N-62-4-5

U. S. ARMY
TRANSPORTATION RESEARCH COMMAND
FORT EUSTIS, VIRGINIA

TCREC TECHNICAL REPORT 62-11

THE STRUCTURE OF TRANSPORTATION NETWORKS

Task 9R98-09-003-01

Contract No. DA 44-177-TC-685

May 1962

by:

TRANSPORTATION CENTER
NORTHWESTERN UNIVERSITY
Evanston, Illinois

DISCLAIMER NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission, to manufacture, use, or sell any patented invention that may in any way be related thereto.

ASTIA AVAILABILITY NOTICE

Qualified requesters may obtain copies of this report from

Armed Services Technical Information Agency
Arlington Hall Station
Arlington 12, Virginia

This report has been released to the Office of Technical Services, U. S. Department of Commerce, Washington 25, D. C., for sale to the general public.

The information contained herein will not be used for advertising purposes.

The findings and recommendations contained in this report are those of the contractor and do not necessarily reflect the views of the Chief of Transportation or the Department of the Army.

HEADQUARTERS

U. S. ARMY TRANSPORTATION RESEARCH COMMAND

TRANSPORTATION CORPS

Fort Eustis, Virginia

FOREWORD

This report was prepared by the Transportation Center, Northwestern University, under the terms of Contract DA 44-177-TC-685. Conclusions and recommendations contained in the report are concurred in by this command.

FOR THE COMMANDER:

EARL A. WIRTH
CWO-4 USA
Adjutant

APPROVED BY:

JOE R. CLOYD
USATRECOM Project Engineer

Contract DA 44-177-TC-685
May 1962

THE STRUCTURE OF TRANSPORTATION NETWORKS

Transportation Forecast and Prediction Study
Progress Report

Prepared by:
The Transportation Center
at
Northwestern University
Evanston, Illinois

for

U. S. ARMY TRANSPORTATION RESEARCH COMMAND
FORT EUSTIS, VIRGINIA

PREFACE

This monograph is the second of a planned series of reports on basic research on the development of transportation systems. The first monograph dealt with the forecasting of transportation stock aggregates, both for nations as a whole and for the distribution of transportation facilities within nations. The present monograph deals with the structure of transportation systems. Attention is given to what literally may be termed the layout, network, geometry, or pattern of transportation facilities. The task of the present monograph is presented in detail in its first chapter. The remainder of this foreword will be used to present the over-all plan of research and to identify research topics and research design problems.

TRANSPORTATION, A FUNCTION OF CHARACTERISTICS OF AREAS

In the introduction to the previous volume* in this series it was pointed out that the surface transportation development of any country depends upon:

1. General economic development.
2. Natural environment.
3. Location of activities.
4. Available technology and relative cost structure.
5. The interests and preferences of those who make decisions affecting transportation.
6. Military and political influences.
7. The historical pattern of development and outlook for the future.

This list is not inviolable and a knowledgeable person could extend or modify it in useful ways. For instance, one could extend the list to the subject of the air transportation by mentioning the position of nations on international air routes. Though incomplete, the list serves an important purpose: it emphasizes that transportation development may be thought of as some function of characteristics of areas. By adopting this systematic view, transportation may be explained in terms of its relation to other factors, and thus capability to forecast and evaluate transportation developments may

*Transportation Geography Research, A report prepared for the U. S. Army Transportation Research Command, Fort Eustis, Virginia, by the Transportation Center at Northwestern University, July 1, 1960, under Contract DA 44-177-TC-574.

be developed.

The notion of functional relationships given in the paragraph above is a very simple one, but its simplicity should not disguise its significance. There has been little work on transportation development topics from this point of view, but this point of view is basic to the present studies. Because of lack of attention to external influences of which transportation is a function, our ability to make incisive statements of how and why transportation systems develop and what we anticipate in the way of development is small. In contrast, much useful work has been done on certain other transportation topics; for instance, knowledge of technical problems of construction is both voluminous and useful.

OPERATIONAL VIEWS OF TRANSPORTATION

The statement of functional relationships above is too general to give much information regarding the substantive character of the research. Indeed, this research may be described as an attempt to give substance to statements of functional relationships, and the character of the research may be displayed by commenting upon the ways it has been found practicable (or is felt will be practicable) to give operational meaning to functional statements.

Transportation may be viewed the following five ways:

1. Stock aggregates. Miles of road, number of cars, etc.
2. Structure or layout of transportation routes. What types of routes go where?
3. Flows. What goes where?
4. Intensity of use, or the transportation activity as a productive activity within an economy. What does transportation use as its inputs, and in what ways does it contribute to the output of the economy?
5. Relationships of transportation networks to each other. For instance, how does the railroad network relate to the highway network?

Transportation was viewed in the first way — as stock aggregates -- in the first monograph of this series. Transportation is viewed in the second manner in the present monograph, and the other views of transportation are the objects of present and planned research.

It might be noted at this time that by viewing transportation in several different ways we have "disaggregated" the concept of transportation, and the various views of transportation are not completely independent. The layout of the transportation network is certainly dependent upon flows, for instance. Thus, viewing transportation in any one of these ways just listed requires that we add other aspects of transportation to the list of conditions controlling the characteristics of transportation systems. For instance, the layout of a transportation network is determined by such factors as topography and the like, listed early in this Foreword, and layout is also influenced by the relations of networks to each other, flows, and other of the operational views of transportation.

Other observations will be made on the structure of the research problem following the discussion below of the operational definitions of transportation activities.

TRANSPORTATION RESEARCH TOPICS

STOCKS

Variations in surface transportation stocks from nation to nation and within nations were the subject of a previous research report. Consequently, the best available statement of how stocks may be studied and the degree to which inter- and intra-nation variations in stocks may be explained by functional relationships between stocks and various explanatory variables is available there. Briefly, it was found that stocks could be codified as mobile equipment (cars, buses, trucks, and railroad equipment) and fixed facilities (miles of railroad and road). The quantity of these transportation stocks is explainable in varying degrees by functional relationships between these stocks and variables such as income and population. Generally, it was found that the distribution of stocks is most closely related to income, population, and size of area.

STRUCTURE

As mentioned earlier, the term structure refers to the layout, geometry, or pattern of the transportation network or system, and these words will be used interchangeably. Research on structure treats such topics as the location of routes; location of intersections and/or terminals; density and length of routes; accessibility of individual points on a network to other points on the network; and distance that must be traveled in order to reach every point or "saturate" a network. Giving such notions empirical content is part of the research problem. Not only must these notions be measable, ways must be found to relate these notions to factors conditioning the development of transportation. Research on some of these subjects is presented in the present monograph.

FLOWS

The term flows refers simply to the movement of persons, commodities, or messages; flows are the activity of transportation. The fact that so many flows may be identified presents a major problem in the systematic study of flow systems. This problem is one of finding incisive measures of flows that are useful in a sense that they are both: (1) related to determining variables, and (2) general and of wide applicability. The ability to relate measures of flows to explanatory variables is necessary if functional statements are to be made about flow systems, and if methods are to be developed for the short and long run forecasting of the flow characteristics of transport systems. Measures of flows must be incisive so that once these measures are accomplished they, in turn, will yield information about any particular type of flow that it might be desired to study.

The study of flow systems is a part of the current research, but it is not reviewed in the present volume nor were flow studies reviewed in the previous report. It is possible to mention here, however, certain expectations of fruitful results from the study of flows. For one, there is a small but useful literature that treats flow systems using the mathematics of linear programming and, in some instances, making use of such economic concepts as comparative advantage, factor price equilibrium, interregional equilibrium, etc. The generality and preciseness of this approach strongly infers that it will provide systematic measures of flow conditions and methods of relating these flow conditions to external variables.

A second approach that promises to be useful makes use of the temporal sequence of flow activities. Transportation activities may be characterised, for instance, by the proportion of time they are operating at capacity. This measure of utilization would seem to be directly related to such external variables as the agricultural industry of a nation, industrial mix, production characteristics, and the regional distribution of these activities. In turn, this measure would provide considerable information on the transportation system itself -- type of system and equipment used, efficiency, etc.

INTENSITY

The functioning of a national economy requires the use of transportation. It is common knowledge that as economies develop specialized production increases and relatively more transportation is required. Related research questions treat the manner in which economies with different structures and levels of development vary in their use of transportation, and the relations between these variations and other determinants of transportation systems.

This problem demands more sophisticated study and introduces more problems than might be assumed at first glance. In part, this is because of the many ways transportation is used in an economy. One might prepare a list of all outputs of the economy -- beef, steel, automobiles, electric motors, etc. -- and the transportation required per unit of output for each of these industries might be stated. A convenient measure might well be the monetary value of transportation required for each dollars' worth of output of each commodity. The summary measure of transportation intensity within a country is, then, a list of numbers showing the input of transportation per unit of output for all activities. A set of these measures would then consist of a set of lists, one list for each nation. We are now faced by the problem of how these lists differ from nation to nation and how these differences may be explained by functional relations.

Each of these lists of numbers may be viewed as a vector, and the technical problem of working with vector measurements is not great. Experience indicates that a more difficult aspect of the research will be that of finding the measures that are the elements of the individual vectors. A set of vector measurements for Japan, Italy, and the United States was presented in the previous monograph (Table IV, page 17), and similar data are available for a number of other nations. However, anyone familiar with statistics of this type would realize that there are probably very serious questions of the comparability of these data. A chief research problem will be that of assuring comparable statistics.

RELATIONS OF NETWORKS TO EACH OTHER

This topic differs somewhat from the preceding four topics in that it is concerned with the relationships among transportation systems rather than being a way in which the notion of transportation may be codified. A well-known topic is that of the relationships between highway networks and railroad networks. In certain areas the highway network is supplementary to the railroad network. In others, the highway network both supplements and complements or even competes with the railroad network. These network relations are evidenced by and are a combination of stock conditions, characteristics of flows, network structure, and other use characteristics described earlier. Research is programmed on this topic because it complements the four research topics discussed above.

TWO DESIGN PROBLEMS

The materials above stress that the present research is oriented toward the functional relationships between transportation and its environment, and state in a brief manner five ways in which transportation may be viewed for purposes of the analysis. These materials give an overview, albeit sketchy, of the orientation and content of the research. Remarks on design problems also should be presented in order to complete this overview. Transportation is so intertwined with its economic and social environment that it poses especially difficult research design problems. This section contains a discussion of the implications of this fact for the design of the research.

Two cases at point will serve to introduce this topic. For one, any change in a transportation system changes the economic and social environment within which the system operates. Replacement of a primitive highway system by a more satisfactory railroad system in a developing nation, for instance, may affect the pattern of development profoundly. Places advantageously located on the railroad system may grow at rates different from those located on the more primitive transportation system; in fact many places on the latter system may actually decline in size. An example of this effect well known to everyone in this country is the set of changes occasioned by the paving of rural roads during the 1920's and 1930's. Relative positions of hamlets and rural market towns were shifted, patterns of trade shifted, as did other of the factors relating to transportation development and use.

A second characteristic of transportation systems is that decision rules for development change as the system develops. Consider a partially completed transportation network. Development of the network may be looked upon as a series of link (or capacity) additions. A link addition is made to the network and a further link addition is made at another time. From a formal point of view the second addition is not a replication of the first addition. The decision rule by which the system changed must also have changed. The situation is rather obvious and easy to see. Though easy to see, it poses tremendous difficulties for analysis. In scientific work, systems in which the operating rules change over time are much more difficult to study than those for which the operating rules are constant as the system evolves.

The two ideas above introduce a number of research complications. For one, they infer that a study of functional relationships stressed earlier must be done in a special way. These functional relationships do not involve variables which may be thought of as strictly explanatory. This is because the environment of the transportation system changes as transportation evolves. Not only do the environmental, or explanatory, variables change in relative magnitude depending upon the character of transportation changes, the rules that relate these variables to the transportation activity also change.

In short, we are faced with a situation where the rules implied by functional relationships differ depending upon the state of the development of the transportation system. Further, the transportation system interacts with so-called determining factors and these shift during the period of transportation development. These complications taken together pose an especially difficult situation for scientific analysis.

ACKNOWLEDGEMENTS

The subjects treated within this monograph have been discussed with many persons, and it is not practicable to acknowledge all of this assistance here. A number of persons assisted in a direct manner with certain of the studies. Mr. Karel Kansky has done extensive work with two of the indices that are used in Chapter 3 of this report and with related work to be reported in a later monograph. Mr. Anthony Burges provided the data and the background note used in the study of the Irish transportation network discussed in Chapter 5 of this study. Mr. Joseph Stowers gave assistance with the computations used in Chapters 3 and 4 of this monograph and in connection with the decision model discussed in Chapter 5 of the monograph. Professor Brian J. L. Berry of the University of Chicago also assisted with computations and with their interpretation. Professor Martin Beckmann of Brown University and Professor Leon N. Moses of Northwestern University have made a variety of suggestions useful in the study. Mrs. Susanne Carow and Mrs. Shelda Ryburn assisted with typing and proof reading. Professors William L. Garrison and Duane F. Marble were responsible for the research and the monograph.

CONTENTS

FIGURES

TABLES

Table | Page

Table Page

CHAPTER 1. INTRODUCTION

OVERVIEW AND EVALUATION

SUBJECT

This monograph is concerned with the structure, geometry, mesh, pattern, or layout of transportation networks. These words convey notions of the arrangements of routes, intersections, and terminals on the earth's surface, and these notions provide an adequate general identification of the concern of this research. Operational definitions are used later in this monograph when particular subjects are treated. For instance, a later section of this discussion treats changes in the structures of networks when transportation systems are expanded and operational definitions are given at that time consistent with the needs of the subject. Elsewhere in this monograph suitable operational definitions are given for the comparison of transportation networks with each other.

