place several reservations upon the actual success rates of these urban areas, and perhaps also further financial justification should be encouraged before any future new town areas are advocated.

These then are the major conclusions to be drawn from this research although several other important results are argued within the relevent sections of the main text of this report. However, before terminating this report some comment must briefly be made concerning the possible direction of future research.

11.2 FUTURE RESEARCH DEVELOPMENTS

Because of the very nature of this research, as a basic introductory study, the lines of possible future research must be regarded as being of the utmost importance. As a result two possible general lines of progress are suggested here.

The results of this research have continually indicated the importance of the human, psychological sub-system and the need for a greater understanding of the human being in respect to the motor vehicle. If systems analysis is therefore ultimately to be applied successfully to road accident analysis, it is obviously essential to open this sub-system and study the relationships between the various elements found within that sub-system. That is, one of the greatest needs at the present time is a better understanding of the human component in road accident causation.

The second major trend for future research, and perhaps the more important as regards this present research, is the extension of the overview approach, to individual units. That is, so far it has only been possible to combine, or group urban areas according to their aggregate responses, and their overall variation. What is now needed is the individual study, in some depth, of the various individual urban areas which are shown to be representative of a certain group of urban areas. By this means the subjective approach normally used in selecting areas for detailed study, can now be replaced by a more objective approach based on previous analysis and computation. However, it must once again be emphasised that before such analysis can take place, perhaps it would be useful to apply more sophisticated grouping techniques than those used in this analysis, probably once again making use of multi-dimensional statistical techniques.

Working along the same lines of thought, the results of the trend surface analyses also indicate urban areas (and in some cases regions) which show abnormal responses to specific sections of the road accident problem. Therefore once again it should be possible to take these individual areas and subject them to intense scrutiny as regards these specific problems. This should be done in order to discover the cause of these anomalies, and therefore perhaps understand road accident inter-relationships more fully.

A final possibility for future research is the extension of the study of the various noted relationships over time. This should be done by the extensive monitoring of accidents either at the local or national level, in order to obtain suitable temporal data which can be analysed with greater confidence due to the smaller levels of error. The results of such work, extended over a longer period of time than used in this research, would give valuable information as regards the changing influence of various variables, and therefore more information as to the type of nature of the road accident system.

In general therefore the suggestions made here as regards future research include the further collection of data, and the study of element inter-relationships, in order to understand the road accident problem somewhat better, whilst at the same time enabling the structuring of the road accident system. As has been reiterated on several occasions in this report, until such a situation has been reached, no degree of prediction or evaluation can be accepted with any degree of certainty, and road accident research must remain within its present constraining limitations of piece-meal, unscientific advancement.

REFERENCES

ABERCROMBIE SIR P., (1933) Town and Country Planning.

ACKOFF R.L., S.K. GUPTA and J.S. MINAS (1962) Scientific method: optimising applied research decisions. New York.

ANDERSON R.L. and T.A. BANCROFT (1952) Statistical Theory in Research. New York: McGraw-Hill.

ANDERSON B.J., S. ERLANDER and J. GUSTAVSSON (1965) Appendix to the application of temporary road speed limits in Sweden 1961-62. Ministry of Communications, Stockholm.

ASHBY W.R. (1963) An introduction to cybernetics (Wiley Science Editions) New York.

ASHBY W.R. (1966) Design for a brain - London.

BARMACK J.E. (1962) Methodological problems in the design of motor vehicle accident research. Am. Jour. Pub. Health 52. 1866 - 1871.

BARTLETT M.C. (1950) Tests of Significance in Factor Analysis. British Journ. of Psychology (Statistical Section) 3. 77 - 85.

BERRY B.J.L. (1968) Metropolitan Area re-definition: a re-evaluation of concept and statistical practice, Working pap. No.28, U.S. Bureau of the Census of Washington D.C.

BICKMORE (1963) Atlas of Great Britain.

BLAIR D.J. and T.H. BLISS (1967) The measurement of shape in geography: An appraisal of methods and techniques. Bulletin of quantitative data for geographers No. 11. Dept. of Geography, Nottingham University.

BOHLIN N.I. (1967) A statistical analysis of 28,000 accident cases with emphasis on occupant restraint value. 11th Stapp. Car Conference Society of Automotive Engineers, U.S.A.

BORCHERT J.R. (1961) The twin cities urbanised area: past, present and future. Geographical Review 51. 47 - 70.

BUCHANAN C. (1963) Traffic in towns. H.M.S.O.

BUNGE W. (1966) Theoretical Geography: Lund studies in geography, Series C. No. 1 1962 revised and enlarged 1966.

CARROLL J.B. (1953) An analytical solution for approximating simple structure in Factor Analysis. Psychometrika 18, 23 - 38.

CHILDE V.G. (1936) Man makes himself. London.

CHORLEY R.J. and P. HAGGETT (1965) Trend-surface mapping in geographical research. Trans. Inst. Br. Geogr. 37, 47-67.

CHURCHMAN C.W. (1961) Prediction and optimal decision. New Jersey.

CIRILLO J.A., DIETZ S.K. and BEATTY R.L. (1969) Analysis and Modelling of Relationships between Accidents and the Geometric and Traffic Characteristics of the Inter-state System. Bureau of Public Roads, U.S. Dept. of Transportation, Washington D.C.

CLARK P.J. (1956) Grouping in Spatial Distributions, Science Vol. 123, 373 - 374.

CLIFF A.D. and ORD J.K. (1970) Spatial autocorrelation: A review of existing and new measures with applications. Econ. Geogr. 46, 269 - 292.

COHEN and PRESTON (1968) Accident causation and prevention.

COLTON T. and BUXBAUM R.C. (1968) Motor Vehicle Inspection and Motor Vehicle Accident Mortality. Am. J. Public Health 58. 1090 - 1099.

COVERDALE and COLPITTS (1967) Report on an Evaluation of Motor Vehicle Inspection, New York.

CUMBERNAULD DEVELOPMENT CORPORATION (1967) Report on Road Accidents.

CURRY L. (1966) A Note on Spatial Association. The Professional Geographer. Vol. 18. 97 - 99.

DACEY M.F. (1967) Description of line patterns. Northwestern University, Studies in Geography 13. 277-287.

DAVIES O.L. (1963) (Ed.) The Design and Analysis of Industrial Experiments. London.

DAWSON R.F.F. (1971) Current Costs of Road Accidents in G.B. Road Research Lab. Report 396.

DEPARTMENT OF THE ENVIRONMENT (1972) Road Accidents 1970. H.M.S.O.

DHRYMES P.J. (1970) Econometrics: Statistical Foundations and Applications. Harper & Row.

DICKINSON R.E. (1947) City, region and regionalism: a geographical contribution to human ecology. London.

DRAPER N.R. and SMITH H. (1966) Applied Regression Analysis. New York.

DUNCAN O.D. et al. (1961) Statistical Geography: Problems in Analyzing Areal Data. New York.

DUNN J.B. (1970) Traffic Census Results for 1970. Road Research Lab. Report 428.

EFLANDER S. (1971) A review of some statistical models used in automobile insurance and in road accident studies. Accident Analysis and Prevention 3. 45 - 75.

ERLANDER S., GUSTAVSSON J, and LARUSSON E. (1968) Some investigations on the relationship between road accidents and estimated traffic. Accid. Anal & Prev. 1. 17 - 64.

FAIRBAIRN K.J. and ROBINSON G. (1967) Towns and Trend-Surfaces in Gippsland, Victoria. Australian Geographical Studies 5. 125 - 134.

GENSLER H.G. (1966) Power steering cars in accidents. ACIR Bulletin No. 9. Cornell Aero Lab., Buffalo, N.Y.

GORDON J.E. (1949) The Epidemiology of Accidents. Amer. Jour. of Pub. Health 39. 504 - 515.

GOULD P.R. (1966) On Mental Maps. Discussion Paper No. 9. Michigan Inter-University Community of Mathematical Geographers.

HAGGETT P. (1965) Locational Analysis in Human Geography, New York.

HARMAN H.H. (1960) Modern Factor Analysis, Chicago.

HARRIS P. (1971) Road Accident Tabulation Language. (Rattle) Road Research Lab. Report 337.

HARVEY D. (1969) Explanation in Geography, London.

HENDRICKSON A.E. and WHITE P.O. (1964) Promax: a quick method for rotation to oblique simple structure. Br. Jour. Statist. Psychol. 17. 65 - 70.

HERRMANN P.G. (1968) Forecasts of Vehicle Ownership in Counties and County Boroughs in Great Britain. Road Research Lab. Report 200.

HOME OFFICE (1971) Criminal Stats. England & Wales (Cmnd. 5020)

HOME OFFICE (1971 - 72) Offences Relating to Motor Vehicles. (H.C. 461).

HODGE G. (1965) The Prediction of Trade Centre Viability in the Great Plains. Papers, The Regional Science Association 15. 87 - 115.

HORST P. (1941) Prediction of Personal Adjustment. Social Science Research Council Bulletin 48.

JEFFERS J.N.R. (1967) Two Case Studies in the Application of Principal Component Analysis. Applied Statistics, 16. 225 - 236.

JORGENSEN N.O. (1969) A Model for Forecasting Traffic Accidents. Paper presented to the O.C.D.E. Symposium on the use of Statistical Methods in the Analysis of Road Accidents. (R.R.L.Crowthorne) Report 10.

KAISER H.F. and DICKMAN K.W.(1959) Analytical Determination of Common Factors - unpublished manuscript.

KATZ A. (1971) Towards a Methodology of Traffic Safety Measurement and Program Evaluation. Accid. Anal. & Prev. 3. 15 - 43.

KENDALL M.G. (1957) A Course in Multivariate Analysis.

KING L.J. (1969) Statistical Analysis in Geography.

KLIR J. and VALACH M. (1967) Cybernetic Modelling, London.

KRUMBEIN W.C. (1963) Confidence Intervals on Low Order Polynomial Trend Surfaces. Jour. of Geo-physical Research 68. 5869 - 78.

KRUMBEIN W.C. and GRAYBILL F.A. (1965) An Introduction to Statistical Models in Geology, New York.

LAMBE C.G. (1967) Statistical Methods and Formulae, London.

LAWLEY D.N. and MAXWELL A.E. (1971) Factor Analysis as a Statistical Method (2nd Ed.) London.