OBJECTIVE

The over-all objective of this research is the development of a workable theory which will explain transportation networks in the sense of the following discussion. Given a series of characteristics of an area on which the intensity and character of transportation depend, a systematic method is desired for stating the nature of transportation in that area. The nature of transportation within an area depends on such matters as the physical characteristics of the area, the economic activities, density of population, income, and the like. Further, given a change in the characteristics of an area, it is desired to know how the transportation system will change. A shift in economic activity, for example, might result in and be dependent upon a shift in the nature of transportation.

The objective of the research discussed in the present monograph is more limited than that of the research as a whole. This monograph presents research relating to one aspect of transportation, the structure of transportation networks. It is desired to know the manner in which structures of transportation networks are dependent upon conditions such as the physical characteristics of the areas served, and it is desired to know how networks will change given changes in the circumstances conditioning transportation. The objective of research reported here is the development of an ability to make such statements, and the objective of the present monograph is that of presenting preliminary work on this subject.

OVERVIEW

This monograph is divided into five parts (Chapters) with subjects as follows:

1. Introductory material.
2. A discussion of pattern analysis from a mathematical point of view.
3. A report of research on nation-to-nation comparative studies of transportation networks.
4. A report of preliminary research on within-country transportation patterns.
5. A discussion of a simple decision model which replicates in an elementary way the extension of transport networks.

The discussion presents two types of materials. In the remainder of the present chapter and in Chapter 2 concern is chiefly with general materials. These materials are deemed to be of great importance by the researchers. They present systematic ideas basic to the evaluation of networks. Chapter 1, the present chapter, is introductory in nature. Chapter 2 takes advantage of the simplicity of mathematical ideas and notation to present certain features of networks. While the mathematics used is extremely simple, every effort is made to substitute expository discussion for mathematical elegance.

Chapters 3, 4, and 5 are empirical in nature. They present studies of certain of the notions identified in the present chapter and in the second chapter of this monograph. It is to be emphasized that these empirical sections are strictly exploratory. These topics are the objects of more elaborate empirical studies currently taking place.

ACCOMPLISHMENTS

Several authors have commented upon the structure of transportation, networks and there have been several theoretical studies of the geometric properties of simple networks (References 2, 3, 8, 9, 14, 15, 25). Also, several studies of network structure have been made from planning points of view (References 7, 17). While these studies suggest directions of useful research, only one of these studies takes advantage of the mathematical apparatus used in the present study and no one of the studies uses extensive empirical information. Consequently, the work discussed in this monograph lies within a relatively unexplored study area.* In spite of the fact that this research has not followed well-explored directions of transportation analysis, it has been possible to develop useful ways to analyze the structure of transportation networks.

This research has developed ways to find partial answers to three questions and it has identified topics that must be investigated in order to find more satisfactory answers to these questions. These questions are discussed below.

1. What measures will describe the structure of transportation networks?

A number of measures were developed and evaluated in this study. While more work needs to be done on this subject, these measures proved successful in that they could be used in the process of answering questions 2 and 3 below.

2. In what ways do these measures or characteristics of a transportation network depend upon the characteristics of the area in which the transportation network lies?

It has been possible to find "strong" relationships between the measures used in this study and the characteristics of areas.

3. Given knowledge of network measures (say, measures estimated from knowledge of the characteristics of an area), can the associated transportation network be mapped?

It has been possible to do some work with the manner in which networks expand subject to measures dependent upon the characteristics of the area.

Put another way, the study provides a way to evaluate transportation networks through a two-stage analysis. First, given some information on the nature of an area (and this may be very crude

---

*This statement is not entirely true. The apparatus used in this research has been used for the study of electrical and communications networks and experience in these fields was available to the researchers. The reader is referred to Seshu and Reed (Reference 20) for a complete discussion of this topic.

information), it is possible to predict values or measures of the structure of transportation in that area. Second, from these values or measures of structure it is possible to map the transportation system. This provides a way of "understanding" the how and why of the development of transportation networks. Because it is possible to postulate changes in the characteristics of areas and solve for the transportation network in light of these changes, the study also provides a forecasting tool.

Any device of this sort developed in this study is subject to certain limitations, as will be noted throughout the ensuing discussion. In the present study these are not so much limitations upon the internal workings of the methods used in the study as they are limitations presented by what the methods do not incorporate. The work reported here is regarded as useful and useable. First priority on increasing its usefulness and useableness should be given to enlarging the scope of the approach.

THE STRUCTURE OF TRANSPORTATION NETWORKS

The materials below present certain features of the structure of transportation networks. All points made here are extremely elementary and serve to introduce topics treated in greater detail later in this monograph as well as to orient the reader to the points of view used in the research.

STRUCTURE

Maps of transportation networks may vary considerably in their content. At a minimum they display lines indicating the location of routes (and, by the absence of lines, indicate where routes are not located), intersections of routes, and angles between and lengths of routes. Because a map is a two-dimensional representation of the earth's surface, and this representation may be made by using a variety of map projections, angles between lines and lengths of lines may vary from map to map. Maps often contain additional information. For instance they may indicate types of routes (say, highways versus railroads or air routes) or route qualities (say, metal-surfaced roads versus dirt roads), and, in addition, maps may distinguish among intersections. Certain intersections may be towns and these named, and whether or not an intersection is also a transfer point may be indicated.

In short, maps vary in the amount of information they contain. In a similar manner, the study of network structure provides alternate choices of the characteristics of the network to be studied. A minimum is that of working with points and lines only. Less simplified work might use information on angles between routes, lengths of routes, capacity, etc.

In addition, one alternate available is to treat networks at different levels of aggregation, either (1) using measures of the characteristics of entire networks or (2) using measures of relationships among links (or nodes) on the network. There is another alternate of course. It is to study individual links without reference to other links on the system. The latter is not a viable alternate at the level of generalization at which the present research takes place.

The ensuing discussion illustrates questions that may be asked from the points of view of these alternates.

THE ENTIRE NETWORK: MINIMAL SYSTEM

Figure 1 shows the internal airline routes of Guatemala and Figure 2 shows the internal airline routes of Honduras.

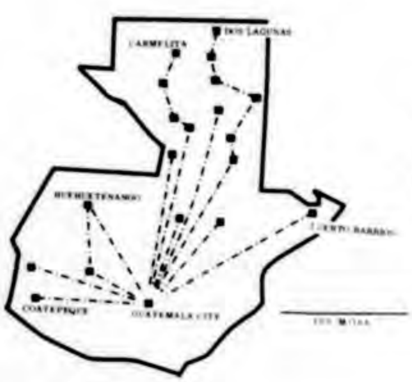


Figure 1. Internal Airline Routes of Guatemala.

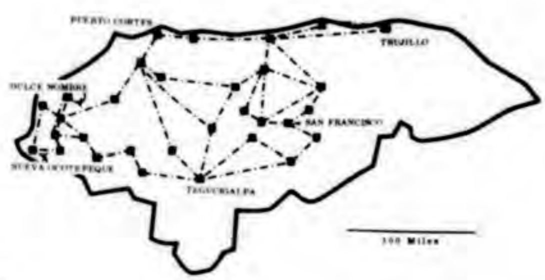


Figure 2. Internal Airline Routes of Honduras.

The maps are minimal in the sense that they show only the existence (or non-existence) of routes, and the location of terminals. The maps also show the lengths of routes, of course, but map information on length of route is somewhat difficult to interpret. Length of route in combination with information on amount hauled over the route provides a metric of the tie between two places on the network. Information is not available on simple maps, such as the sample maps reproduced here, to indicate amount of traffic, so length of route does not show the strength of ties between places. Also, it is widely known that cost of transportation is a non-linear function of distance. Stated another way, the cost of moving the first mile on a trip between two points is not the same as the cost of moving the second mile. Generally, terminal and other fixed costs make it much more costly to move the first mile than to move each succeeding mile. Still a further observation might be made regarding the distance measure. In the case of air transportation, the distance between two points is not a fixed number, it depends upon the choice of route by the pilot. In the case of North Atlantic routes, for instance, flight path distances may vary greatly from day to day and season to season. For similar reasons the time required to travel between two points is a variable and the cost of moving between two points is a variable. The interpretation of distance on a map, then, would seem to require transforming the distance onto some useable scale. The transformation would depend upon whether or not distance is given a cost interpretation, and it might require regarding the distance as a variable and recording its mean and variance.

The more elementary consideration of the existence or non-existence of terminals and the existence or non-existence of routes seems somewhat easier to interpret than the questions of route length. For the moment, attention will be given to questions of the layout of transportation networks viewed in the simplest manner -- the existence of terminals and the existence or non-existence of routes.

Surely analysts would agree that the sample networks (Figures 1 and 2) differ in layout. It is equally clear that different analysts would have difficulty communicating just what they meant by their statements that the systems differ. This difficulty reveals one of the central problems in this research; that is, the establishment of meaningful ways to codify structures of transportation networks.

This problem of the codification of structure is treated extensively in Chapter 2 of this monograph. However, a few introductory remarks will be made on the problem of codification. For one, we may note that the sample maps differ in the degree to which the systems are complete. In the case of airline routes, a complete system may be defined as one in which each terminal has a direct route to every other terminal. (This definition of a complete system is not satisfactory for certain other kinds of systems; this point is discussed in Chapter 2 when the difference between planar and non-planar graphs is discussed.) A measure of completeness may be concocted by comparing the actual number of routes with the number of routes that would be required to link every terminal directly with every other terminal.

It might also be observed that paths between places differ. Path differences may be measured in a number of ways. Note that in going from Dos LaGunas to Carmelita in Guatemala one must use ten links. A trip from San Francisco to Nueva Ocotepeque in Honduras would require the use of nine links, though longer paths could be used. The networks are similar when the lengths of their longest paths are compared. However, there is a very striking difference in the "centrality" of the two networks. The paths in Guatemala give the impression of focusing on the city of Guatemala while those in Honduras do not give the impression of focusing on any particular city. Other differences are the number of loops and in the number of endpoints (places with only one link) on the systems.

PLACES ON NETWORKS; MINIMUM SYSTEM

In the section above attention was given to certain apparent over-all differences between networks. The map patterns also reveal differences between particular places on the networks. On the Guatemala map, for example, it may be noted that most terminals are at the intersection of two links, a number of terminals have only one link, and no terminals have more than two links except for Guatemala City which has eleven links. The distribution of links by terminals is quite different for Honduras where a number of places have three or more links. This observation suggests that measures may be made of the accessibility of individual places to the entire network. A simple measure would be the number of links to a place. A more general measure might be the number of links that must be traversed in order to reach every other place on the network from a particular place.

RESEARCH QUESTIONS

The materials above introduce the notion of the structure of transportation networks and illustrate: (1) how questions about structure may be raised regarding the network as a whole, (2) how questions of structure may be raised regarding the relations between individual intersections (or links) and the network as a whole, and (3) that structure may be treated using different levels of information. In regard to the latter point, it was noted that minimal identification of structure requires working with a system where intersections or terminal points are identified and routes between these intersections or terminal points are stated. It was also noted that structure may be given more precise meaning by using information on links or routes, angles between routes, and by differentiating between types of links (say, identifying different capacities or modes) and types of terminals or intersections (intersections without terminal facilities, interchanges between modes, interchanges between the same mode). The possibility of looking at structure from two different points of view and the possibility of using different levels of information conditions research on structural questions, as will be noted in the ensuing discussion of research questions.

A THEORETICAL APPARATUS

One salient feature of the voluminous material on transportation is its heavy dependence upon descriptive verbal expression and the lack of exact definition and generality in this expression. (A notable exception to this statement, of course, is the technical literature of engineering.) Descriptive materials varying in completeness are available on most transportation systems, but the nature of these materials is such that they are not suitable for systematic study.

It must be stressed that this is a criticism of the literature from the point of view of its use for theoretical study. The literature has other uses and its value from other points of view is not denied.

The most serious deficit of the literature from our point of view is its preoccupation with incommensurate specifics. Fragments of information are available on capacity, orientation, and other characteristics of individual routes in nearly every system. Yet there is no calculus available through which these fragments of information may be manipulated. There are no good ways to compare information among systems or to put together pieces of information about an individual system and make statements about the system as a whole.

Another characteristic of the literature is its heavy dependence upon primitives. For instance, there is much use in the literature of such expressions as main line, feeder routes, tributary areas, service, capacity, circumferential routes, bypasses, and the like. Use of jargon of this type is not to be decried: it is a way to give verbal description to system characteristics. It is tempting to say that the problem with these primitives arises because of lack of their exact definition, but this is not quite true. There has been considerable work with exact definition of some of these terms, especially capacity. Experience elsewhere with primitives has shown that they may be used and used effectively without their ever having an exact identification. Electricity and gravity are examples of primitives that have been used without anyone ever deciding exactly what they are. The difference between these primitives and those illustrated earlier relating to transportation systems is that the latter are operationally useful. This observation reveals one of the high priority problems in developing more incisive material on transportation systems; it is to give primitive notions about transportation systems operational meaning.

The problem is more than that of giving operational meaning to primitives already available. Upon examination some of the notions currently available would not seem to warrant further concern, and there is need for notions of wider scope than many notions currently available.

In order to meet needs for general and useable notions regarding transportation systems, a large proportion of the present research has been given to development of a suitable theoretical apparatus for the study of transportation structure. An appropriate mathematics is available — graph theory — for the study of minimal elements of structure; namely, routes and intersections. The representation of transportation systems using this type of mathematics is presented in Chapter 2 of this monograph. This work is regarded by researchers as of great importance. Experience indicates that general and incisive results are to be expected only if they are set within a general framework of this type. Development of an appropriate calculus and language for the study of transportation systems is urgently needed.

COMPLEMENTARY EMPIRICAL STUDIES

There is quite literally no end to the alternates available for the development of a general apparatus for this study of transportation systems, and a choice must be made from among these alternates. Empirical work utilizing the implications from the theoretical work provides a guide for the selection of alternates. Chapters 3, 4, and 5 of this monograph present the results of certain empirical studies complementing the work at the theoretical level. This work: (1) makes a direct examination of certain of the simple notions that may be derived from elementary theoretical considerations and (2) compares elementary theoretical considerations with alternates. These alternates are in part forced from empirical evidence and are in part the empirical expression of informal ideas.

The empirical work is presented in three chapters. In Chapter 3 regression studies are presented that compare alternate notions about the structure of transportation systems and relate these to determining variables. These regression studies provide an empirical measure of the efficacy of various ways of measuring and considering transportation networks, as well as statements of how these characteristics of transportation networks are related to the environment of the transportation network. The effect of level of economic development on the structure of networks is considered, for example.

Chapter 4 is used to present the result of components analyses of two networks of air routes. These analyses may be thought of as forcing ideas about networks from empirical materials.

Examination of maps and the connection matrices of the maps (such matrices will be discussed in Chapter 2) reveals differences in the degree which networks are dominated by places with high levels of service. The components analyses represent a method of forcing a concept of relative dominance from empirical materials.

The final chapter, Chapter 5, presents the results of an exploratory study using a decision model designed to replicate the development of transportation networks. This study had two general purposes. For one, the study was an investigation of whether or not decision rules (represented by the transformation matrices of Chapter 2) are subject to systematic changes (represented by changes in the transformation matrices) as transportation networks develop.

The second reason for undertaking a study using the decision model is revealed by the following discussion. Studies of the history of transportation networks reveal that each network addition may be thought of as unique. The availability of financing, topographic conditions, market considerations, and a long list of other considerations were brought to bear on the decision to develop individual links. The decision model tests the speculation that these complex decisions might be reproducible to a large degree by a few simple considerations or rules. To the degree that this is true, these simple rules may be substituted for the complex set of circumstances that surround the development of particular routes, and may be thought of as an explanation for the pattern of system development.

CHAPTER 2. MATHEMATICS OF STRUCTURE

A basic concern in this research is finding adequate simplified representations of transportation systems that may be readily manipulated in the quest for information and understanding. A simplified representation of a real world system is commonly called a model. A model may be in one of two classes: physical or symbolic. Perhaps the most familiar of these are the physical models such as a three-dimensional relief representation of an area, or an aeronautical engineer's wind tunnel design model. Somewhat less familiar, but of much greater importance, are the symbolic or mathematical models. In this class of models the system under study is represented by a series of symbolic or mathematical statements. These symbolic statements may then be manipulated in order to determine the way in which the real world system will react to specific changes in system parameters. The model cannot be expected to reproduce all the details of the original system since construction of any model of a system involves a certain loss of information. In model construction, this information loss must be balanced against the increased difficulty of manipulation encountered as the model's complexity, and hence its reproduction of detail in the original system, increases.

In the Foreword of this monograph it was noted that there are many characteristics of transport systems such as stocks, flows, structure or layout, etc. This chapter will briefly examine one form of mathematics which appears to provide a symbolic language especially appropriate for the analysis of the structure or layout of transportation systems. The branch of mathematics to be examined here is linear graph theory, a branch of the field known as combinatorial topology. Graph theory is concerned with systems of lines and points, and the underlying notion utilized in the following discussion is that a transportation system may be considered as a series of nodes or points, representing either urban centers or junction points on the system, and a series of routes or lines connecting them. At the most abstract level these might be the only features of the system admitted in the model, and information such as the distance between nodes would not be considered. Admittedly, this basic structure is simplified considerably from the real world system, but it is only through this process that comparatively simple systems can be obtained and manipulated to establish basic parameters. Then, through a process of continuing relaxation of assumptions, it is possible to proceed to the more complex cases which may appear to be characteristic of real world transportation systems.

BASIC CONCEPTS OF GRAPH THEORY

Linear graph theory is a branch of a larger field known as combinatorial topology. Much of the work in this area has been done by European researchers and currently the basic books in the field are the works of Konig (Reference 13) and Berge (Reference 4). At present the only extensive discussion available in English on this topic is one by Seshu and Reed (Reference 20) which deals with linear graphs and their application to electrical networks. More limited discussions are also available in the earlier works of Whitney (References 27, 28, 29,) and Veblen (Reference 26). In order to permit easy reference to a comprehensive discussion in English, the following outline of the basic concepts of graph theory will follow that established by Seshu and Reed.

PRIMITIVE NOTIONS

Basically a linear graph is a collection, or set, of line segments and points. The line segments are commonly known as edges, while the other basic elements of the graph are normally known as points or vertices. A variety of labels are frequently attached to these simple concepts. Edges, for instance, are also known as arcs, or routes, or one-cells, while vertices are frequently called points, nodes, junctions, or sometimes zero-cells. The two primitive concepts, edges and vertices, are combined to form what is called a linear graph. A linear graph is a collection, or set, of edges, no two of which have a point in common that is not a vertex. Collections or sets of this nature may be either finite or infinite depending upon the number of elements that they contain. In the following discussion only finite graphs will be considered.

CLASSIFICATION

Graphs may be broken down into several different classes. One of the most basic breakdowns is that of non-oriented and oriented graphs. In non-oriented graphs, the only operational concept is that of incidence, that is, the notion that the end-points of one or more edges may coincide with a vertex, i.e., that the edges are incident upon that vertex. The oriented, or directed, graph on the other hand also recognizes a sense of direction of the edges. In this case it is recognized that an edge is incident upon two vertices, and also that a sense of direction is implied from one vertex to the other. These concepts are illustrated in Figure 3.

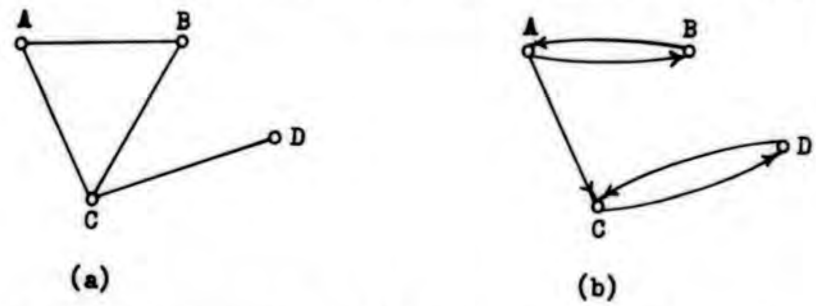


Figure 3. Non-oriented (a) and oriented (b) graphs.

Another classification of linear graphs, which may be treated as a further subdivision of two classes mentioned previously, is that of <u>weighted</u> or <u>non-weighted</u> graphs. So far, operations have been discussed in terms of a very simple set of concepts. Measurement has dealt only with binary relations, such as existence or non-existence and incidence or non-incidence, with no notions of quantity or any other type of metrization. It is possible to introduce metrics into the system so that a specific numerical value is associated with each edge and/or vertex. For instance distances between urban centers on a transport network might be associated with the edges of the graph representing that network. Each edge would then have associated with it a specific numerical value, or weight, and the resulting graph would be a weighted graph, or net. (See the discussion by Hohn, Seshu, and Aufenkamp (Reference 12).) In an even more complex case, it would be possible to assign weights to the nodes or vertices of the graph, as well as the edges. When this is done, the system's configuration closely resembles that encountered in the study of stochastic processes. (See Bartlett (Reference 1).)

ISOMORPHISMS

It is easy to see that a given linear graph may be structured in several different ways, e.g., relabeling the nodes and edges. In such a case it would be useful to have some precise way of recognizing that the graphs are really identical even though they may be arranged differently, and that their vertices and edges may bear different labels. This situation represents a mathematical <u>isomorphism</u>. Two graphs, say G and G*, can be said to be isomorphic if there is a one-to-one correspondence between the edges of G and G*, and a one-to-one correspondence between the edges of G and G* which preserves the incidence relationships. A one-to-one correspondence between the vertices of the two graphs means that each vertex in G can be associated with one, and only one, vertex in G^* and vice versa. A similar explanation holds for the one-to-one correspondence between the edges. It should be noted, of course, that this definition of isomorphic graphs contains two important conditions: the necessary condition that a one-to-one correspondence exists between vertices and edges of the two graphs, and the sufficient condition that the incidence relationships be preserved. (See Figure 4.)

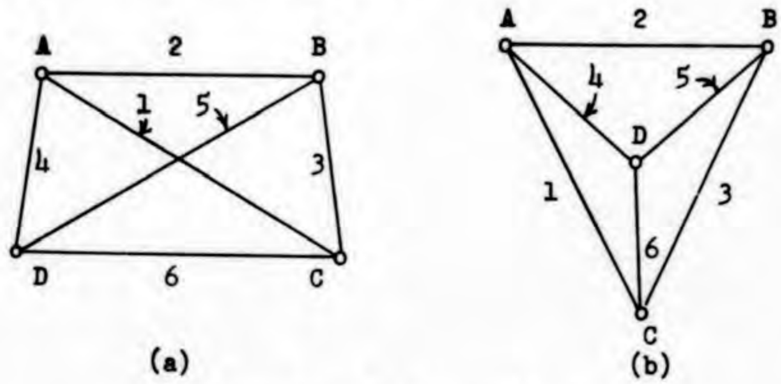
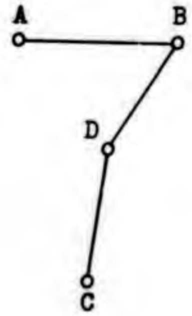


Figure 4. Two Isomorphic Graphs.

SUBGRAPHS AND EDGE SEQUENCES

Frequently it is desired to discuss not a complete graph, but rather some smaller portion of it. This subgraph normally consists of a set of the edges of the graph. If the edges of a subgraph can be arranged so that each edge has a vertex in common with the preceding edge in the ordered sequence, and the other vertex in common with the succeeding edge, then the subgraph so defined is known as an edge sequence. Of course it is possible that an edge may appear several times in an edge sequence. Where this occurs, the number of times the edge appears in the edge sequence is known as the multiplicity of the edge. If each edge in the edge sequence has a multiplicity of one, that is, it appears only once in the edge sequence, then the sequence is known as an edge train. The terminal vertices of the edge train are known as initial and final vertices, where the vertex of the first edge of the edge sequence that is not shared by the second edge is known as the initial vertex and, similarly, the vertex of the last edge that is not common to the previous edge is known as the final vertex. If the terminal vertices coincide, the edge train is closed, and if they do not coincide, it is open. (See Figure 5.)



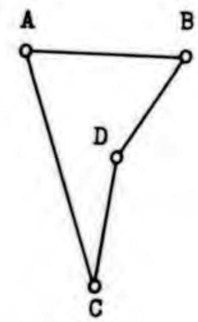


Figure 5. Open Edge (a) and Closed Edge (b) Trains.

CONNECTIVITY

Looking at the vertices that appear in an edge sequence, or for that matter elsewhere in the graph, it is possible to count the number of edges that are incident at a particular vertex. This number is known as the degree of the vertex. Using this notion in combination with the notion of an edge train, it is possible to define a path as an edge train where the degree of each internal vertex is two and the degree of each terminal vertex is one. Similarly, a circuit or loop is defined as a closed edge train where all vertices are of degree two.

Now, using these concepts the very important notion of connectivity may be introduced. A graph G is said to be connected if there exists a path between any two vertices of the graph. Thus from an intuitive standpoint, we may feel that a graph is connected if it is in "one piece." Suppose the graph is not connected; then this means that there are pairs of points or vertices in the system which cannot be joined in a path. The graph is then unconnected and it is intuitively obvious that it must consist of a number of "connected pieces." These "pieces" of the larger graph are, by previous definition, subgraphs and are known by the special name of maximal connected subgraphs. The number of these maximal connected subgraphs in any finite graph, G, is denoted by p and, as a consequence $p = 1$ for a graph G if, and only if, G is connected. This count of the number of maximal connected subgraphs present represents one of the simplest descriptions of the structure of a graph, and provides an index that remains invariant under all isomorphic transformations. The technical name for this index is the zeroth Betti number. (See Figure 6.)

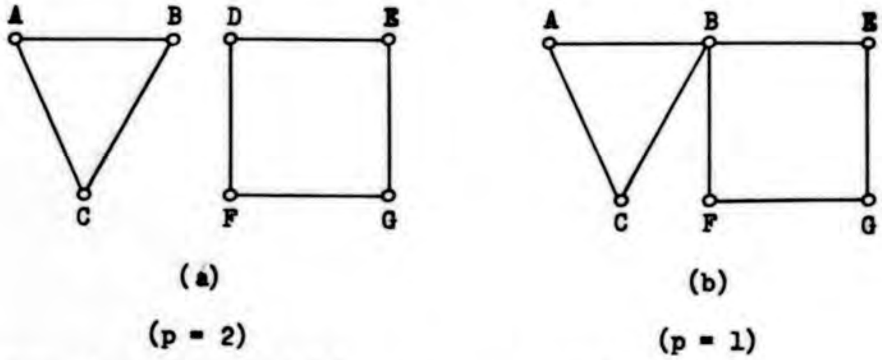


Figure 6. Unconnected (a) and Connected (b) Graphs.