LIKERT R. (1960) The Dual Function of Statistics. Jour. Amer. Statist. Assoc. 55. 1 - 7.

LOWRY I.S. (1965) A short course in model design. J. Amer. Inst. Planners 31. No. 2 158 - 166.

MACKAY G.M. et al. (1969) Cause and effects of Traffic Accidents. Dept. of Transp. and Environmental Planning. University of Birmingham.

MASSEY W.F. (1965) Principal Components Regression in Exploratory Statistical Research. J. Amer. Statist. Assoc. 60. 234 - 256.

McCARTY H.H., HOOK J.C. & KNOS D.S. (1956) The Measurement of Association in Industrial Geography. Dept. of Geography, University of Iowa.

McCUTCHEON R.W. and SHERMAN H.W. (1969) The Influence of Periodic Motor Vehicle Inspection on Mechanical Condition. J. Safety Research 1. 184 - 193

McFARLAND R.A. (1963) A critique of accident research. Annals of the New York Academy of Sciences 107, Art. 2. 686 - 695.

McGLADE F.S. and ABERCROMBIE S.A. (1965) Accident classification for research purposes. Traff. Quarterly 19(4)

MILL J.S. (1874) System of Logic.

MINISTRY OF TRANSPORT (1967) Highway Statistics, H.M.S.O.

MINISTRY OF TRANSPORT (1967) Road Safety - A Fresh Approach. H.M.S.O.

MORGAN J.N. (1967) Who uses seat belts? Behavioural Sciences 12. 463 - 465.

MUNT P.W. and WEBSTER F.V. (1968) A study of road traffic crossing the cordons round the central areas of eight towns with populations ranging from 18,000 to 8 million. Road Research Lab. Report 209.

NADER R. (1965) "Unsafe at any speed."

ORGANISATION FOR ECONOMICS CO-OPERATION AND DEVELOPMENT (1969) Symposium on the use of statistical methods in the analysis of road accidents. Paris.

PERANIO A. (1970) Conceptualisation and use of road safety and traffic engineering formulas. Traff. Quarterly 25.(3) 429 - 446.

PFUNDT K. (1969) Three difficulties in the comparison of accident rates. Accid. Anal. & Prev. 1. 253 - 260.

POOLE M.A. and O'FARRELL P.N. (1970) The assumptions of the linear regression model. Trans. Instit. of Br. Geogr. No. 52.

RAPAPORT A. (1967) Research on the alcoholic driver: A critique in Selzer, Gikes and Huelke. The Prevention of Highway Injury. Univ. of Michigan.

REGISTRAR GENERAL'S OFFICE (1961) Census Returns.

REGISTRAR GENERAL'S OFFICE (1966) Census Returns

REGISTRAR GENERAL'S OFFICE (1971) Census Returns.

RECHT J.C. (1966) Multiple Regression of the Effects of Safety Activities on the Traffic Accident Pattern. National Safety Council, Chicago

REMPLER S.A. and BLOMQUIST N. (1956) Index I. Jämförelse mellan motorfordonsbestand, trafikintensitet och polisrapporterade vägtrafikolyckor; ars - och manadsvis inom hela landet utan uppdelning pa olika olyckstyper. Stockholm

REMPLER S.A. and WESTERGREN K. (1966) Utrecklingen av polisrapporterade Vägtrafikolyckor 1950 - 1964. Statistiska forskningsgruppen. Stockholm.

RIDLEY T.M. and TURNER E.D. (1968) Value of Traffic Management: A reply. J. Transp. Economics and Policy II 367-388.

ROBERTSON J.S., RYAN G.A. and MACLEAN A.J. (1966) Traffic Accidents in Adelaide, South Australia. A.R.R.B. Special Report No. 1.

ROBINSON A.H. (1956) The Necessity of Weighting Values in Correlation of Areal Data. Annals Assoc. of Amer. Geogr. 46. 233 - 236.

ROBINSON A.H., LINDBERG J.B., and BRINKMAN L.W. (1961) A Correlation and Regression Analysis Applied to Rural Farm Population Densities in the Great Plains. Annals Assoc. of Amer. Geogr. 51. 211 - 221.

ROBSON B.T. (1973) Urban Growth. An Approach. London.

ROSEN, R. (1967) Optimality Principles in Biology. London.

SKILLING H. (1964) An operational view. American Scientist, 52. 388A - 396A.

SMAILES A.E. (1947) The analysis and delimitation of urban fields. Geography 32. 151 - 161.

SMEED R.J. (1949) Some statistical aspects of road safety research. J. Royal Statist. Soc. Series A. (general) Part 1.

SMEED R.J. (1972) The usefulness of formulae in traffic engineering and road safety. Accid. Anal. and Prev. 4. 303 - 312.

SMEED R.J. and JEFFCOATE G.O. (1969) International Comparison of Road Accident Statistics. Paper presented to the O.C.D.E. Symposium on the use of statistical methods in the analysis of road accidents (R.R.L. Crowthorne). Report No. 34.

SMEED R.J. and JEFFCOATE G.O. (1970) Effects of changes in motorisation in various countries on the number of road fatalities. Traff. Engng. Control 12 (3) 150 - 151.

STONE H.W.D. (1914) The principals of urban traffic.

STOTT P.F. (1968) Report from London. Traff. Engng. & Control 9 592 - 595.

SUCHMAN E.A. (1961) A conceptional analysis of the accident phenomenon. Social Problems 8, 241 - 253.

TAAFFE E.J., MORRILL R.L. and GOULD P.R. (1963) Transport Expansion in Underdeveloped Countries: A Comparative Analysis. Geographical Review 53, 503 - 529.

TAMBURRI T.N. (1969) Wrong-Way Driving Accidents are Reduced. Highway Res. Rec. No. 292. Highway Res. Board, Washington D.C.

TANNER J.C. (1958) A problem in the combination of accident frequencies. Biometrika 45. 331 - 342.

THOMAS E.N. (1960) Maps of Residuals from Regressions: Their Characteristics and uses in Geographic Research. Monograph No. 2, Dept. of Geography, Univ. of Iowa.

THOMAS E.N., and ANDERSON D.L. (1965) Additional comments on weighting values in correlation analysis of areal data. Annal. Assoc. of Amer. Geographers 55, 492 - 505.

THOMPSON G.H. (1951) The Factorial Analysis of Human Ability, London.

THOMSON, J.M. (1968) The value of traffic management. Jour. Transp. Econ. and Policy II, 3 - 32.

THURSTONE L.L. (1947) Multiple Factor Analysis, Chicago.

TISDALE H. (1942) The process of urbanisation. Social Forces 20, 311 - 316.

TOBLER W.R. (1966) Numerical Map Generalisation and Notes on the Analysis of Geographical Distributions. Discussion Paper No. 8. Michigan Interstate University Community of Mathematical Geographers

TRIPP H.A. (1942) Town Planning and Road Traffic.

WEBER M. (1958) The City. New York.

WHITLEN E.H.T. (1959) Composition trends in a granite: modal variation and ghost stratigraphy in part of the Donegal granite Eire. J. Geophys. Res. 64. 835 - 848.

WOHL M. (1969) Putting the analysis and evaluation of traffic safety measures into perspective. Accid. Anal. and Prev. 1, 77 - 101.

WONG S.T. (1963) A multi-variate statistical model for predicting mean annual flood in New England. Annals. Assoc. of Amer. Geogr. 53. 298 - 311.

APPENDIX 1

THE TOWN CLASSIFICATION INDEX

Appendix 1. THE TOWN CLASSIFICATION INDEX

The derivation of the town classification index has already been explained (page 113), and therefore the only important point to reiterate at this point is that the index is a compound variable consisting of the product of Haggett's shape index and Clark's network structure. Consequently this compound index enables one to observe the type of network system present within our existing urban areas. In order to extend this idea the T.C.I. variable was plotted against Haggett's shape index. (Fig. 12.1.1.).

As can be seen from this diagram the Haggett shape index can be divided into three regions:-

1) 0.20 - 0.40 which can be said to tend towards linearity,
2) 0.40 - 0.83 random.
3) 0.83 - 1.00 which can be said to tend towards circularity.

Similarly since Clark's work shows that the number of nearest neighbour points can indicate the kind of network structure, it also becomes possible to divide each of the afore-mentioned three groups (shape) into five further groups, according to the T.C.I. values. The limiting values for each of the network grids were:-

1) Manhattan $\geqslant$ 0.75
2) Random 0.66
3) Radial $\leqslant$ 0.45

and therefore the calculated T.C.I. limits for the shape group 0.20 - 0.40 can be given as:-

1) Manhattan 0.05 - 0.1050
2) Random 0.0660 - 0.1386
3) Radial 0.1100 - 0.2310

FIG. 12.1.1. PLOT OF T.C.I. AGAINST HAGGETT'S SHAPE INDEX

Correspondingly, when the towns are plotted upon this graph framework one is able to study these areas according to both their shape and kind of network structure (Fig. 12.1.1.).