TREES

One notion commonly encountered in graph theory is that of a tree. Intuitively a tree is thought of as a structure in one piece, with branches and branches off other branches, containing no closed paths or circuits. The term tree is used in graph theory in a very similar manner. In graph theory, a tree is defined as a connected subgraph of a connected graph which contains all the vertices of the graph but which does not contain any circuits. (This is a complete tree, and its definition differs somewhat from the strict definition of a tree utilized in mathematics.) A given finite graph is a tree if, and only if, there exists exactly one path between any two vertices of the graph. (See Figure 7.)

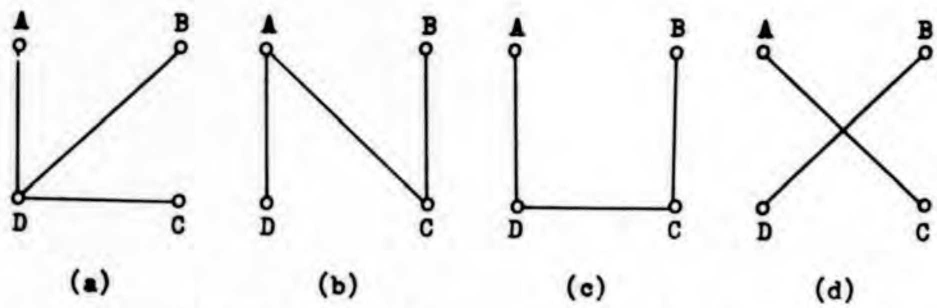
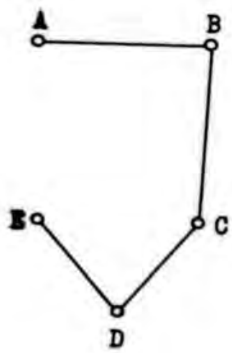


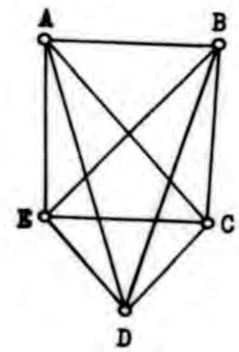
Figure 7. Some Trees of Figure 4(a).

It can be shown by induction that if a tree contains v vertices, it contains v-1 edges. For instance, in a transport system the smallest number of routes that will completely connect five urban places is four. Conversely, it can be shown that the maximum number of routes between n points is (n (n-1))/2; that is, ten transport routes would be required to completely connect a system containing five urban places. (See Figure 8.)



(a)

($v = 5$, $e = 4$)



(b)

($v = 5$, $e = 10$)

Figure 8. Tree (a) and Completely Connected Network (b).

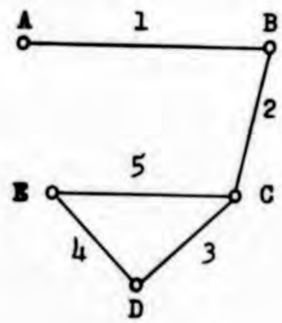
FUNDAMENTAL CIRCUITS

For a given graph and a given tree defined on the graph, elements of the graph may be divided into two classes, branches and cords. Branches are those elements which are contained in the tree, whereas cords are elements that are not in the tree and are therefore in its complement (or co-tree). It also may be shown that a connected graph consisting of v vertices and e edges contains v-1 branches and e-v+1 cords. The sum of branches plus cords should equal the total number of edges, i.e., e-v+1+v-1 should equal e, as it does. If one cord should be added to a tree, a graph is obtained that is no longer a tree. The cord and the path in the tree between the vertices of the cord constitutes a circuit. This is, moreover, a unique circuit and the only circuit of the resulting graph.

The fundamental circuits of a connected graph G for a tree T are the e-v+1 circuits consisting of each cord and its unique tree path. In a more general sense, this number is given by $\mu = e-v+p$, where v is the number of vertices, e the number of edges, and p the number of maximal connected subgraphs. The index μ is invariant under isomorphic transformations and is known as nullity, cyclomatic number, or first Betti number.

MATRIX REPRESENTATION

It has already been observed that the most fundamental characteristic of a graph is the relationship between the edges and the vertices. The graph is completely specified as soon as it is known which edges are incident on which vertices. Such a specification can be made through a simple diagram, such as has been used previously in the discussion, or even more conveniently by means of a matrix. In this matrix each row corresponds to a vertex and each column to an edge. If the edge is connected to or incident at a vertex, a one is placed in the cell entry; otherwise a zero. Precisely, the matrix is defined as follows: The vertex or <u>incidence matrix</u>, A is a matrix of v rows and e columns for a graph of v vertices and e edges, where $a_{ij} = 1$ if edge j is incident at vertex i, and $a_{ij} = 0$ if edge j is not incident at vertex i. A simple graph and its incidence matrix are shown in Figure 9.



(a)

$$A = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

(b)

Figure 9. A Graph and its Incidence Matrix.

From inspection of the matrix it can be seen that every column of the incidence matrix contains exactly two non-zero elements. This is a fundamental characteristic of the incidence matrix. The incidence matrix is equivalent to a given graph in the sense that each is determined completely by the other. Seshu and Reed note that if two graphs, G_1 and G_2, have incidence matrices which differ only by a permutation of rows and columns, then G_1 and G_2 are isomorphic; and conversely. The notion of isomorphism was introduced earlier in this discussion, and since a permutation of rows and columns merely involves an interchange of these items, it is now possible to list all possible isomorphic forms of a given graph by means of a relatively simple series of matrix operations.

<u>THE CONNECTION MATRIX</u>

The matrix representation which has proven most useful in many types of network analysis has not been the incidence matrix described above, but another matrix known as the <u>connection matrix</u>. (See Shimbel (References 21 and 22).) In a graph with v vertices, the connection matrix is a v x v matrix where each row and each column correspond to a specific vertex in the graph. The elements of the matrix are again zero or one depending upon the existence or non-existence of an edge directly connecting the two vertices. That is, $c_{ij} = 1$ if there is an edge which is incident at one end upon vertex i and at the other upon vertex j. The element $c_{ij} = 0$ if no direct connection exists. The elements upon the principal diagonal, the c_{ii}, which represent internal or self linkages are usually defined as either all zeros or all ones depending upon the structure of the problem being investigated. (See Figure 10.)

$$
C = \begin{array}{c c}
 & \begin{array}{ccccc} A & B & C & D & E \end{array} \\
\begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} &
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 0
\end{bmatrix}
\end{array}
$$

Figure 10. The Connection Matrix of Figure 9(a).

The connection matrix, usually denoted by C, is intimately related to and can be derived from the incidence matrix discussed above. This relationship involves the notion of matrix transposition. Matrix transposition is an operation wherein the rows and columns of the matrix have been interchanged; those entries that made up the first row of the matrix have now become the first column; those of the second row have become the second column; etc. The operation "transpose" on a matrix is normally denoted by placing a T as a superscript on the matrix notation so that, for example, the transpose of the incidence matrix is V^T. The relationship between the incidence matrix and the connection matrix is then: $C = VV^T$. The arithmetic used in this operation is not the usual base ten arithmetic, but is an arithmetic wherein the only elements that exist are zero and one, and one plus one is defined as equal to zero.

It is not necessary to have the incidence matrix in order to determine the connection matrix. Under normal circumstances it is possible to establish the elements of the connection matrix by simple inspection of the existing network. Looking at the connection matrix it is possible to see that the pattern of zeros and ones present represents the existence or non-existence of direct linkages between vertices in the system. The connection matrix may then be constructed by direct observation of the graph. (Compare Figures 9 and 10.)

In a system that is not completely connected there will be many places between which no direct link will exist. However, it is quite possible that these places may be reached by moving through one or more intermediate vertices, i.e., via some indirect route such as A to E via B and C in Figure 9. Given the connection matrix it is possible to determine how many indirect routes of any given length connect any two vertices in the system. For instance, if it is desired to know the number of two-link routes that exist between two vertices in the system, say i and j, we may do so by finding the square of the original connection matrix. The ijth element of the matrix C^2 is then interpreted as the number of two-link routes connecting vertex i and vertex j. (See Figure 11.)

$$C^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 1 \\ 1 & 0 & 3 & 1 & 1 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & 2 \end{bmatrix} \end{array}$$

Figure 11. Square of the Matrix C of Figure 10.

A similar procedure is followed for routes of greater number of lengths. C^5, for instance, will indicate how many five-link routes exist between each vertex and every other vertex. The entries in these cells, however, contain a number of redundant paths since many paths have been counted which contain edges with a multiplicity greater than one. However, it is possible to calculate the number of non-redundant paths (i.e., those containing only edges of multiplicity one) of any given length by means of a relatively complex mathematical manipulation. (See the work by Ross and Harary (Reference 19).)

DIAMETER AND SOLUTION TIME

Imagine that there are two vertices in the system, X and Y, which are quite remote from each other. That is, there are no direct links between them, no two-stage, no three-stage, no four-stage, etc., links. In this case, the corresponding element of successive powers of the connection matrix c_{xy} will remain at zero. Eventually, the entry in this cell only will change from zero to some non-zero number. If the two vertices are the "most remote" on the system, it can be seen that the matrix now contains no zeros. The power to which the original connection matrix has been raised to obtain this situation, is then known as the solution time of the network, or the diameter of the system. Put another way, the diameter or solution time of the system may be found by listing the number of links in the shortest path between each pair of nodes and selecting the largest of these numbers.

PLANAR AND NON-PLANAR GRAPHS

Up to this point, the discussion has been in terms of graphs as abstract objects. One especially important problem that arises in the application of graph theory to transportation networks is that of mapping the graph onto a plane. A graph which can be mapped onto a plane such that no two edges have a point in common that is not a vertex is known as a planar graph. Graphs which cannot be so mapped are known as non-planar graphs.

Why is this distinction important in the study of transportation networks? The answer to this question lies in the problem of involuntary intersections; that is, intersections created by the physical crossing of two or more routes being constructed to connect nodes in the system. In general, planar graphs correspond to that class of systems which may be constructed without creating involuntary intersections. Empirical examination of transportation systems indicates that surface routes, rail and highway, tend to have the characteristics of planar graphs, while airline routes appear more like non-planar graphs. This appears to be a reasonable conclusion since these systems normally tend to operate in two and three dimensions respectively.

THE DEVELOPMENT OF INDICES FOR GRAPHS

Given the fact that a transportation network may be regarded as a graph, it becomes interesting to develop certain summary indices which may provide useful information relating to the structure of the network. A certain amount of information has been lost in passing from the real-world system to its graph or matrix representation and further loss becomes necessary in order to assist in information handling. The problem here is similar to one encountered in most forms of statistics where data have been gathered into a frequency distribution. In order to obtain a readily comprehensible summary index of the distribution, and in order to be able to distinguish among different distributions, certain summary measures such as mean, variance, etc., have been developed. The problem in the present case is to compile a set of measures pertaining to the structure of the graph which will serve a similar purpose.

THE BETTI NUMBERS

In the study of the theory of linear graphs mathematicians have developed certain indices which are regarded as invariant, that is, their values are not changed by isomorphic transformations of the graph. Perhaps the easiest to comprehend of these are the $0^{\underline{th}}$ and $1^{\underline{st}}$ Betti numbers which were mentioned earlier in the discussion. The $0^{\underline{th}}$ Betti number, it will be recalled, was merely a count of the number of disconnected parts of the network. This summary index certainly represents a simple and basic statement about network structure.

The $1^{\underline{st}}$ Betti number, or the cyclomatic number as it is commonly known, presents a more sophisticated index pertaining to network structure. As pointed out earlier, if one cord is added to a tree, a graph is obtained with a unique single circuit known as a fundamental circuit. If this operation of cord addition is repeated for a graph with v vertices, e-v+1 circuits consisting of each cord and its unique tree path are obtained. In a more general fashion, if the graph is not connected, it consists of maximal connected subgraphs. A tree can be defined for each subgraph and a set of these trees is called a <u>forest</u> of G. It follows that there are v-p elements in the forest and e-v+p elements not in the forest. This index was earlier noted as μ, or the cyclomatic number, and it is essentially a count of the number of fundamental circuits existing in the system. To an extent, the cyclomatic number may be considered to be a measure of redundancy in the system. Since it was noted that a tree provides one and only one path between any pair of points, it can be seen that additional paths provided by circuits are redundant and that the total number of circuits present in the graph may be considered as a crude measure of the redundancy of the system. As may be seen from the structure of this index, any tree or disconnected graph has a cyclomatic number of 0, whereas as the graph moves closer and closer to the completely connected state, the cyclomatic number increases. (See Figure 12.) Applying this notion to the structure of transportation networks, it might be hypothesized that the magnitude of the cyclomatic number which characterizes a nation's transportation system would bear a direct relationship to the level of social and economic development of the nation.