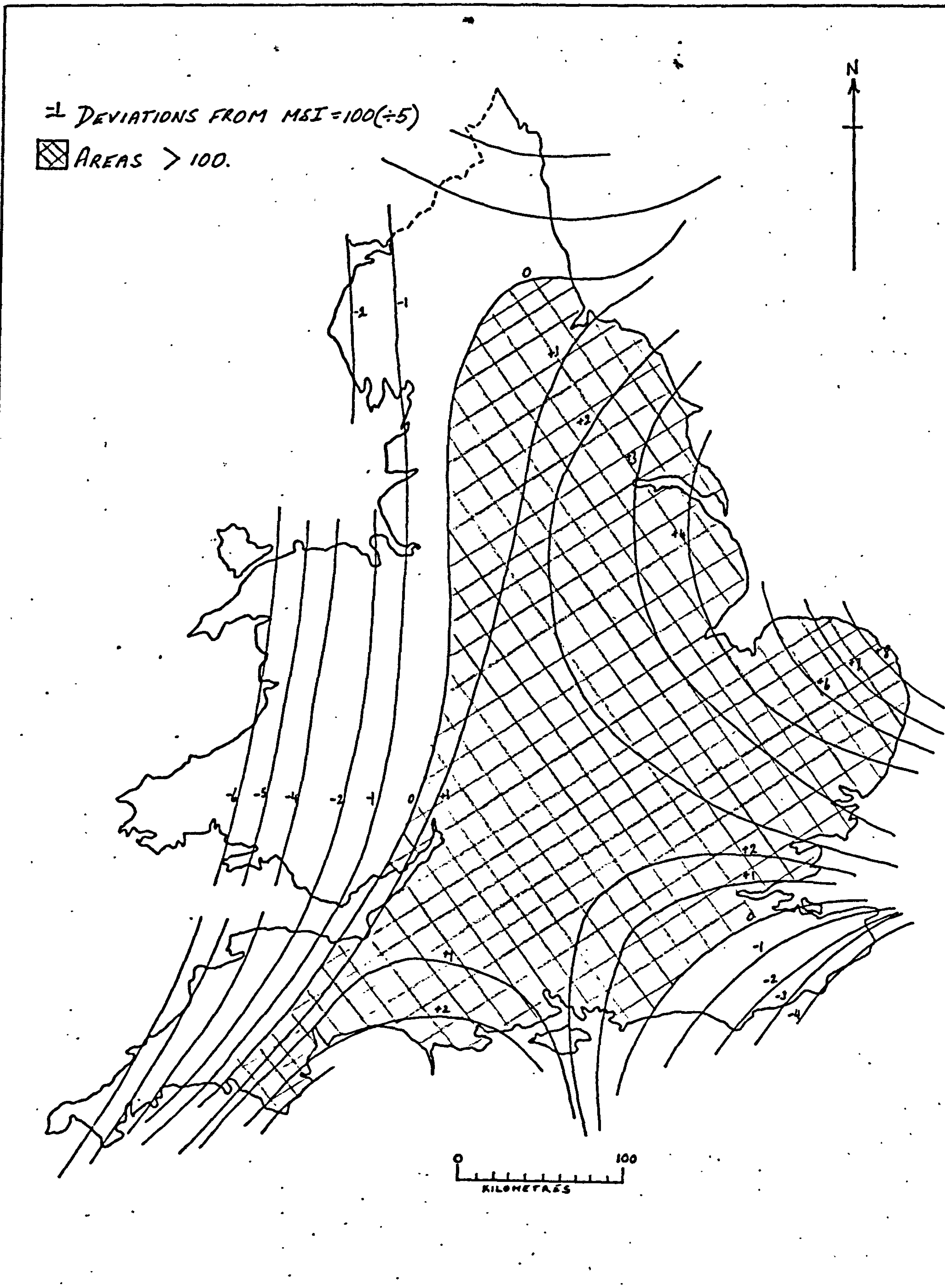
The most obvious feature of this graph plot is the seemingly linear relationship, which is of little surprise, since it simply means that the shape of the urban area becomes more circular there is a positive tendency to develop a radial network system. However what is of more interest is the manner in which the network systems vary within each shape group. Thus within the lower shape group Southend has a distinct Manhattan grid whilst Manchester would seem to have a strong tendency towards a radial network.

Within the middle shape range group, there is once again a wide divergence. Thus Crewe, Canterbury and York all have strong radial network systems whilst Swansea, Birmingham and Cambridge would seem to have Manhattan grid systems.

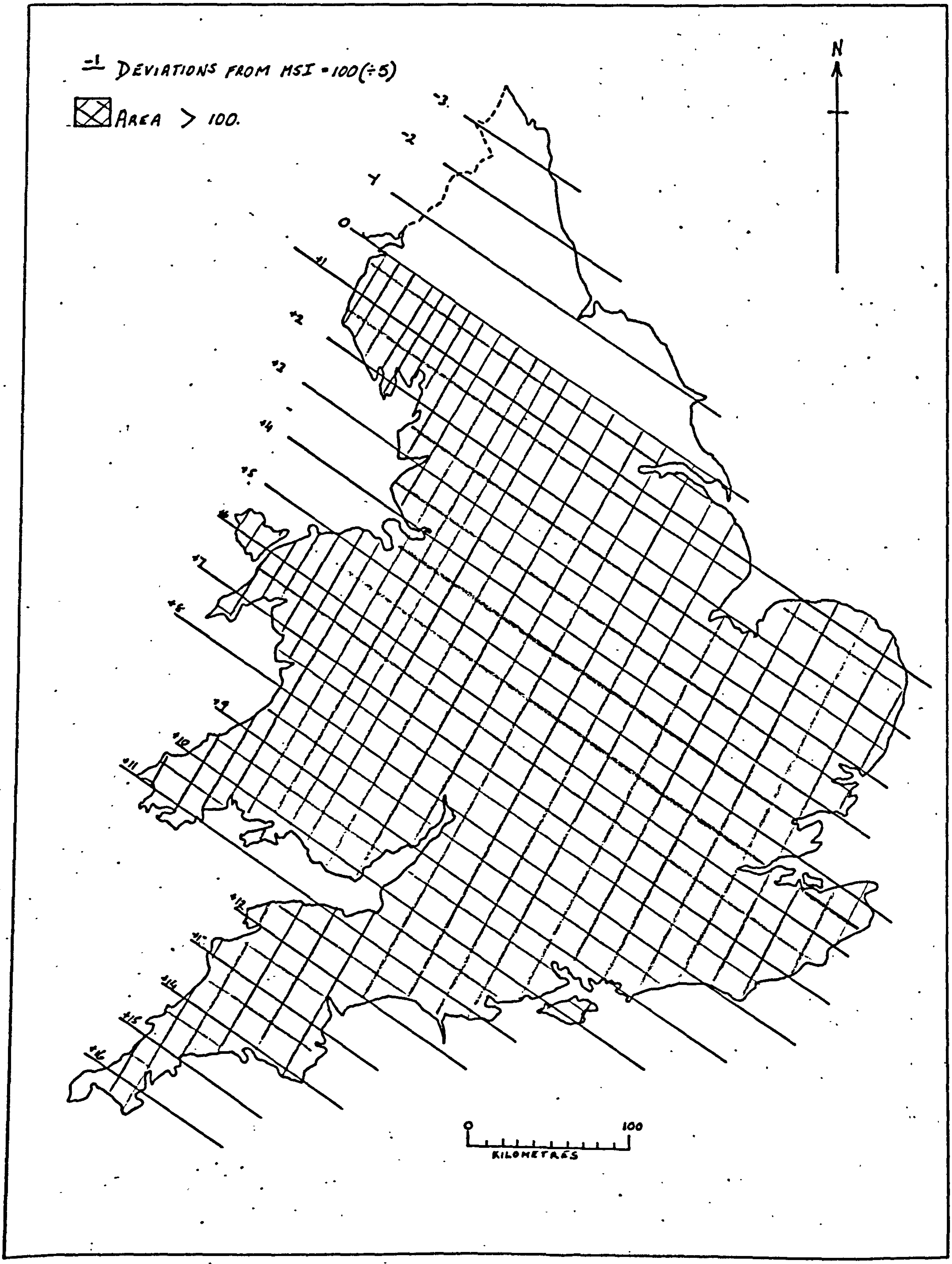
Although this line of research is only marginal to the project reported here, it is suggested that as a descriptive device this index and method of representation could perhaps repay further investigation.