THE ALPHA AND GAMMA INDICES

While the cyclomatic number does provide an index of network structure that is invariant under isomorphic transformations, it does not provide a readily intelligible measure of structure since it is bounded below by zero and bounded above only by some number which is a function of the number of nodes in the system. Ideally, it would be desirable to transform this index in such a manner that it will have common upper and lower bounds for all networks. Two

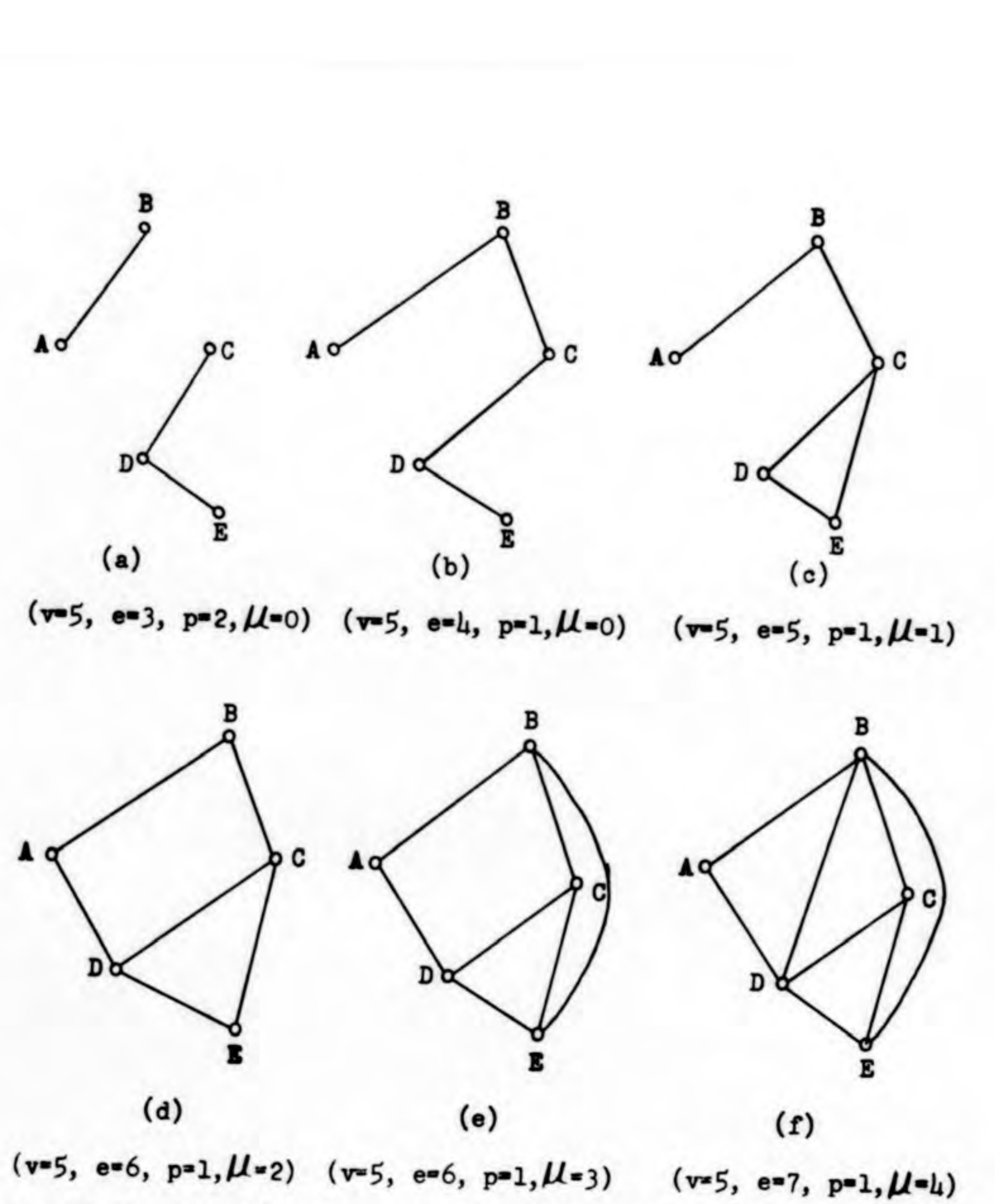


Figure 12. The Relation of the Cyclomatic Number to Network Structure.

additional measures are suggested which have this property, and which also appear to remain invariant under isomorphic transformations. They are known respectively as the alpha and gamma indices.

The gamma index for a planar network with e edges and v vertices is defined as $\gamma = e/(3(v-1))$. This is the ratio of the observed number of edges, e to the maximum number of edges in a planar graph. Obviously, this index would have a slightly different structure where non-planar graphs were under consideration (e.g., airline routes). In any network, the maximum number of direct connections is strictly a function of the number of nodes present. For a given network, as the number of edges in the system decreases, the gamma index will approach zero as a lower limit. On the other hand, as the number of edges in the system approaches the maximum, the gamma index will approach one as an upper limit. It appears to be most convenient to express this index as a percentage and it is therefore multiplied by 100, giving it a range from 0 to 100 and it is then interpreted as per cent connected.

The alpha index is somewhat similar, consisting of the ratio of the observed number of fundamental circuits to the maximum number of fundamental circuits which may exist in the system. The observed number of fundamental circuits is, of course, the cyclomatic number. Whereas the maximum number of fundamental circuits is equal to the number of edges present in a completely connected planar graph minus the number of edges contained in the complete tree. Therefore, for a planar graph, the alpha index will have the form $\alpha = \mu /(2v-5)$. This index is also multiplied by 100 to give it a range from 0 to 100 and an interpretation as per cent redundant.

The two indices then provide two percentage measures of the structure of a network. The alpha index may be interpreted as a per cent redundant, with a tree having zero redundancy and a maximally connected network having 100 per cent redundancy. The gamma index, on the other hand, may be interpreted as per cent connected with a completely unconnected system having a zero value and a completely connected system having a value of 100 per cent.

SUMMARY

This chapter has outlined some of the elementary notions contained in the theory of linear graphs and has indicated ways these notions may be applied to the construction of precise statements about the structural characteristics of transportation networks. The reader should realize that the mathematical material presented here lacks both rigor and completeness. The material as presented here is intended to serve only as the most cursory review of the material, and the interested reader is referred to the work by Seshu and Reed (Reference 20), or those of Konig (Reference 13) and Berge (Reference 4), for further amplification and greater rigor.

The indices which have been suggested in this chapter are far from exhaustive and, in the present form, are far from entirely adequate. Perhaps their major deficiency, arising out of the use of graph-theoretic models, is the lack of a precise statement pertaining to the angular structure of the network. The work by Beckmann (References 2,3) has shown how important this latter measure is in the analysis of transportation systems. However, this appears to be a measure which will be difficult to incorporate into the analysis since little existing mathematics appears to be relevant to the topic. The studies reported in the following chapter make use of the indices developed here and are felt to be extremely encouraging. However, it is planned that further work will be undertaken on indices incorporating the more complex notions of angular structure of the network.

CHAPTER 3. INTERNATION COMPARISONS OF TRANSPORTATION NETWORK STRUCTURES

The previous chapter discussed how the structure of transportation networks may be viewed from a rather simple mathematical point of view and presented a number of network measures. The measures represent succinct, though perhaps unfamiliar, ways to summarize the structure of transportation networks. The present chapter reports the use of regression studies of these measures to answer the question, "Can the structures of transportation systems be related to the features of the areas within which they are located?" In addition to the measures developed in a previous chapter, certain other measures developed for other aspects of the research were used in the regression studies. This chapter presents, in turn, the measurements used in the regression analyses, the computations made, and the results of the computations.

It was found that measures of network structure are related rather closely to the characteristics of the areas within which the networks lie, and this fact will be discussed in this chapter.

INPUTS: THE INDEPENDENT VARIABLES

The characteristics of areas containing transportation networks were represented by independent or explanatory variables. These variables fall into two categories -- (1) characteristics of areas that are functions of the level and nature of economic, social, and resource development and (2) characteristics of the physical make-up of areas. The former may be described as development variables and the latter as physical variables.

DEVELOPMENT VARIABLES

There is widespread knowledge that transportation development is closely correlated with the level of national development. In the preceding chapter of this monograph, indices were developed to measure the structure of transportation development. Before the relation between the level of national development and these transportation measures can be investigated, measures must be established for the notion "national development." Much work has been done on the notion of development and the problem of measurement has been solved elsewhere to a degree sufficient for the current needs of this research. The ensuing discussion takes advantage of a detailed statistical analysis of the development measurement question by Brian J. L. Berry (References 5,6).

Statistics are available for many nations treating such matters as the value of foreign trade, value of imports, development of energy resources, population density, and newspaper circulation. Berry found some 43 such measures for some 95 nations. Any one of these measures might serve as an index of development in an approximate fashion. The notion is tempting that if one measure is a useful one, two or more indices should be better. However, an inspection of the set of statistics will reveal quickly that various measures are redundant on each other. Berry's contribution was the combining of these statistics in a manner that established the basic factors that underlie variations of the statistics.

Berry's work revealed that four basic factors underlie variations in measures of degrees of development.* These factors were, namely: technological level, demographic level, income and external relations level, and size level. Nearly all of the variability that occurred in the 43 statistics could be attributed to the first two of the factors. The first of the four factors had by far the stronger relations and the second factor was next. The remaining two factors were relatively unimportant. Technological level takes account of various measures that may be made on the degree of urbanization, industrialization, transportation, trade, income, and the like. Demographic level reflects largely birth and death rates, population densities, population per unit of cultivated land, and similar measures. Berry has provided formulas for computing the technological and demographic levels of nations using weighted sums of the 43 statistics mentioned earlier.

A simple alternate method also has been provided for obtaining the technological and demographic scores. A study of the distribution of scores has revealed that they may be generated by answers to a series of simple questions. For example, by specifying the location of the nation, its climate, governmental status, presence or absence of subsistence or commercial sectors of the economy, and presence or absence of trade in raw materials, it is possible to use equations provided by Berry to estimate technological and demographic

*These results were obtained from a direct factor analysis of a table showing the ranks of 95 countries on the 43 available statistical measures. Nations were ranked from 1 to 95. The nation with a score of one would be that nation with the highest value. The nation with the score of 95 was that nation with the lowest value of the statistic.

levels.*

In summary, development may be measured by synthesizing available statistics. The study of these statistics has revealed that the development of a nation may be measured on two scales: a technological scale and a demographic scale. Also, it was noted that although these scales were derived from a rather complex set of operations on a large number of statistics, it is possible to assign values appropriate to individual areas using simple information and a relatively simple technique. Thus a relatively simple set of information will yield measures of development similar to measures that may be obtained from a large set of statistical indices.

PHYSICAL VARIABLES

The nature of the transportation network may depend upon the physical properties of the area it traverses. Three physical properties of areas were measured -- size, shape, and relief.

Information on the size of areas is available in a number of places. The data used here were taken from tables in the Encyclopedia Britannica and rounded by two decimal points (the tables in the Encyclopedia Britannica were derived from a variety of documents, all generally available). For example, Tunisia was reported as 48,332 square miles in size. This was recorded as 483 for purposes of this study.

Shape was measured on maps. The longest axis across each nation was determined by inspection and a perpendicular constructed across the nation at the midpoint of the longest axis. The measure of shape was obtained by dividing the airline distance along the longest axis by the airline distance along the perpendicular. It might be noted that this is not a completely satisfactory measure of shape. A nation shaped like a rectangle may have the same measure of shape as a nation shaped like an ellipse if the ratio is the same for the two areas. It might also be noted that this measure of shape is a pure number and, so far as the measure is concerned, is independent of size.

As is true of shape there is no entirely satisfactory measure of the relief of areas. The measure used here was constructed in an

*Berry obtained these results by a regression analysis on each factor. The factor scores were related to a set of independent variables. Each measure of an independent variable was in nominal form — the yes or no answer to a question such as, "Is the country tropical?" The regression model for technological scale had a R^2 of 86.5 and the regression model for demographic scale had R^2 of 81.9.

ad hoc manner and proved to be suitable to the study. Three lines were drawn at random across each area of study and the airline length of each line was measured. The distance along each route was also measured along the surface using profiles in the Times Atlas. These surface routes, of course, were greater than or equal to the airline distances. The airline route was taken to be 100 per cent and the surface route a percentage larger than 100 per cent. For each country, the "per cent larger" sums were added and divided by three, and the resulting value used to express the relief.

THE DATA

Transportation networks within twenty-five nations were selected for analysis and values of the five independent variables described above were computed for each nation. Table 1 presents the list of nations studied and values of the independent variables. Preliminary investigation of the sizes of areas and the characteristics of their transportation systems revealed existence of nonlinear relationships. Before computations were made, the raw data on size were converted into natural logarithms.

The size data in Table 1 and in ensuing tables are the natural logarithms of observed values.

Table 2 shows the means and standard deviations of the independent variables. The standard deviation is a measure of the variability of the observations about the mean. Approximately two-thirds of the observations lie within plus or minus one standard deviation of the mean.

Table 3 and 4 display the associations among the independent variables. Table 3 gives the correlation coefficients between variables taken two at a time. For example, the correlation between size (number 3 reading down the left hand column) and technological development is .35. This indicates that nations large in size tend to have higher values on the technological development scale. Note that the values on the technological development scale are high for the less-developed countries and low for the more developed countries. The correlation coefficient indicates that the larger countries tend to be less developed than small.

Table 4 shows coefficients of determination. These are the squares of the correlation coefficients and may be thought of as a percentage measure of association. The entry .12 relating to technological development and size indicates that 12 per cent of the variability in technological development is associated with variation in size (or vice versa, 12 per cent of the variation in size is associated with variation in technological development). Table 4 reveals that with the exception of the relationship between technological

TABLE 1

OBSERVED VALUES OF INDEPENDENT VARIABLES

		Technological* Development	Demographic* Level	Size	Shape	Relief
1.	Tunisia	351	32	4.683	2.677	5.51
2.	Ceylon	323	14	4.403	2.510	3.54
3.	Ghana	355	15	4.962	2.506	2.83
4.	Bolivia	370	18	5.627	2.135	20.00
5.	Iraq	344	25	5.234	2.376	2.21
6.	Nigeria	394	0	5.530	2.164	3.60
7.	Sudan	410	6	5.985	2.093	4.48
8.	Thailand	400	9	5.297	2.245	5.41
9.	France	125	38	5.327	1.468	12.04
10.	Mexico	222	19	5.881	2.422	20.96
11.	Yugoslavia	241	16	4.994	2.553	23.25
12.	Sweden	154	55	5.239	2.872	8.34
13.	Poland	182	25	5.080	2.155	3.15
14.	Czechoslovakia	159	38	4.693	2.706	13.93
15.	Hungary	221	29	4.555	2.521	5.94
16.	Bulgaria	279	47	4.631	2.438	12.97
17.	Finland	202	46	5.114	2.780	0.35
18.	Angola	438	28	5.682	1.809	1.42
19.	Algeria	323	26	5.963	1.230	1.52
20.	Cuba	256	37	4.645	2.979	12.78
21.	Rumania	258	23	4.962	2.135	25.01
22.	Malaya	256	17	4.705	2.159	19.51
23.	Iran	372	12	5.803	2.422	8.50
24.	Turkey	283	8	5.481	2.692	19.45
25.	Chile	239	24	5.456	3.100	66.80

*From Berry (Reference 5, p. 110, Table VIII-1, cols. 1 and 2, "Second Values"). Twenty was added to each Demographic Level entry.

TABLE 2

MEANS AND STANDARD DEVIATIONS, INDEPENDENT VARIABLES

	Variable	Mean	Standard Deviation
1.	Technological Development	286.280	87.849
2.	Demographic Level	24.280	13.731
3.	Size	5.197	.472
4.	Shape	2.367	.433
5.	Relief	12.140	13.738

TABLE 3

CORRELATION MATRIX, INDEPENDENT VARIABLES

	1. Technological Development	2. Demographic Level	3. Size	4. Shape	5. Relief
1.	1				
2.	-.61	1			
3.	.35	-.40	1		
4.	-.24	.21	-.43	1	
5.	-.27	-.04	.05	.36	1

TABLE 4

COEFFICIENTS OF DETERMINATION, INDEPENDENT VARIABLES

	1. Technological Development	2. Demographic Level	3. Size	4. Shape	5. Relief
1.	1				
2.	.38	1			
3.	.12	.13	1		
4.	.06	.04	.19	1	
5.	.07	.00	.00	.13	1

development and demographic level, relationships among the dependent variables are nil. This is highly desirable. In the analysis it is possible to measure association between any one of the independent variables and variation in transportation, without being forced to qualify the statement by the condition that other independent variables would vary in association with the variation in the independent variable at point. The association between technological development and demographic level is notably higher than the association among other of the independent variables and this point should be kept in mind when interpreting the results.

INPUTS: THE DEPENDENT VARIABLES

The rationale underlying this study is that transportation structure is dependent upon the characteristics of the area containing the network. The characteristics of areas have been summarized in terms of the independent variables just discussed. The dependent variables to be discussed now are those measures of transportation that are pertinent to the notion of transportation structure. These measures fall into two categories: (1) the measures discussed in Chapter 2 based on graph-theoretic considerations and (2) measures based on certain other work to be mentioned below.

GRAPH-THEORETIC MEASURES

Six measures of a graph-theoretic type were made on transportation networks in each of the 25 nations selected for study. These were:

1. The number of vertices, nodes, or places.
2. The number of edges, links, or routes. A variety of sources were used for the vertex and edge measurements. Sources included World Railways (Reference 30) and maps published in various issues of the journal Road International. Definitions of vertices and thus edges were partly topological and partly based on certain information contained on the maps. Any intersection of routes defined a vertex. Also, any place on the network deemed significant by the person who drafted the map was taken to be a vertex. End points were always treated as vertices.
3. Alpha index. This is the cyclomatic number (see item 5 below), divided by the maximum possible number of fundamental circuits, or $\mu/(2v-5)$.
4. Gamma index. This is the number of observed edges divided by the maximum possible number of edges in a planar graph with the observed number of vertices, or $e/3(v-2)$.
5. Cyclomatic number. This is the measure of the number of circuits in the transportation system, or the number of links in the system excess to the number required to tie the vertices together in a minimal way. $\mu = e-v+p$.
6. Diameter. As discussed in Chapter 2, this is a measure of the "span" of the transportation system. It is the minimum number of links that must be traversed in order to move between the two points that are the greatest distance apart on the network (as measured by the number of edges).

OTHER DEPENDENT VARIABLES

The above are six invariant characteristics of the structure of transportation networks. These were supplemented by two additional measures that were adopted after extensive empirical measurements of transportation networks. Much of the empirical work was in connection with studies which are to be reported at a later date and, consequently, this work will not be reviewed here. Two of the measures suggested by the empirical work were used in the present study, however, and they will be discussed briefly.

Measurements of the lengths of edges in miles proved practicable and preliminary correlations indicated that such measures were related to such independent variables as technological development. Consequently, measures were made in the nations under study of the average length of edges. Two measures were made, one for highway and one for railroads.

Also, a structure index was computed for both highway and rail networks in the nations under study. This structure index was computed by summing the number of endpoints in the system (weighted by two when endpoints were at an intersection) and dividing this number into the total length of the network. This, then, is a measure of length per weighted unit of structure. This measure was developed in connection with empirical work to be reported in a later monograph, as was mentioned before. However, preliminary work with the index indicated that it would be highly correlated with various independent variables and it was decided to use the index in this study.

THE DATA

On the basis of preliminary graphic analysis it was decided to transform many of the dependent variables to their natural logarithms. These transformations are listed below:

1. Vertices, transformed to the natural logarithm of the observed number.
2. Edges, transformed to the natural logarithm of the observed number.
3. Cyclomatic number, one was added to the observed value and the result transformed to its natural logarithm.
4. Average highway edge length, transformed to the natural logarithm of the natural logarithm of the observed value.
5. Average railroad edge length, transformed to the natural logarithm of the natural logarithm of the observed value.
6. Highway structure index, transformed to the natural logarithm of the observed value.
7. Railroad structure index, transformed to the natural logarithm of the observed value.

The presence of these transformations is not indicated in the tables in this chapter.

Table 5 presents the means and standard deviations of the dependent variables, and Tables 6 and 7 indicate the relationships between the dependent and independent variables taken two at a time. It may be noted, for example, that there is a correlation of -.86 between the natural logarithm of the number of vertices and the index of technological development. This indicates, as would be expected, that the more developed the area the greater the number of nodes or vertices

TABLE 5

SUMMARY, MEANS AND STANDARD DEVIATIONS, DEPENDENT VARIABLES

Variable	Mean	Standard Deviation
1. Number of Vertices	4.592	.857
2. Number of Edges	4.700	.953
3. Alpha Index	7.039	6.234
4. Gamma Index	38.366	4.048
5. Cyclomatic Number	2.322	1.630
6. Diameter	25.320	10.754
7. Average Edge Length (Highway)	.190	.039
8. Average Edge Length (Rail)	.225	.081
9. Structure Index (Highway)	.945	.172
10. Structure Index (Rail)	1.193	.342

TABLE 6

SUMMARY, SIMPLE COEFFICIENTS OF CORRELATION BETWEEN THE DEPENDENT AND INDEPENDENT VARIABLES

	Technological Development	Demographic Level	Size	Shape	Relief
1. Vertices	-.86	.53	-.15	.08	.30
2. Edges	-.85	.55	-.16	.05	.26
3. Alpha Index	-.64	.50	-.12	-.14	-.09
4. Gamma Index	-.61	.49	-.14	-.18	-.15
5. Cyclomatic Number	-.73	.56	-.13	-.04	.11
6. Diameter	-.79	.44	-.02	.24	.49
7. Average Edge Length (Highway)	.67	-.34	.64	-.27	-.10
8. Average Edge Length (Rail)	.66	-.62	.61	-.40	-.40
9. Structure Index (Highway)	.84	-.52	.58	-.29	-.11
10. Structure Index (Rail)	.73	-.59	.57	-.26	-.36

TABLE 7

SUMMARY, COEFFICIENTS OF DETERMINATION BETWEEN

THE DEPENDENT AND INDEPENDENT VARIABLES

		Technological Development	Demographic Level	Size	Shape	Relief
1.	Vertices	.73	.29	.02	.00	.09
2.	Edges	.73	.30	.02	.00	.07
3.	Alpha Index	.41	.25	.01	.02	.01
4.	Gamma Index	.37	.24	.02	.03	.02
5.	Cyclomatic Number	.54	.32	.02	.00	.01
6.	Diameter	.62	.20	.00	.06	.24
7.	Average Edge Length (Highway)	.45	.11	.41	.07	.01
8.	Average Edge Length (Rail)	.43	.39	.37	.16	.16
9.	Structure Index (Highway	.71	.27	.34	.08	.01
10.	Structure Index (Rail)	.54	.34	.33	.06	.13

on the transportation system. It also may be noted that the more developed the country, the shorter is the average edge length.

An inspection of Table 7 reveals associations measured in percentages. Some 73 per cent of the variation in number of vertices, for example, may be associated with the level of technological development. The relationship between the number of vertices and the demographic level is less strong. It may be noted that the six measures of structure derived from graph-theoretic points of view are almost completely independent of the size, shape and relief of areas, while those developed from empirical considerations generally have some relation with one or more of these physical variables.

THE COMPUTATIONS

Ten regression analyses were made. In each analysis the value of a dependent variable was assigned to the independent variables to the extent that variations in the data indicated that assignments were warranted. This section contains a discussion of the regression model used, the steps in computation, and the outputs from the regressions. It will be noted in this section that these computations were very good in the sense that much of the variability in individual dependent variables could be associated with the independent variables.

MODEL USED

The model used in this study was the standard linear equation used in regression studies, namely:

$$y = b_o + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + e$$

y represents the value of one of the dependent variables for a nation -- say, number of vertices. x_1 through x_5 represent the values of the independent variables for that nation. These were respectively: technological development, demographic level, size, shape, and relief. Values b_o through b_5 are the regression coefficients estimated from the data. The value of b_1, for instance, represents the contribution of a unit on the technological development scale to the value of a dependent variable. In the case of vertices, b_1 would be in units of number of vertices per unit of technological development. b_o is a constant term to which incremental values are added or subtracted in the calculation of the value of y. e is a measure of error. The b's were computed from a consideration of all of the data, they may be thought of as average relationships. When a computation is made for an individual country these average effects may not hold exactly. e is that amount necessary to correct y to its observed value. If the e's are small relative to the variability in the y's and if they are "well-behaved" -- e.g., normally distributed -- then the regressions may be though of as good.

STEPS

Each of the ten regressions was done in the following way. First, estimates were made of values of the regression coefficients using least-squares methods, the standard method for regression analysis. While the necessary matrix multiplications and inversions were accomplished for the data as a whole, the regressions differed somewhat from usual computations in that the regression coefficients were determined one at a time. In the first regression, for instance, the value of the coefficient b_1 was established and certain auxiliary computations made of correlation and error. Next, the value of regression coefficient b_2 was established and the estimates of error and correlation recomputed. These computations were continued in this incremental fashion until all five of the regression coefficients were determined. The preliminary four regressions of each of the dependent variables are of greatest interest in regards to the estimates of error to be discussed below.

OUTPUTS

The outputs from the calculations fall into two categories: (1) the regression coefficients, and (2) various estimates of error or reliability of the regressions. The regression coefficients are given in Table 8; these are numerical values of the b's in the linear

TABLE 8

SUMMARY, REGRESSION COEFFICIENTS OF THE TEN REGRESSIONS

		Techno-logical Develop-ment	Demo-graphic Level	Size	Shape	Relief	Constant Term
1.	Vertices	-.008	-.006	.201	-.267	-.007	6.30
2.	Edges	-.009	-.007	.213	-.335	-.006	6.74
3.	Alpha Index	-.049	.068	1.208	-3.306	-.090	22.123
4.	Gamma Index	-.031	.043	.412	-2.419	-.071	50.543
5.	Cyclomatic Number	-.012	.027	.354	-.799	.001	5.278
6.	Diameter	-.0918	.012	10.101	7.447	.347	-21.565
7.	Average Edge Length (Highway)	.0003	.0007	.0423	.0027	.0002	-.1424
8.	Average Edge Length (Rail)	.0001	-.0023	.0802	.0138	-.0025	-.1757
9.	Structure Index (Highway)	.0016	.0013	.1169	-.0062	.0014	-.1547
10.	Structure Index (Rail)	.0014	-.0064	.3366	.1738	-.0093	-1.0992

regression equation given above. For the first regression, for instance, we have:

$$y = 6.30 - .008x_1 - .006x_2 + .201x_3 - .267x_4 - .007x_5 + e$$

or:

$$\left\{\begin{array}{l}\text{number of}\\ \text{vertices}\\ \text{on network}\end{array}\right\} = 6.30 - .008\left\{\begin{array}{l}\text{index of}\\ \text{technological}\\ \text{development}\end{array}\right\} - .006\left\{\begin{array}{l}\text{index of}\\ \text{demographic}\\ \text{level}\end{array}\right\} + .201\left\{\begin{array}{l}\text{size}\\ \text{meas-}\\ \text{ure}\end{array}\right\}$$

$$- .267\left\{\begin{array}{l}\text{shape}\\ \text{measure}\end{array}\right\} - .007\left\{\begin{array}{l}\text{relief}\\ \text{measure}\end{array}\right\} + \begin{array}{l}\text{Error of}\\ \text{the estimate}\end{array}$$

and equations may be written for the nine remaining regressions. The regression coefficients are in the units described in the column and row headings on the table. -.008 is the number of vertices added per unit of technological development, for instance. The negative sign indicates that the higher the technological development score, the smaller is the number of vertices. This is to be expected since the technological development scale is inverted in the sense that the more developed countries have lower values on the scale.