APPENDIX 2

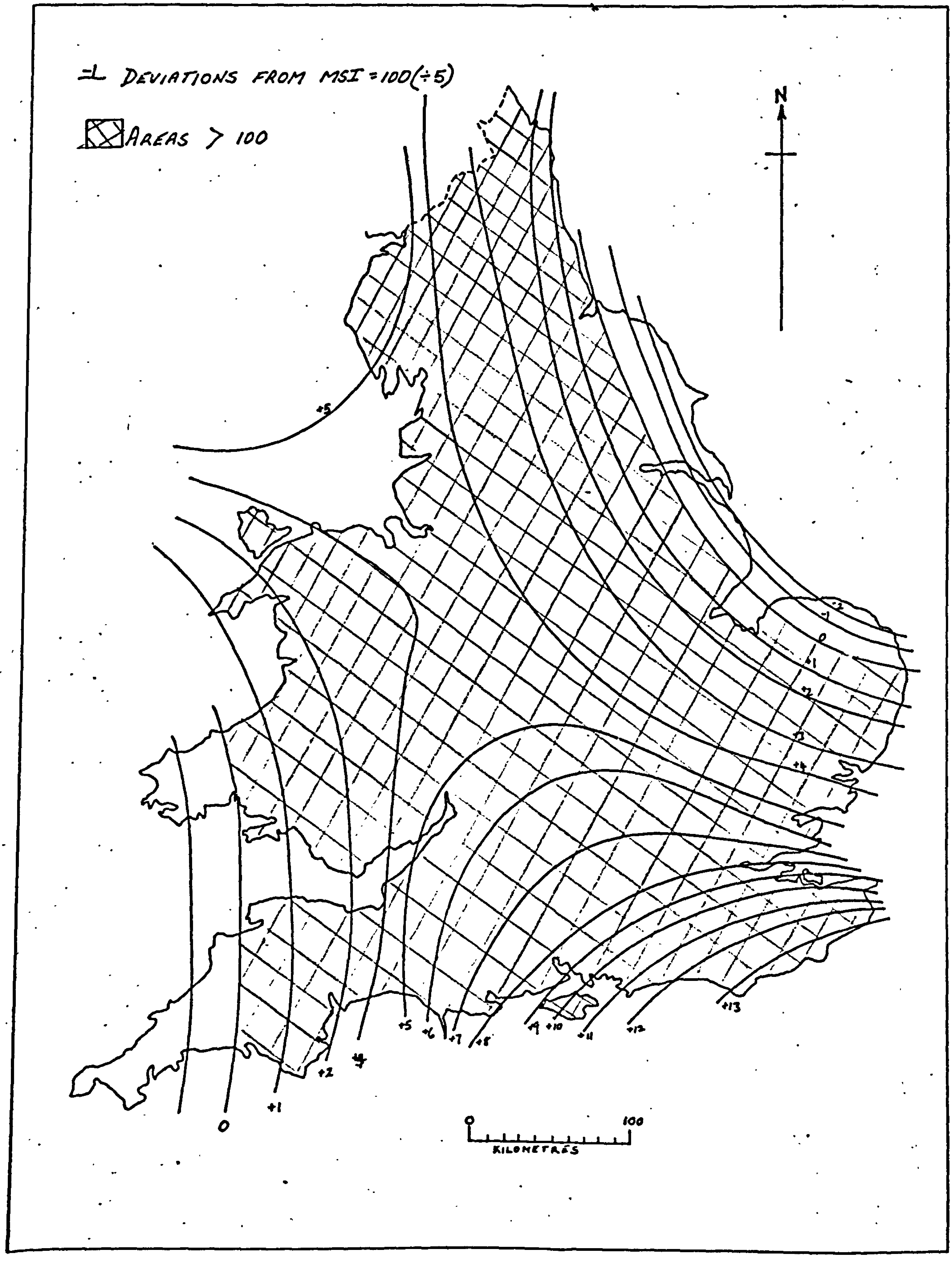
SIGNIFICANT TREND SURFACE MAPS



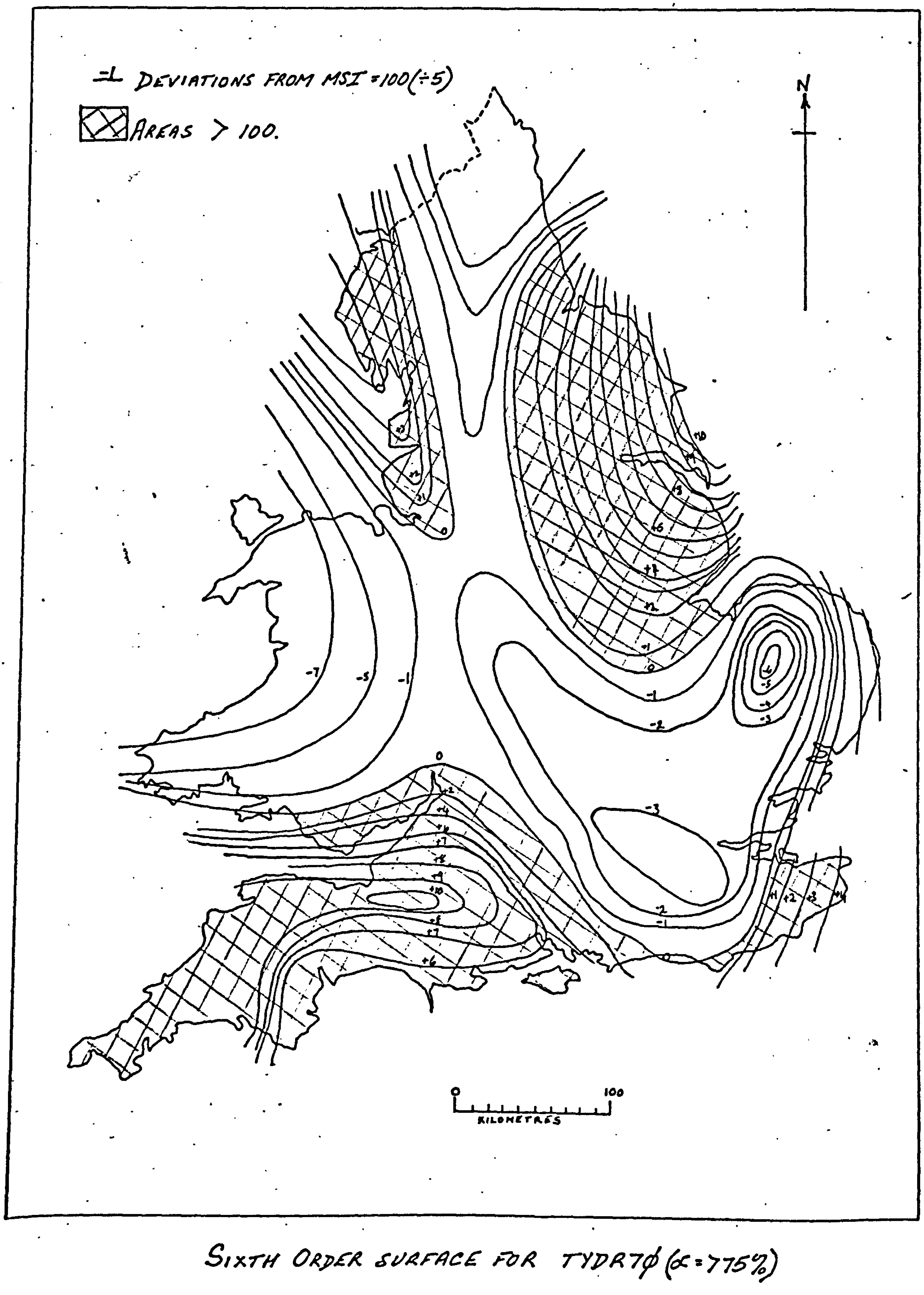
CUBIC SURFACE FOR TA7Ø (α = >75%).



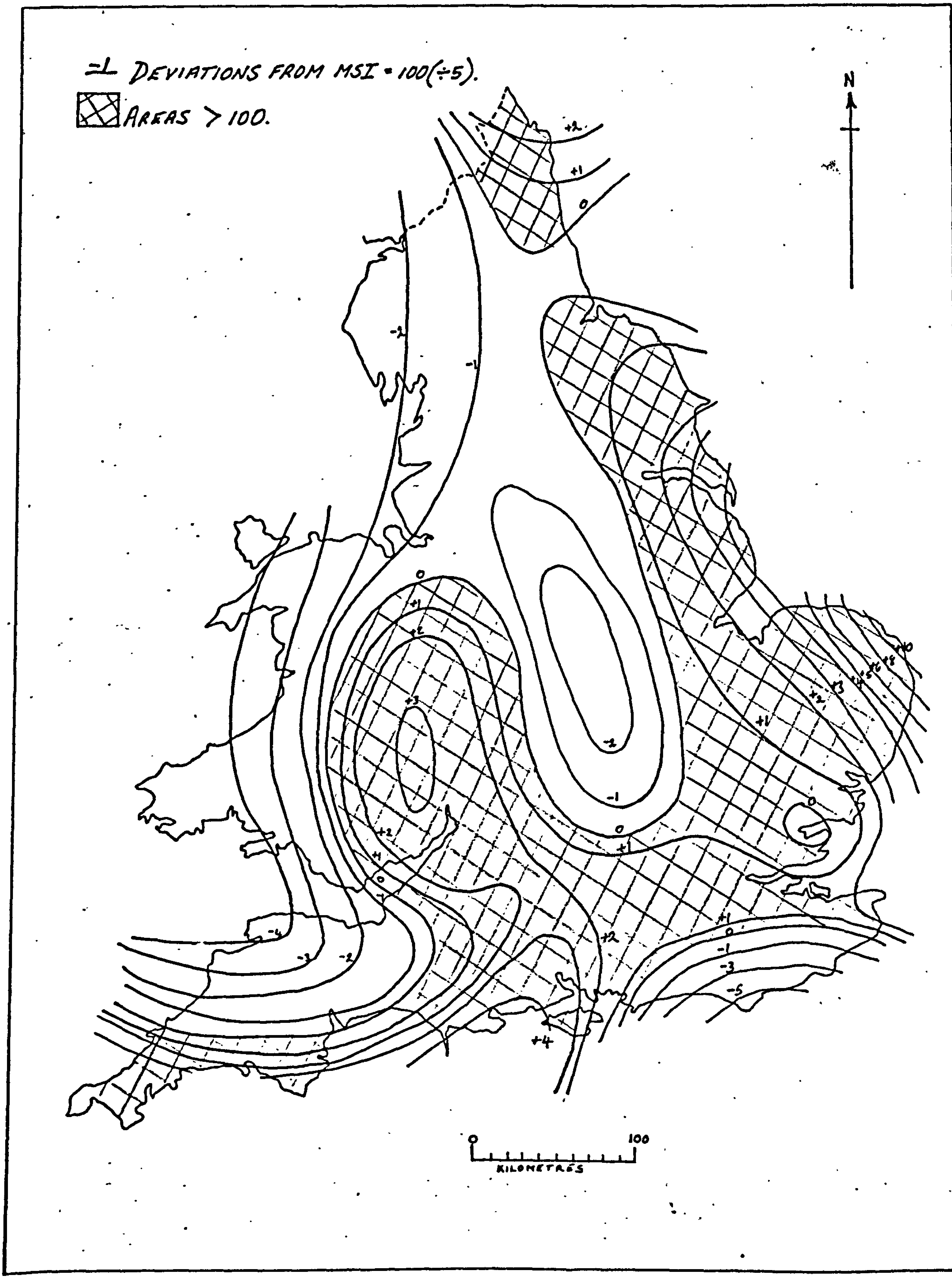
Linear Surface for TFS70 (α = 7.75%).