A variety of measures were made of the reliability of the regressions. Tables 9 and 10 present measures of how well the regressions work.

TABLE 9

SUMMARY, MULTIPLE CORRELATION COEFFICIENTS OF THE TEN REGRESSIONS

	Techno-logical Develop-ment	Demo-graphic Level	Size	Shape	Relief
1. Vertices	.86	.86	.87	.87	.88
2. Edges	.85	.85	.87	.87	.88
3. Alpha Index	.64	.66	.68	.73	.75
4. Gamma Index	.61	.63	.64	.72	.75
5. Cyclomatic Number	.73	.75	.77	.79	.79
6. Diameter	.79	.79	.83	.87	.89
7. Average Edge Length (Highway)	.67	.68	.82	.82	.82
8. Average Edge Length (Rail)	.66	.71	.80	.80	.88
9. Structure Index (Highway)	.84	.84	.90	.90	.90
10. Structure Index (Rail)	.73	.75	.81	.81	.87

The multiple correlation coefficients shown in Table 9 may be interpreted as the correlation between the dependent variables and the independent variables taken one, two, three, four and five at a time. It may be noted that the correlation between the number of vertices and the level of technological development is .86. When the level of demographic development is also considered the correlation is the same. When the correlation between the number of vertices and all five of the independent variables is considered, the correlation coefficient rises to .88.

Table 10 contains the coefficients of determination, sometimes termed the power of the model. These are the squares of the multiple correlation coefficients, and they may be interpreted as the per cent of the variability in the dependent variable associated with variation in the independent variables. The first entry indicates that some 73 per cent of nation-to-nation variation in the number of vertices may

TABLE 10

SUMMARY OF COEFFICIENTS OF DETERMINATION OF THE TEN REGRESSIONS

		Techno-logical Develop-ment	Demo-graphic Level	Size	Shape	Relief
1.	Vertices	.73	.73	.76	.77	.77
2.	Edges	.73	.73	.75	.76	.77
3.	Alpha Index	.42	.44	.46	.54	.57
4.	Gamma Index	.37	.40	.41	.52	.56
5.	Cyclomatic Number	.54	.56	.59	.62	.62
6.	Diameter	.62	.62	.67	.75	.80
7.	Average Edge Length (Highway)	.45	.46	.66	.67	.67
8.	Average Edge Length (Rail)	.43	.51	.63	.65	.77
9.	Structure Index (Highway)	.71	.71	.80	.80	.81
10.	Structure Index (Rail)	.54	.57	.66	.66	.76

be associated with the level of technological development, and the right-hand entry of the first column indicates that some 77 per cent of the variability from nation-to-nation in number of vertices may be associated with the five independent variables taken altogether.

A number of other outputs from the regression bear on the "goodness" of the regressions. The twenty-two to twenty-five nations used in this study may be viewed as a sample from a larger set of nations, though it might be somewhat difficult to decide on the number of nations in the world. The United Nations' Statistical Yearbooks list approximately 260 political divisions, but about 100 of these are sub-divisions of larger units. Ginsburg (Reference 10) found it practicable to consider about 140 countries. Thus, the computations made here might be regarded as a sample from approximately 140 areas. Still a broader view might be adopted. The 140 areas might be regarded as displaying patterns from a larger universe of transportation network structures that might have developed given the conditions that control network development. This set of possible patterns might be very large or even unlimited in number. Reasoning this way, the sample might be regarded as one from an extremely large universe of possible transportation network structures.

When the regressions are viewed as a sample, the question of the reliability of the regression coefficients for the universe as a whole arises. Two measures bearing on this question were obtained from the regressions. Table 11 shows the estimated standard deviations of the regression coefficients. The regression coefficient

TABLE 11

SUMMARY, STANDARD DEVIATIONS OF THE REGRESSION COEFFICIENTS

		Techno-logical Develop-ment	Demo graphic Level	Size	Shape	Relief
1.	Vertices	.001	.009	.243	.267	.008
2.	Edges	.002	.016	.273	.300	.009
3.	Alpha Index	.015	.092	2.452	2.693	.083
4.	Gamma Index	.009	.060	1.610	1.769	.056
5.	Cyclomatic Number	.004	.022	.598	.657	.020
6.	Diameter	.019	.119	3.210	3.536	.169
7.	Average Edge Length (Highway)	.00008	.00051	.01365	.01500	.00046
8.	Average Edge Length (Rail)	.00014	.00087	.01231	.01238	.00078
9.	Structure Index (Highway)	.00028	.00169	.04476	.04918	.00152
10.	Structure Index (Rail)	.00062	.00380	.10080	.11080	.00342

between vertices and technological development, which is -.008, has the associated standard deviation of .001. Suppose there is a true regression coefficient for the universe. (-.008 is the best linear unbiased estimate of this true regression coefficient.) The number .001 may be interpreted as a description of the scatter of estimated regression coefficients around the true regression coefficient when samples of size 25 are taken and b's computed. Plus or minus two standard deviations would include some 95 per cent of the regression coefficients obtained in a large set of samples from the universe. Attaching the plus or minus two standard deviations to the -.008 allows the statement that one may be reasonably confident that the true regression coefficient lies between -.006 and -.010.

Table 12 presents the variance ratios for the regressions. The variance ratios permit significance tests of the increments in the variation explained by the regressions when individual regression coefficients are added. It may be seen that the reduction in the variance associated with technological development is extremely significant and that other reductions in variance are significant on occasion.

The regression equations may be viewed as forecasting devices. Observations on the independent variables may be entered into the

TABLE 12

VARIANCE RATIOS WITH 1 AND m-j DEGREES OF FREEDOM

Variable \ j equal	Technological Development 1	Demographic Level 2	Size 3	Shape 4	Relief 5	m No. of Observations
1. Vertices	63.31**	.01	2.32	.48	.76	25
2. Edges	61.27**	.07	2.23	.90	.39	25
3. Alpha Index	16.43**	.75	.83	3.53	1.18	25
4. Gamma Index	13.73**	.83	.45	4.44*	1.67	25
5. Cyclomatic Number	26.62**	1.03	1.51	1.84	.00	25
6. Diameter	32.33	.15	3.73	4.30*	4.21	22
7. Average Edge Length (Highway)	18.69**	.33	13.08**	.14	.14	25
8. Average Edge Length (Rail)	17.30**	3.45	7.38*	.61	10.26**	25
9. Structure Index (Highway)	55.00**	.01	10.25**	.09	.86	25
10. Structure Index (Rail)	26.52**	1.49	5.79*	.14	7.35*	25

**Significant at the 1% level.
*Significant at the 5% level.

equations and estimates of transportation structure derived. A change in some character of an area, such as level of technological development, may be postulated and new values of transportation structure estimated. Persons using regression equations of this sort for estimating purposes should proceed with caution, of course. For one thing, estimates that extend beyond the range of variability in the original data should be made only with great care. Too, the regressions limit themselves to measures of structure and do not display exactly how a change will take place on the map. (This point is investigated in Chapter 5 of this monograph.) In addition to the preceding points, the behavior of errors of estimation are critical where estimates are made. For this reason, a number of measures bearing on errors of estimation were obtained in the course of the regression calculations.

Table 13 presents the standard errors of estimate for the ten regressions. In discussing the regression equation earlier in this section, mention was made of the error term that must be introduced into the equation in order to correct the computed structure value to the actual observed value. The standard error of the estimate is a measure of the distribution of these error terms (their means are zeros). For the first regression, for example, about 66 per cent of the computed values from the regressions lie within plus or minus one

TABLE 13

STANDARD ERRORS OF ESTIMATE*

Variable	Standard Error of the Estimate
1. Vertices	.457
2. Edges	.514
3. Alpha Index	4.617
4. Gamma Index	3.033
5. Cyclomatic Number	1.127
6. Diameter	5.488
7. Average Edge Length (Highway)	.026
8. Average Edge Length (Rail)	.043
9. Structure Index (Highway)	.007
10. Structure Index (Rail)	.190

*Unbiased

standard error or plus or minus .457 of the observed values. 95 per cent lie within plus or minus two standard deviations, or plus or minus .914.

Because the behavior of these errors of estimate is critical to the use of the estimating equations, Tables 14 through 23 and Figure 13 have been provided showing the error of estimate for each nation and equation. Table 14, for instance, shows the observed number of vertices, the number estimated for each nation and the difference between the number observed and the number estimated by the equation. For Tunisia the estimated equation missed the number of vertices by .099, or, taking the antilog of .099, by about one vertex. Figure 13 shows for each regression how the errors were distributed. The plots are in standard error units.

RESULTS

Because the results of the regressions must be judged within the overall structure of the study, it might be wise to repeat certain comments made earlier on the structure of this study before going on to a summary of the results of the regressions. In the first paragraph of this chapter the question was asked, "Can the structures of transportation systems be related to the features of the areas within which they are located?" In terms of certain measures of structure and certain ways of measuring the "characteristics of the features of the areas," the answer to this question is a definite yes. But beyond this it is desired to answer this question in the affirmative so far as the actual networks or maps of the transportation systems of areas are concerned. That is, the ultimate answer to the question requires generating the

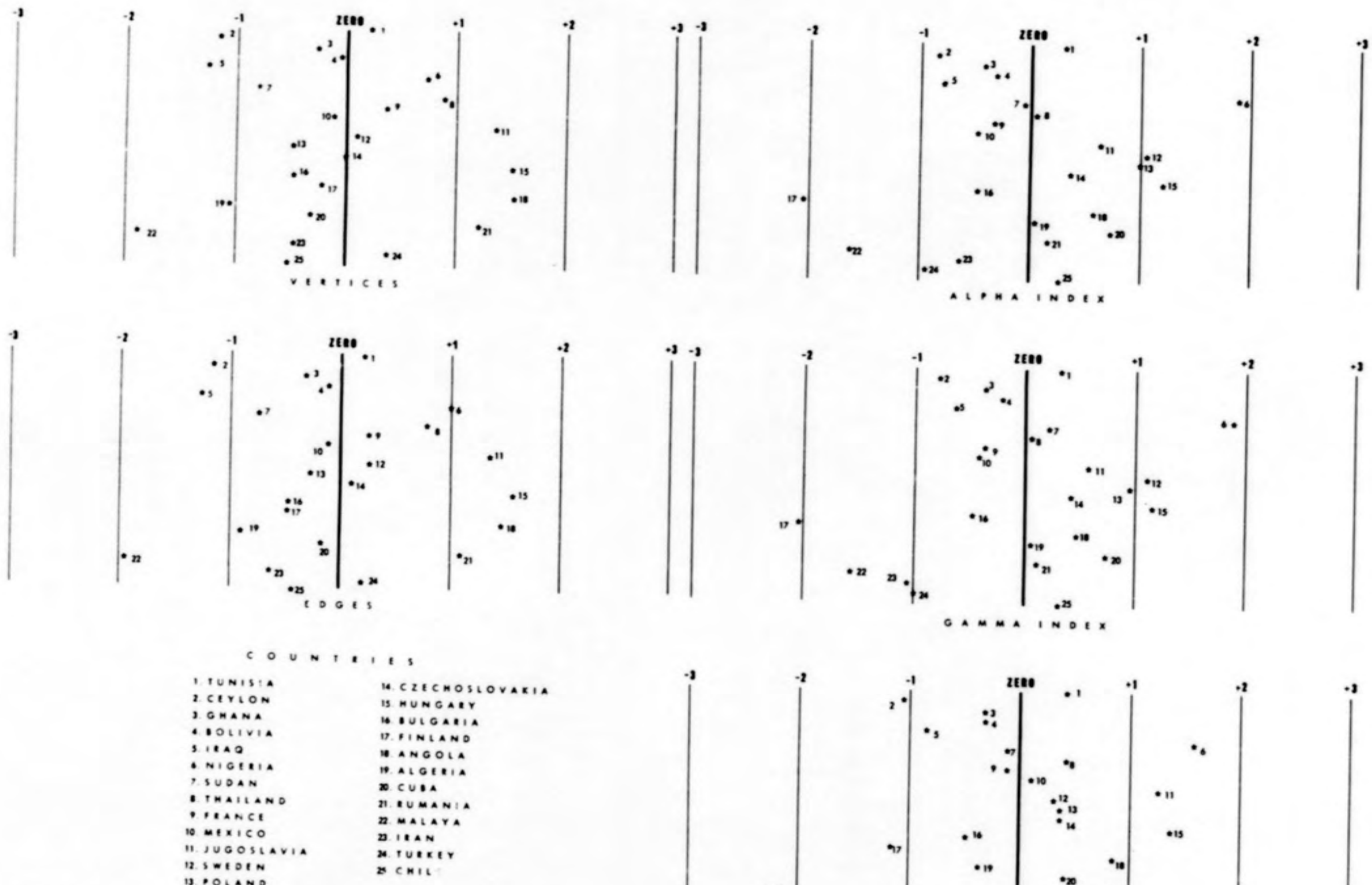


Figure 13. Deviation of observed point from calculated point plotted as ratio of residual to adjusted standard error.