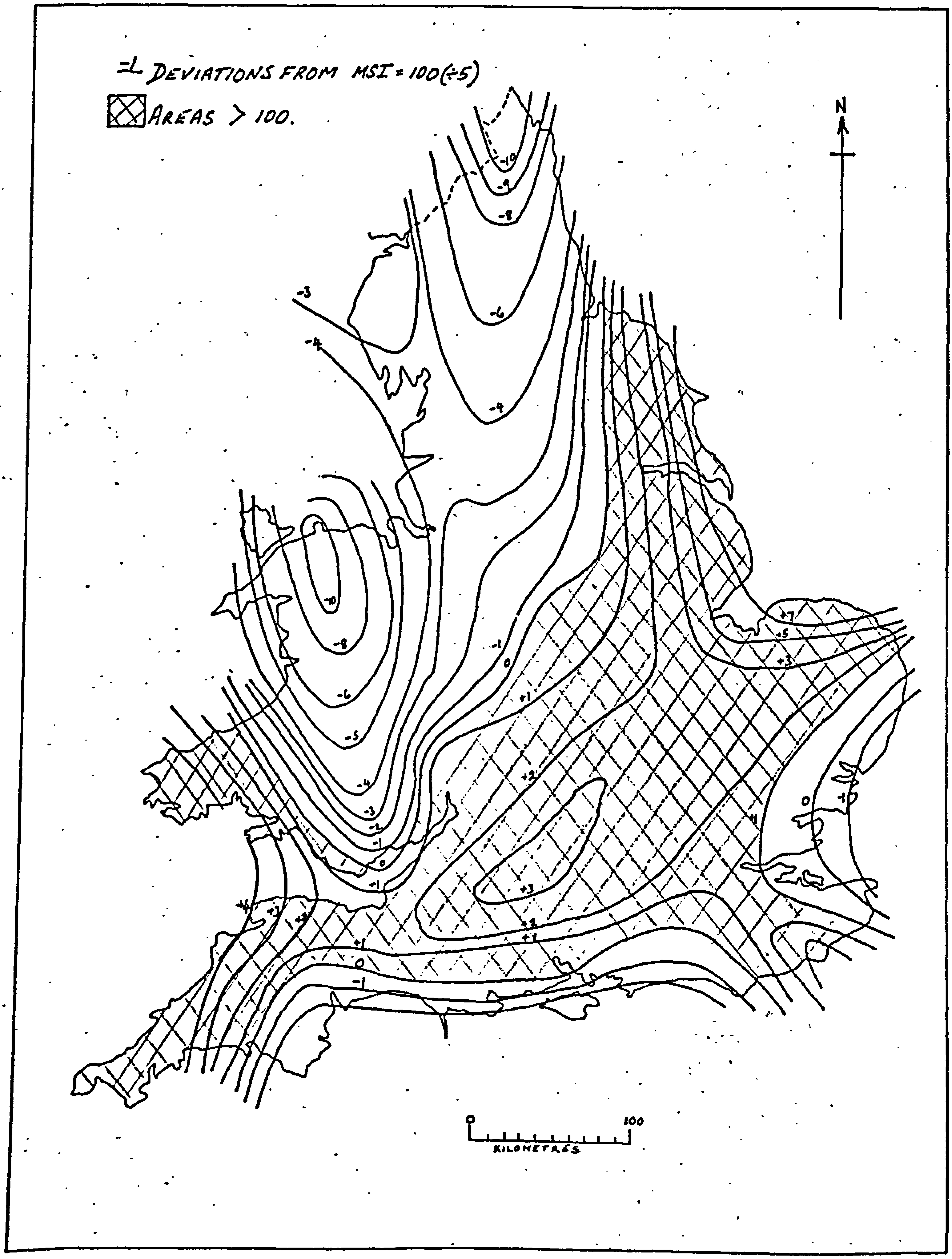
QUADRATIC SURFACE FOR TCPED70 (α = >75%).



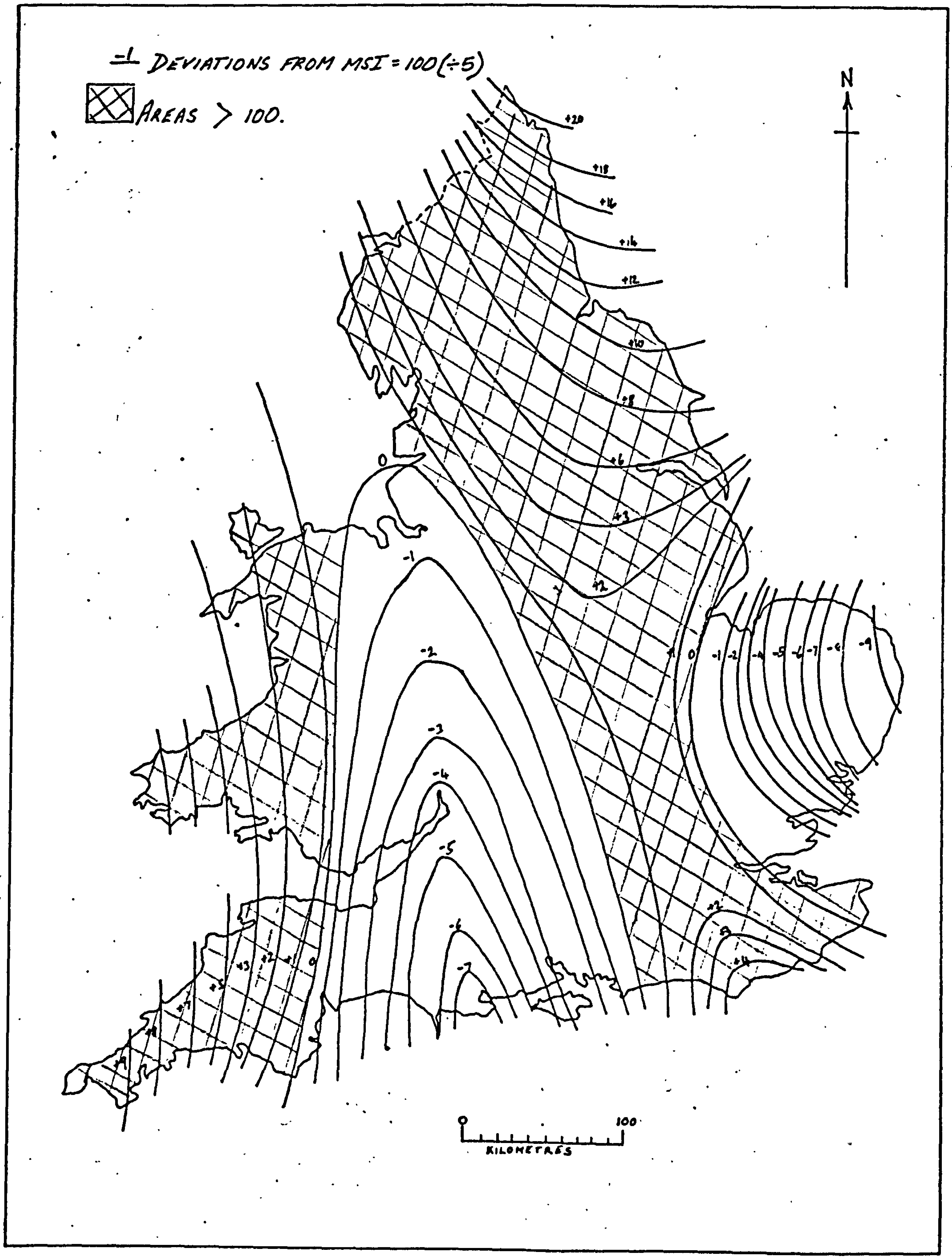
SIXTH ORDER SURFACE FOR TYDR70 (α = 77.5%)



SIXTH ORDER SURFACE FOR TCPP70 (α = > 75%)



FIFTH ORDER SURFACE FOR TAMURD70 (α = >75%).



CUBIC SURFACE FOR TPCWT7Ø (α = 77.5%).

APPENDIX 3

MULTIPLE REGRESSION RESULTS

Y	REGRESSION: TOTAL CASUALTIES				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	779.09		$\bar{Y}$	804.43
	X1	1011.31	16.48	S.D.	1015.21
				S.E.E.	123.32
	X2	-36.20	14.91	R^2	0.986
	X8	-26.67	14.25	F	827.25
	X10	-44.02	12.64	D.F.	8,91
	X15	-45.62	14.53		
	X17	-92.94	21.23		
	X18	-157.86	20.90		
	X19	-38.44	12.89		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	787.37		$\bar{Y}$	787.37
				S.D.	1062.76
	X1	1012.61	20.25	S.E.E.	200.63
	X4	-55.43	20.27	R^2	0.966
	X15	134.12	20.32	F	670.71
	X19	-59.93	20.70	D.F.	4,95

Y	REGRESSION: TOTAL ACCIDENTS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	.REGRESSION	STATISTICS
	X_o	616.77		$\bar{Y}$	636.72
	X1	812.23	13.10	S.D.	808.17
				S.E.E.	98.06
	X2	-31.74	11.86		
	X8	-26.93	11.33	R^2	0.897
	X10	-32.58	10.05	F	829.23
				D.F.	8,91
	X15	34.62	11.55		
	X17	-68.53	16.88		
	X18	-135.06	16.62		
	X19	-31.31	10.25		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	622.46		$\bar{Y}$	622.46
				S.D.	813.12
	X1	818.36	14.83		
				S.E.E.	108.81
	X4	-28.41	11.13		
	X7	20.46	11.30	R^2	0.983
	X15	73.90	11.91	F	776.59
	X17	-52.13	18.46	D.F.	7,92
	X18	-100.84	19.54		
	X19	-40.91	11.54		

Y	REGRESSION: TOTAL FATAL AND SERIOUS CASUALTIES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	216.30		$\bar{Y}$	223.75
	X1	278.74		S.D.	281.57
	X8	-12.82	5.77	S.E.E.	57.57
	X9	14.32	7.12	R^2	0.956
	X16	17.50	5.92	F	568.37
				D.F.	4,95
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	212.09		$\bar{Y}$	212.09
	X1	283.91	8.46	S.D.	279.09
	X7	14.60	6.73	S.E.E.	65.77
	X17	23.54	8.31	R^2	0.947
	X20	-16.51	6.67	F	421.89
				D.F.	4,95

Y	REGRESSION: TOTAL PEDESTRIAN CASUALTIES.				
	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
1970 RESULTS	X_o	284.39		$\bar{Y}$	257.41
	X1	375.84	5.83	S.D.	369.55
	X3	-19.62	4.22	S.E.E.	41.69
	X10	-12.78	4.25	R^2	0.988
	X15	22.70	4.90	F	1098.25
	X16	-14.31	4.80	D.F.	7,92
	X17	-11.97	6.92		
	X18	-55.86	7.40		
	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
1969 RESULTS	X_o	233.61		$\bar{Y}$	233.61
	X1	273.90	11.53	S.D.	309.06
	X3	-22.29	12.08	S.E.E.	109.98
	X4	-19.65	11.20	R^2	0.881
	X5	23.96	11.47	F	114.80
	X16	46.17	11.71	D.F.	6,93
	X19	-25.25	11.46		

Y	REGRESSION: TOTAL ACCIDENTS INVOLVING MOTOR CYCLES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	.REGRESSION	STATISTICS
	X_o	70.01		$\bar{Y}$	72.80
	X1	73.20	3.93	S.D.	84.11
				S.E.E.	27.48
	X3	6.03	3.04	R^2	0.901
	X7	-8.99	3.95	F	119.34
	X15	18.02	3.27	D.F.	7,92
	X16	14.60	3.18		
	X17	-24.54	4.52		
	X18	-12.58	4.93		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	88.59		$\bar{Y}$	88.59
				S.D.	80.01
	X1	37.90	6.13		
				S.E.E.	47.58
	X9	12.25	5.05	R^2	0.671
	X15	10.97	5.11	F	26.85
	X16	28.28	5.30	D.F.	7,92
	X17	-20.74	6.71		
	X19	-12.18	4.94		
	X21	13.35	4.94		

Y	REGRESSION: TOTAL ACCIDENTS INVOLVING PEDAL CYCLISTS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	84.72		$\bar{Y}$	86.73
	X1	35.99	8.29	S.D.	81.03
				S.E.E.	49.40
	X12	25.91	6.58	R^2	0.641
	X17	-15.96	7.55	F	42.58
	X21	19.48	4.98	D.F.	4,95
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	72.75		$\bar{Y}$	72.75
				S.D.	79.33
	X1	61.10	4.16	S.E.E.	31.55
	X6	7.56	3.40	R^2	0.851
	X7	6.97	3.34	F	88.82
	X16	22.29	3.46	D.F.	6,93
	X17	-13.97	4.46		
	X19	-11.37	3.27		

Y	REGRESSION: TOTAL DRIVER AND RIDER CASUALTIES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	320.88		$\bar{Y}$	329.04
	X1	346.39	10.65	S.D.	362.64
				S.E.E.	76.26
	X3	26.81	8.29	R^2	0.959
	X15	30.51	9.24	F	306.67
	X16	40.21	8.99	D.F.	7,92
	X17	-48.80	12.55		
	X18	-36.09	13.73		
	X19	-22.91	8.10		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	308.78		$\bar{Y}$	308.78
				S.D.	312.25
	X1	248.98	13.50	S.E.E.	105.57
	X4	-19.88	10.87	R^2	0.893
	X9	23.53	10.82	F	128.83
	X16	78.68	11.50	D.F.	6,93
	X17	-43.73	14.38		
	X19	-41.75	10.91		

Y	REGRESSION: TOTAL CHILD PEDESTRIAN CASUALTIES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	.REGRESSION	STATISTICS
	X_o	116.91		$\bar{Y}$	120.75
	X1	178.24	4.37	S.D.	117.06
				S.E.E.	33.01
	X10	-6.99	3.33	R^2	0.967
	X13	11.49	3.35	F	459.27
	X15	12.55	3.81	D.F.	6,93
	X17	-12.75	5.39		
	X18	-33.22	5.53		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	109.93		$\bar{Y}$	109.93
				S.D.	151.19
	X1	131.15	6.12	S.E.E.	60.92
	X14	15.97	6.23	R^2	0.850
	X15	14.45	6.23	F	106.63
	X16	25.46	6.11	D.F.	5,94
	X19	-11.95	6.22		

Y	REGRESSION: TOTAL YOUNG DRIVER AND RIDER CASUALTIES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	160.40		$\bar{Y}$	166.04
	X1	175.20	6.66	S.D.	180.16
				S.E.E.	46.56
	X3	16.35	5.15	R^2	0.938
	X7	-14.38	6.69	F	198.61
	X15	12.89	5.54	D.F.	7,92
	X16	26.94	5.39		
	X17	-22.23	7.66		
	X18	-18.57	8.35		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	163.30		$\bar{Y}$	163.30
				S.D.	164.51
	X1	138.71	7.70	S.E.E.	57.61
	X3	11.01	6.24	R^2	0.885
	X9	12.66	5.85	F	119.06
	X16	26.94	5.39	D.F.	6,93
	X17	-22.23	7.66		
	X19	-20.57	5.89		

Y	REGRESSION: TOTAL JUNCTION ACCIDENTS.			

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	.REGRESSION	STATISTICS
X_o	413.68		$\bar{Y}$	431.16
X1	484.77	15.82	S.D.	557.98
X2	-53.62	14.31	S.E.E.	115.62
X8	-35.92	14.72	R^2	0.961
X9	43.02	17.31	F	241.24
X10	-32.95	11.93	D.F.	9,88
X15	147.37	13.73		
X16	77.91	13.49		
X17	-122.40	19.75		
X19	-32.70	12.88		

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	434.63		$\bar{Y}$	437.42
X1	593.78	10.41	S.D.	629.55
X4	-28.34	10.39	S.E.E.	102.83
X15	112.58	10.43	R^2	0.974
X19	-23.27	10.64	F	885.60
			D.F.	4,95

Y — REGRESSION: TOTAL ACCIDENTS AT PEDESTRIAN CROSSINGS.

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
X_o	18.59		$\bar{Y}$	18.93
X1	29.57	2.70	S.D.	33.81
			S.E.E.	12.12
X12	3.56	2.12	R^2	0.878
X16	-4.79	1.56	F	127.05
X17	-5.77	2.26	D.F.	5,88
X18	-9.49	2.06		

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	15.68		$\bar{Y}$	15.36
			S.D.	24.53
X1	21.17	1.09	S.E.E.	10.67
X16	2.72	1.12	R^2	0.817
X19	-3.22	1.09	F	138.28
			D.F.	3,93

Y	REGRESSION: TOTAL ACCIDENTS DUE TO A TURNING MOVEMENT.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	148.53		$\bar{Y}$	156.42
	X1	118.57	5.83	S.D.	145.77
				S.E.E.	45.82
	X8	-9.27	4.69	R^2	0.907
	X16	48.50	4.77	F	163.32
	X17	-13.79	5.01	D.F.	5,84
	X19	-12.54	6.28		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	147.96		$\bar{Y}$	157.67
				S.D.	147.78
	X1	110.87	7.23	S.E.E.	56.16
	X4	-12.61	6.09	R^2	0.865
	X9	16.95	6.04	F	88.87
	X16	43.09	6.27	D.F.	6,83
	X17	-19.59	7.81		
	X19	-11.25	5.82		

Y	REGRESSION: TOTAL ACCIDENTS AT ROUNDABOUTS.			

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
X_o	20.41		$\bar{Y}$	21.06
X1	27.21	1.31	S.D.	26.72
			S.E.E.	10.54
X15	-2.65	1.19	R^2	0.851
X18	-7.96	1.48	F	132.66
X21	3.38	1.06	D.F.	4,93

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	20.16		$\bar{Y}$	20.29
			S.D.	28.89
X1	30.42	1.42	S.E.E.	10.64
X3	4.51	1.45	R^2	0.870
X14	2.34	1.38	F	155.81
X18	-6.24	1.41	D.F.	4,93

Y	REGRESSION: TOTAL ACCIDENTS AT T JUNCTIONS.			

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
X_o	219.88		$\bar{Y}$	227.86
X1	245.34	7.89	S.D.	269.18
			S.E.E.	61.82
X10	-20.96	6.26	R^2	0.950
X15	34.94	6.79	F	291.37
X16	39.16	6.72	D.F.	6,91
X17	-18.78	9.12		
X19	-17.63	6.53		

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	220.90		$\bar{Y}$	220.01
			S.D.	277.78
X1	240.40	9.19		
X4	-23.76	7.24	S.E.E.	70.97
X15	46.78	7.50	R^2	0.939
X16	32.29	7.89	F	232.50
X17	-24.35	10.01	D.F.	6,91
X19	-24.21	7.46		

Y	REGRESSION: TOTAL ACCIDENTS AT Y JUNCTIONS.			
	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION STATISTICS	
1970 RESULTS	X_o	13.74		$\bar{Y}$	14.30
	X1	15.29	1.06	S.D.	19.90
				S.E.E.	10.48
	X15	-1.80	1.05	R^2	0.731
	X19	-3.96	1.08	F	85.18
				D.F.	3,94

	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
1969 RESULTS	X_o	14.09		$\bar{Y}$	14.24
				S.D.	18.37
	X1	15.54	0.81	S.E.E.	8.01
	X6	2.33	0.89	R^2	0.820
	X15	-3.02	0.90	F	98.39
	X19	-2.69	0.82	D.F.	4,93
	X21	-1.91	0.81		

Y	REGRESSION: TOTAL ACCIDENTS AT X JUNCTIONS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	.REGRESSION	STATISTICS
	X_o	117.95		$\bar{Y}$	124.10
	X1	174.71	6.21	S.D.	209.16
				S.E.E.	46.63
	X2	-16.63	5.05	R^2	0.954
	X10	-12.81	4.75	F	265.99
	X15	64.73	5.45	D.F.	7,90
	X17	-48.28	8.00		
	X18	-18.29	7.93		
	X19	-11.01	4.89		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	126.17		$\bar{Y}$	127.32
				S.D.	224.44
	X1	201.05	6.82		
				S.E.E.	48.40
	X2	-15.24	5.24	R^2	0.957
	X4	-12.58	4.97	F	249.59
	X15	50.46	5.61	D.F.	8,89
	X16	-18.24	5.74		
	X17	-30.10	8.56		
	X18	-27.54	9.32		
	X19	-13.83	5.21		

Y	REGRESSION: TOTAL ACCIDENTS ON UNCLASSIFIED ROADS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	281.72		$\bar{Y}$	290.50
	X1	387.13	7.90	S.D.	376.20
				S.E.E.	63.88
	X15	22.27	7.18	R^2	0.972
	X18	-42.69	8.93	F	834.76
	X21	14.02	6.42	D.F.	4,95
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	286.71		$\bar{Y}$	286.71
				S.D.	382.05
	X1	389.57	7.78		
				S.E.E.	59.73
	X15	38.01	6.40	R^2	0.977
	X18	-31.30	8.08	F	988.69
	X21	15.39	6.01	D.F.	4,95

Y	REGRESSION: TOTAL ACCIDENTS ON CLASSIFIED ROADS.			

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
X_o	335.24		$\bar{Y}$	347.63
X1	372.48	16.29	S.D.	440.91
			S.E.E.	129.03
X2	-27.24	13.34	R^2	0.921
X15	107.31	13.77	F	220.11
X16	50.56	13.54	D.F.	5,94
X17	-86.98	19.33		

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	335.75		$\bar{Y}$	335.75
			S.D.	441.00
X1	431.05	13.79	S.E.E.	101.18
X4	-25.27	10.35	R^2	0.951
X7	21.04	10.51	F	255.55
X15	37.20	11.08	D.F.	7,92
X17	-41.60	17.15		
X18	-62.10	18.17		
X19	-35.76	10.73		

Y	REGRESSION: TOTAL CASUALTIES PER THOUSAND POPULATION.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	5.68		$\bar{Y}$	5.74
	X3	0.167	0.099	S.D.	1.102
				S.E.E.	0.952
	X7	-0.416	0.130	R^2	0.292
	X9	-0.292	0.105	F	7.770
	X16	0.222	0.109	D.F.	5,94
	X18	0.279	0.117		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	5.560		$\bar{Y}$	5.560
				S.D.	1.234
	X3	0.336	0.108	S.E.E.	0978
	X4	-0.225	0.102	R^2	0.416
	X5	-0.306	0.105	F	9.357
	X6	0.244	0.109	D.F.	6,93
	X9	0.183	0.102		
	X12	0.344	0.115		
	X18	0334	0.114		

Y	REGRESSION: TOTAL ACCIDENTS PER THOUSAND POPULATION.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	4.466		$\bar{Y}$	4.492
	X5	-0.205	0.082	S.D.	0.897
				S.E.E.	0.775
	X9	-0.216	0.088	R^2	0.285
	X16	0.211	0.088	F	9.475
	X18	0.226	0.093	D.F.	4,95
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	4.417		$\bar{Y}$	4.417
				S.D.	0.964
	X5	-0.184	0.083	S.E.E.	0.806
	X12	0.226	0.090	R^2	0.362
	X13	-0.196	0.082	F	10.65
	X19	-0.188	0.083	D.F.	5,94
	X21	0.193	0.083		

Y	REGRESSION: TOTAL FATAL AND SERIOUS CASUALTIES PER THOUS. POP.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	1.577		$\bar{Y}$	1.585
	X18	0.181	0.050	S.D.	0.530
				S.E.E.	0.490
	X21	0.108	0.050	R^2	0.162
				F	9.406
				D.F.	2,97
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	1.512		$\bar{Y}$	1.512
				S.D.	0.561
	X7	0.115	0.051		
				S.E.E.	0.449
	X18	0.190	0.050		
				R^2	0.232
	X21	0.132	0.051		
				F	9.663
				D.F.	3,96

Y	REGRESSION: CASUALTY RATE PER MILLION VEHICLE MILES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	18.83		$\bar{Y}$	18.48
	X1	-8.02	1.19	S.D.	13.14
				S.E.E.	6.38
	X5	1.70	0.71	R^2	0.784
	X10	-2.47	0.64	F	41.23
	X12	5.44	1.79	D.F.	8,91
	X15	-2.61	1.03		
	X16	-9.33	0.97		
	X18	-4.56	1.02		
	X22	3.96	1.15		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	18.99		$\bar{Y}$	18.99
				S.D.	14.91
	X1	-3.58	1.21	S.E.E.	8.49
	X4	-2.46	0.86	R^2	0.694
	X16	-7.29	1.01	F	35.39
	X18	-6.25	1.22	D.F.	6,93
	X21	-1.87	0.89		
	X22	4.68	1.00		

Y	REGRESSION: ACCIDENT RATE PER MILLION VEHICLE MILES.			

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
X_o	14.89		$\bar{Y}$	14.47
X1	-4.17	0.78	S.D.	10.45
			S.E.E.	5.67
X10	-1.94	0.56	R^2	0.739
X11	-1.47	0.61	F	37.30
X14	-1.59	0.61	D.F.	7,92
X16	-5.74	0.80		
X18	-3.80	0.80		
X22	1.99	0.67		

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	14.53		$\bar{Y}$	14.53
			S.D.	9.88
X1	-3.40	0.69	S.E.E.	4.79
X4	-2.46	0.61	R^2	0.779
X11	-1.10	0.62	F	54.64
X16	-5.58	0.51	D.F.	6,93
X18	-3.89	0.69		
X22	1.89	0.55		

Y	REGRESSION: FATAL AND SERIOUS CAS. RATE PER MILLION VEH/MILES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	5.31		$\bar{Y}$	5.21
	X1	-1.45	0.40	S.D.	4.41
				S.E.E.	2.89
	X10	-0.82	0.29	R^2	0.594
	X16	-2.07	0.34	F	27.47
	X18	-1.25	0.41	D.F.	5,94
	X22	1.17	0.34		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	5.08		$\bar{Y}$	5.08
				S.D.	4.08
	X1	-1.28	0.36		
	X4	-0.85	0.26	S.E.E.	2.54
	X16	-2.01	0.30	R^2	0.634
	X18	-1.19	0.37	F	32.54
	X22	1.00	0.29	D.F.	5,94

Y	REGRESSION: FATAL AND SERIOUS CASUALTIES/TOTAL CASUALTIES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.276		$\bar{Y}$	0.276
	X15	-0.018	0.007	S.D.	0.077
				S.E.E.	0.075
				R^2	0.059
				F	6.156
				D.F.	1,98
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.272		$\bar{Y}$	0.272
				S.D.	0.081
	X22	0.199	0.007	S.E.E.	0.078
				R^2	0.071
				F	7.447
				D.F.	1,98

Y	REGRESSION: CHILD PED. CAS./TOTAL PEDESTRIAN CASUALTIES.			

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
X_o	0.484		$\bar{Y}$	0.475
X2	0.036	0.01	S.D.	0.099
X6	-0.022	0.01	S.E.E.	0.066
X7	0.058	0.01	R^2	0.589
X8	0.032	0.01	F	22.22
X9	0.021	0.01	D.F.	6,93
X13	0.027	0.01		

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	0.470		$\bar{Y}$	0.470
X5	0.046	0.008	S.D.	0.103
X2	0.035	0.010	S.E.E.	0.076
X9	-0.042	0.009	R^2	0.476
X14	0.024	0.008	F	21.57
			D.F.	4,95

Y	REGRESSION: YOUNG DRIVER AND RIDER CAS./TOTAL DR. AND RID. CAS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.504		$\bar{Y}$	0.505
	X5	-0.024	0.006	S.D.	0.064
				S.E.E.	0.059
	X99	0.015	0.006		
	X14	0.011	0.006	R^2	0.202
				F	6.000
	X21	0.012	0.006	D.F.	4,95
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.527		$\bar{Y}$	0.527
				S.D.	0.083
	X5	-0.023	0.008		
				S.E.E.	0.079
	X6	0.015	0.008		
				R^2	0.121
				F	6.676
				D.F.	2,97

Y	REGRESSION: TOTAL TURNING ACCIDENTS/TOTAL ACCIDENTS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.269		$\bar{Y}$	0.273
	X1	-0.033	0.006	S.D.	0.075
				S.E.E.	0.058
	X5	-0.020	0.006	R^2	0.408
	X21	0.032	0.008	F	19.736
				D.F.	3,86

1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.274		$\bar{Y}$	0.279
				S.D.	0.078
	X1	-0.034	0.007		
				S.E.E.	0.061
	X3	0.021	0.007		
				R^2	0.430
	X9	0.011	0.007		
				F	12.660
	X21	0.026	0.009		
				D.F.	5,84
	X22	-0.015	0.007		

Y	REGRESSION: TOTAL ACCIDS. AT PED. CROSSINGS/TOTAL ACCIDENTS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.027		$\bar{Y}$	0.028
	X6	0.007	0.004	S.D.	0.040
				S.E.E.	0.035
	X9	0.010	0.004	R^2	0.262
	X10	-0.008	0.004	F	6.461
	X11	0.014	0.004	D.F.	5,91
	X17	-0.017	0.004		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.020		$\bar{Y}$	0.019
				S.D.	0.016
	X1	0.003	0.002	S.E.E.	0.014
	X16	0.006	0.002	R^2	0.240
	X18	0.004	0.002	F	7.245
	X21	-0.003	0.001	D.F.	4,92

Y	REGRESSION: TOTAL ACCS.ON UNCLASSIFIED ROADS/MILE UNCLASS.ROAD.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	1.694		$\bar{Y}$	1.696
	X1	0.272	0.050	S.D.	0.607
				S.E.E.	0.474
	X11	-0.095	0.050	R^2	0.414
	X16	0.159	0.050	F	16.745
	X21	0.210	0.050	D.F.	4,95
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	1.665		$\bar{Y}$	1.664
				S.D.	0.623
	X3	0.123	0.062		
	X5	0.107	0.046	S.E.E.	0.411
	X8	0.517	0.073	R^2	0.600
	X14	0.165	0.054	F	17.067
	X16	0.089	0.051	D.F.	8,91
	X18	0.130	0.046		
	X21	0.191	0.049		
	X22	0.256	0.067		

Y	REGRESSION: TOTAL ACCIDENTS ON CLASS. RDS./MILE OF CLASS. RDS.

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
X_o	7.72		$\bar{Y}$	7.75
X5	0.829	0.290	S.D.	3.22
			S.E.E.	2.64
X12	2.882	0.490	R^2	0.356
X18	0.661	0.280	F	13.13
X22	1.289	0.470	D.F.	4,94

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	7.465		$\bar{Y}$	7.465
X4	-0.507	0.262	S.D.	3.220
X7	0.471	0.272	S.E.E.	2.600
X8	2.797	0.422	R^2	0.372
X22	1.260	0.388	F	14.044
			D.F.	4,95

Y	REGRESSION: TOTAL ROUND. ACCIDENTS/TOTAL JUNCTION ACCIDENTS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.053		$\bar{Y}$	0.051
	X6	-0.007	0.003	S.D.	0.037
				S.E.E.	0.032
	X7	0.011	0.004	R^2	0.259
	X8	-0.009	0.003	F	8.134
	X21	0.012	0.003	D.F.	4,93
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.047		$\bar{Y}$	0.047
				S.D.	0.034
	X5	0.010	0.003	S.E.E.	0.031
	X6	-0.006	0.003	R^2	0.216
	X9	0.009	0.004	F	6.415
	X21	0.007	0.003	D.F.	4,93

Y	REGRESSION: TOTAL T JUNCTION ACCIDS./TOTAL JUNCTION ACCIDENTS.			

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
X_o	0.582		$\bar{Y}$	0.574
X1	-0.039	0.010	S.D.	0.101
X7	0.052	0.010	S.E.E.	0.083
X16	-0.022	0.010	R^2	0.348
X21	-0.019	0.008	F	12.420
			D.F.	4,93

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	0.552		$\bar{Y}$	0.552
X1	-0.028	0.008	S.D.	0.093
X4	-0.014	0.008	S.E.E.	0.075
X5	0.039	0.008	R^2	0.377
X10	0.013	0.008	F	14.081
			D.F.	4,93

Y	REGRESSION: TOTAL Y JUNCTION ACCIDS./TOTAL JUNCTION ACCIDENTS.			

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
X_o	0.037		$\bar{Y}$	0.038
X2	-0.007	0.003	S.D.	0.028
			S.E.E.	0.027
X12	-0.007	0.002	R^2	0.131
X14	0.009	0.003	F	4.705
			D.F.	3,94

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	0.037		$\bar{Y}$	0.038
			S.D.	0.022
X2	-0.007	0.002	S.E.E.	0.019
X5	-0.007	0.002	R^2	0.285
X8	-0.007	0.002	F	7.329
X10	0.008	0.002	D.F.	5,92
X21	-0.004	0.002		

Y	REGRESSION: TOTAL X JUNCTION ACCIDS./TOTAL JUNCTION ACCIDENTS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.219		$\bar{Y}$	0.227
	X1	0.023	0.008	S.D.	0.098
				S.E.E.	0.069
	X2	-0.014	0.008	R^2	0.544
	X6	0.018	0.008	F	15.330
	X7	-0.033	0.010	D.F.	7,90
	X9	-0.017	0.009		
	X11	0.018	0.008		
	X12	0.030	0.009		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.242		$\bar{Y}$	0.242
				S.D.	0.100
	X1	0.024	0.008		
	X5	-0.035	0.008	S.E.E.	0.077
	X10	-0.028	0.008	R^2	0.430
	X22	-0.022	0.007	F	17.506
				D.F.	4,93

Y	REGRESSION: TOTAL PED.CASUALTIES PER THOUSAND POPULATION.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	1.644		$\bar{Y}$	1.666
	X1	0.112	0.040	S.D.	0.444
				S.E.E.	0.278
	X3	-0.131	0.030	R^2	0.640
	X7	-0.137	0.040	F	20.184
	X10	-0.056	0.030	D.F.	8,91
	X11	-0.106	0.030		
	X16	0.098	0.030		
	X18	0.092	0.040		
	X22	-0.128	0.040		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	1.580		$\bar{Y}$	1.580
				S.D.	0.459
	X7	0.061	0.035	S.E.E.	0.332
	X8	0.112	0.041	R^2	0.508
	X14	0.172	0.035	F	15.995
	X16	0.161	0.040	D.F.	6,93
	X18	0.144	0.037		
	X19	-0.076	0.036		

Y	REGRESSION: TOTAL PEDESTRIAN CASUALTIES/TOTAL CASUALTIES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.291		$\bar{Y}$	0.292
	X3	-0.033	0.010	S.D.	0.068
				S.E.E.	0.045
	X11	-0.010	0.005	R^2	0.581
	X12	0.019	0.005	F	32.915
	X21	-0.012	0.005	D.F.	4,95
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.287		$\bar{Y}$	0.287
				S.D.	0.072
	X3	-0.027	0.006	S.E.E.	0.051
	X10	0.020	0.005	R^2	0.511
	X16	0.021	0.005	F	24.770
	X21	-0.018	0.006	D.F.	4,95

Y	REGRESSION: TOTAL P/CYCLE ACCIDS./PER THOUS. P/CYCLE WORK TRIPS.			

1970 RESULTS

SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
X_o	31.01		$\bar{Y}$	31.059
X3	-5.09	2.07	S.D.	22.740
			S.E.E.	17.974
X4	-4.15	1.79	R^2	0.407
X5	4.24	1.94	F	12.890
X15	-7.90	1.80	D.F.	5,94
X21	-7.28	2.00		

1969 RESULTS

SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
X_o	31.28		$\bar{Y}$	31.280
			S.D.	23.710
X10	3.90	1.939	S.E.E.	19.140
X15	-11.60	2.136	R^2	0.375
X21	-8.31	1.984	F	14.251
X22	-5.91	2.051	D.F.	4,95

Y	REGRESSION: TOTAL PEDESTRIAN CAS.PER HUNDRED WALK WORK TRIPS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	1.908		$\bar{Y}$	1.918
	X1	0.309	0.060	S.D.	0.686
	X2	-0.111	0.050	S.E.E.	0.450
	X4	-0.093	0.050	R^2	0.599
	X5	0.090	0.050	F	19.667
	X6	-0.106	0.050	D.F.	7,92
	X16	0.322	0.050		
	X17	-0.196	0.050		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	1.806		$\bar{Y}$	1.806
	X1	0.227	0.066	S.D.	0.661
	X6	-0.104	0.054	S.E.E.	0.503
	X16	0.330	0.054	R^2	0.449
	X17	-0.146	0.071	F	15.322
	X21	0.087	0.051	D.F.	5,94

Y	REGRESSION: TOTAL PEDAL CYCLE ACCIDENTS/TOTAL ACCIDENTS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.180		$\bar{Y}$	0.180
	X1	-0.045	0.010	S.D.	0.115
				S.E.E.	0.098
	X5	-0.022	0.010	R^2	0.302
	X16	-0.017	0.010	F	10.250
	X21	0.042	0.010	D.F.	4,95
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.187		$\bar{Y}$	0.187
				S.D.	0.117
	X1	-0.040	0.010	S.E.E.	0.102
	X21	0.043	0.010	R^2	0.248
				F	15.965
				D.F.	2,97

Y	REGRESSION: TOTAL MOTOR CYCLE ACCIDENTS/TOTAL ACCIDENTS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.121		$\bar{Y}$	0.123
	X3	0.015	0.004	S.D.	0.042
				S.E.E.	0.036
	X4	0.007	0.004	R^2	0.301
	X5	-0.014	0.004	F	9.200
	X6	0.007	0.004	D.F.	4,95
	X7	-0.009	0.005		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.127		$\bar{Y}$	0.127
				S.D.	0.043
	X1	-0.018	0.005	S.E.E.	0.038
	X5	-0.010	0.005	R^2	0.277
	X6	0.011	0.004	F	5.932
	X17	-0.009	0.005	D.F.	6,93
	X19	-0.007	0.004		
	X21	0.010	0.004		

Y	REGRESSION: TOTAL DRIVER AND RIDER CAS. PER MILLION VEH/MILES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	8.92		$\bar{Y}$	8.726
	X1	-2.36	0.52	S.D.	6.770
				S.E.E.	3.712
	X3	0.69	0.41	R^2	0.718
	X10	-1.03	0.37	F	39.380
	X16	-3.50	044	D.F.	6,93
	X18	-2.43	0.53		
	X22	1.60	0.45		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	8.62		$\bar{Y}$	8.620
				S.D.	6.430
	X1	-2.93	0.613	S.E.E.	3.240
	X3	1.36	0.301	R^2	0.772
	X4	-1.80	0.421	F	30.040
	X6	0.710	0.394	D.F.	10,89
	X8	2.44	0.889		
	X11	-0.99	0.430		
	X15	-1.36	0.580		
	X16	-4.19	0.541		
	X18	-2.21	0.509		
	X22	1.86	0.507		

Y	REGRESSION: TOTAL DRIVER AND RIDER CASUALTIES/TOTAL CASUALTIES.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.448		$\bar{Y}$	0.447
	X1	-0.018	0.006	S.D.	0.084
				S.E.E.	0.059
	X3	0.037	0.007	R^2	0.547
	X5	-0.021	0.007	F	18.750
	X6	0.015	0.006	D.F.	6,93
	X12	-0.014	0.007		
	X21	0.026	0.007		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.449		$\bar{Y}$	0.449
				S.D.	0.095
	X1	-0.039	0.007	S.E.E.	0.068
	X3	0.031	0.008	R^2	0.506
	X6	0.014	0.007	F	24.310
	X21	0.031	0.008	D.F.	4,95

Y	REGRESSION: TOTAL CASUALTIES/TOTAL ACCIDENTS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	1.270		$\bar{Y}$	1.270
	X5	0.047	0.010	S.D.	0.085
				S.E.E.	0.067
	X6	-0.016	0.006		
				R^2	0.423
	X7	-0.022	0.009	F	13.790
	X16	-0.012	0.006	D.F.	5,94
	X21	-0.015	0.006		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	1.262		$\bar{Y}$	1.262
				S.D.	0.080
	X6	-0.015	0.007		
				S.E.E.	0.065
	X7	-0.014	0.007		
				R^2	0.407
	X8	-0.025	0.009		
				F	6.851
	X10	-0.015	0.007		
				D.F.	9,90
	X11	0.014	0.007		
	X12	0.016	0.008		
	X16	-0.021	0.007		
	X20	-0.022	0.007		
	X21	-0.018	0.007		

Y	REGRESSION: TOTAL JUNCTION ACCIDENTS/TOTAL ACCIDENTS.				
1970 RESULTS	SIGNIF VARS.	B. COEFFICIENT	STAND. ERROR.	REGRESSION	STATISTICS
	X_o	0.643		$\bar{Y}$	0.647
	X2	-0.017	0.007	S.D.	0.093
				S.E.E.	0.070
	X5	-0.031	0.007	R^2	0.474
	X6	0.015	0.007	F	16.557
	X12	0.032	0.007	D.F.	5,92
	X21	-0.014	0.007		
1969 RESULTS	SIGNIF. VARS.	B. COEFFICIENT	STAND. ERROR	REGRESSION	STATISTICS
	X_o	0.660		$\bar{Y}$	0.661
				S.D.	0.087
	X8	0.045	0.007	S.E.E.	0.067
	X10	-0.013	0.007	R^2	0.425
	X19	-0.015	0.007	F	17.169
	X21	-0.016	0.007	D.F.	4,93

ACKNOWLEDGMENTS

The author wishes to express his gratitude to all those people who have helped in the completion and presentation of this research, both as regards academic and financial considerations, within the Department of Civil Engineering, Salford University.

In particular the author would like to thank Dr. S. Raymond for his unswerving support throughout this research and also Mr. R. Thomas for his often invaluable advice as regards the statistical sections of this research.

The author would also like to express his gratitude to the Transport and Road Research Laboratory for the provision of most of the data as regards road accident statistics.

Finally, the author would also like to thank Mrs. F. Ashburn for her total commitment to the art of typing, without which the typing of this manuscript could have been long delayed.