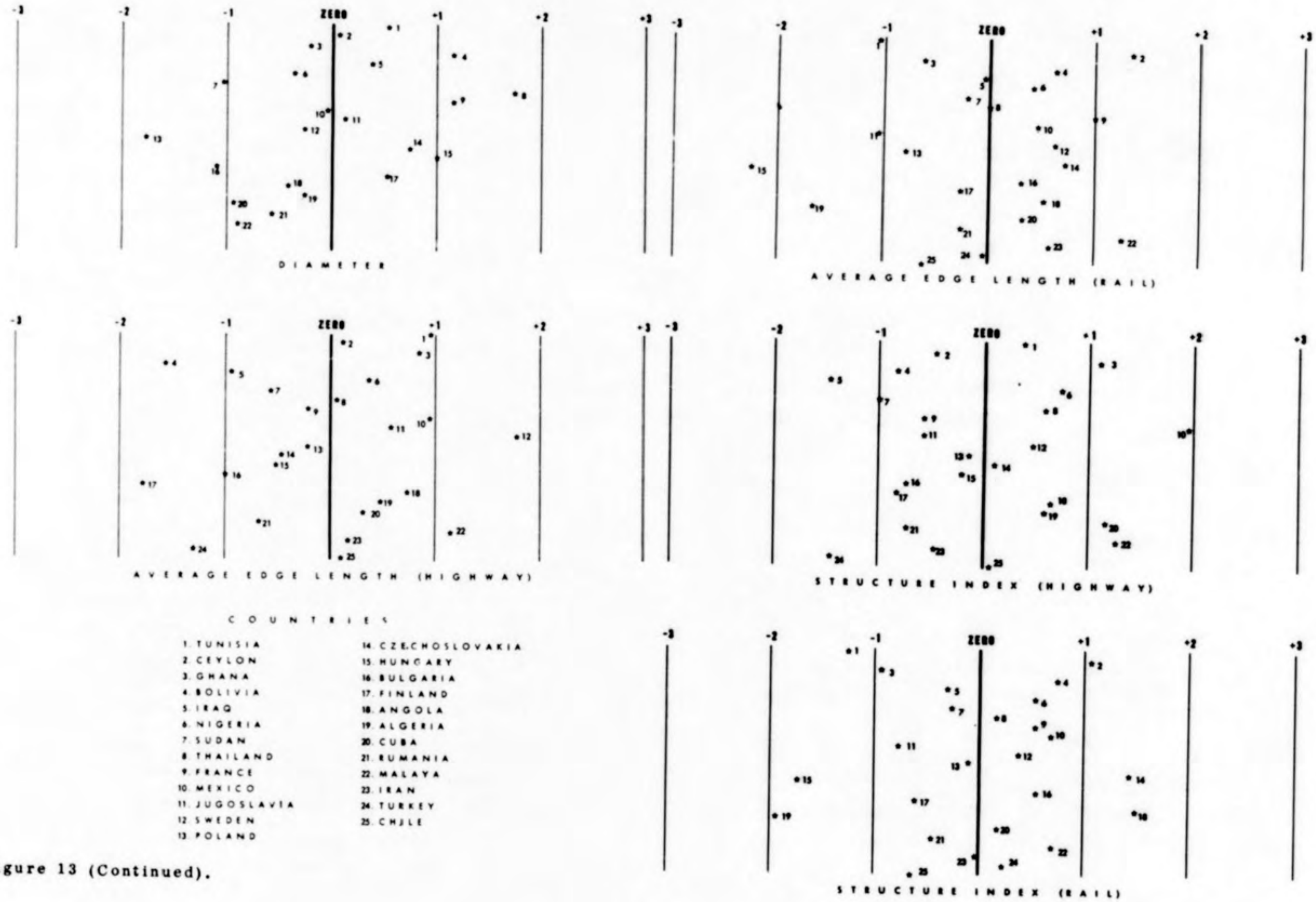


Figure 13 (Continued).

TABLE 14

OBSERVED NUMBER OF VERTICES AND NUMBER ESTIMATED BY THE FIRST REGRESSION

	Nation	Observed	Estimated	Residual (Observed minus Estimated)
1.	Tunisia	3.970	3.873	.099
2.	Ceylon	3.466	3.972	-.506
3.	Ghana	3.714	3.824	-.110
4.	Bolivia	4.060	4.074	-.014
5.	Iraq	3.496	4.057	-.561
6.	Nigeria	3.989	3.629	.360
7.	Sudan	3.296	3.649	-.353
8.	Thailand	3.989	3.576	.412
9.	France	6.433	6.258	.175
10.	Mexico	5.236	5.274	-.038
11.	Yugoslavia	5.553	4.904	.648
12.	Sweden	5.771	5.699	.071
13.	Poland	5.226	5.419	-.193
14.	Czechoslovakia	5.553	5.535	.018
15.	Hungary	5.645	4.939	.705
16.	Bulgaria	4.454	4.656	-.202
17.	Finland	5.118	5.196	-.079
18.	Angola	4.248	3.540	.708
19.	Algeria	4.220	4.682	-.462
20.	Cuba	4.511	4.644	-.133
21.	Rumania	5.493	4.923	.569
22.	Malaya	3.951	4.807	-.856
23.	Iran	3.689	3.899	-.210
24.	Turkey	4.727	4.547	.180
25.	Chile	5.004	5.225	-.221

TABLE 15

OBSERVED NUMBER OF EDGES AND NUMBER ESTIMATED BY THE SECOND REGRESSION

Nation	Observed	Estimated	Residual (Observed minus Estimated)
1.	4.043	3.915	.128
2.	3.434	4.023	-.590
3.	3.714	3.856	-.142
4.	4.078	4.107	-.030
5.	3.496	4.127	-.631
6.	4.159	3.632	.527
7.	3.296	3.656	-.360
8.	4.007	3.577	.430
9.	6.733	6.595	.137
10.	5.371	5.425	-.054
11.	5.727	5.010	.716
12.	6.094	5.944	.149
13.	5.529	5.647	-.119
14.	5.802	5.747	.055
15.	5.916	5.103	.812
16.	4.554	4.792	-.238
17.	5.170	5.400	-.230
18.	4.344	3.574	.769
19.	4.407	4.861	-.454
20.	4.682	4.749	-.067
21.	5.645	5.050	.594
22.	3.931	4.930	-.999
23.	3.611	3.920	-.309
24.	4.718	4.605	.1126
25.	5.050	5.256	-.206

TABLE 16

OBSERVED ALPHA INDEX AND ALPHA INDEX ESTIMATED BY THE THIRD REGRESSION

	Nation	Ob-served	Esti-mated	Residual (Observed minus Estimated)
1.	Tunisia	5.00	3.33	1.67
2.	Ceylon	0.00	3.87	-3.87
3.	Ghana	1.30	3.11	-1.81
4.	Bolivia	1.80	3.06	-1.26
5.	Iraq	1.63	5.15	-3.52
6.	Nigeria	10.70	1.91	8.79
7.	Sudan	2.04	2.24	-0.20
8.	Thailand	1.94	1.52	0.42
9.	France	17.67	19.05	-1.38
10.	Mexico	7.54	9.69	-2.15
11.	Yugoslavia	9.78	6.83	2.95
12.	Sweden	19.30	14.37	4.93
13.	Poland	18.25	13.59	4.66
14.	Czechoslovakia	14.14	12.35	1.79
15.	Hungary	15.68	9.85	5.83
16.	Bulgaria	6.00	7.96	-1.96
17.	Finland	3.00	12.27	-9.27
18.	Angola	5.90	3.22	2.68
19.	Algeria	11.40	10.98	0.42
20.	Cuba	10.20	6.65	3.55
21.	Rumania	8.50	7.66	0.84
22.	Malaya	0.00	7.46	-7.46
23.	Iran	0.00	2.85	-2.85
24.	Turkey	0.45	4.69	-4.24
25.	Chile	3.75	2.29	1.46

TABLE 17

OBSERVED GAMMA INDEX AND GAMMA INDEX ESTIMATED BY THE FOURTH REGRESSION

Na-tion	Ob-served	Esti-mated	Residual (Observed minus Estimated)
1.	37.30	36.21	1.09
2.	34.40	36.72	-2.32
3.	35.00	36.08	-1.08
4.	35.11	35.70	-0.59
5.	35.48	37.31	-1.83
6.	41.00	35.24	5.76
7.	36.00	35.30	0.70
8.	35.25	35.02	0.23
9.	45.16	46.13	-0.97
10.	38.50	39.63	-1.13
11.	39.90	38.07	1.83
12.	46.20	42.80	3.40
13.	45.60	42.69	2.91
14.	43.00	41.70	1.30
15.	44.00	40.37	3.63
16.	37.70	39.10	-1.40
17.	35.60	41.68	-6.08
18.	37.70	36.17	1.53
19.	41.40	41.12	0.28
20.	40.40	38.08	2.32
21.	39.10	38.73	0.37
22.	34.00	38.75	-4.75
23.	32.45	35.57	-3.12
24.	33.60	36.57	-2.97
25.	35.30	34.27	1.03

TABLE 18

THE OBSERVED CYCLOMATIC NUMBER AND THE CYCLOMATIC NUMBER ESTIMATED BY THE FIFTH REGRESSION

	Nation	Observed	Estimated	Residual (Observed minus Estimated)
1.	Tunisia	1.792	1.284	0.508
2.	Ceylon	0.000	1.174	-1.174
3.	Ghana	0.693	1.001	-0.308
4.	Bolivia	1.099	1.448	-0.349
5.	Iraq	0.693	1.612	-0.919
6.	Nigeria	2.398	0.578	1.820
7.	Sudan	0.693	0.761	-0.068
8.	Thailand	1.099	0.604	0.495
9.	France	5.389	5.479	-0.090
10.	Mexico	3.332	3.190	0.142
11.	Yugoslavia	3.912	2.454	1.458
12.	Sweden	4.812	4.424	0.388
13.	Poland	4.205	3.763	0.442
14.	Czechoslovakia	4.290	3.843	0.447
15.	Hungary	4.477	2.909	1.568
16.	Bulgaria	2.303	2.777	-0.474
17.	Finland	2.303	3.597	-1.294
18.	Angola	2.079	1.128	0.951
19.	Algeria	2.708	3.075	-0.367
20.	Cuba	2.890	2.364	0.526
21.	Rumania	3.714	2.758	0.956
22.	Malaya	0.000	2.502	-2.502
23.	Iran	0.000	1.078	-1.078
24.	Turkey	0.693	1.766	-1.073
25.	Chile	2.485	2.477	0.008

TABLE 19

OBSERVED DIAMETER AND DIAMETER ESTIMATED BY THE SIXTH REGRESSION

Nation	Observed	Estimated	Residual (Observed minus Estimated)
1.	19	16.0	3.0
2.	14	13.6	0.4
3.	15	16.0	-1.0
4.	31	24.6	6.4
5.	21	18.7	2.3
6.	14	15.8	-1.8
7.	13	18.8	-5.8
8.	24	14.2	9.8
9.	43	36.4	6.6
10.	43	43.2	-0.2
11.	35	34.2	0.8
12.	41	42.3	-1.3
13.	21	30.6	-9.6
14.	41	36.8	4.2
15.	31	25.5	5.5
16.	17	23.0	-6.0
17.	36	33.1	2.9
18.	8	10.2	-2.2
19.	18	19.2	-1.2
20.	24	29.1	-5.1
21.	27	29.9	-2.9
22.	21	25.7	-4.7

TABLE 20

THE OBSERVED AVERAGE EDGE LENGTH (HIGHWAY) AND THE AVERAGE EDGE LENGTH (HIGHWAY) ESTIMATED BY THE SEVENTH REGRESSION

Nation	Observed	Estimated	Residual (Observed minus Estimated)
1. Tunisia	0.218	0.193	0.025
2. Ceylon	0.162	0.159	0.003
3. Ghana	0.215	0.193	0.022
4. Bolivia	0.190	0.230	-0.040
5. Iraq	0.185	0.208	-0.023
6. Nigeria	0.227	0.217	0.010
7. Sudan	0.232	0.246	-0.014
8. Thailand	0.219	0.216	0.003
9. France	0.149	0.154	-0.005
10. Mexico	0.223	0.198	0.025
11. Yugoslavia	0.180	0.164	0.016
12. Sweden	0.222	0.175	0.047
13. Poland	0.148	0.152	-0.004
14. Czechoslovakia	0.130	0.141	-0.011
15. Hungary	0.133	0.146	-0.013
16. Bulgaria	0.155	0.181	-0.026
17. Finland	0.131	0.176	-0.045
18. Angola	0.275	0.256	0.019
19. Algeria	0.243	0.230	0.013
20. Cuba	0.178	0.169	0.009
21. Rumania	0.156	0.172	-0.016
22. Malaya	0.186	0.156	0.030
23. Iran	0.237	0.233	0.004
24. Turkey	0.159	0.192	-0.033
25. Chile	0.202	0.198	0.004

TABLE 21

THE OBSERVED AVERAGE EDGE LENGTH (RAILROAD) AND THE AVERAGE EDGE LENGTH (RAILROAD) ESTIMATED BY THE EIGHTH REGRESSION

Nation	Observed	Estimated	Residual (Observed minus Estimated)
1.	0.150	0.195	-0.045
2.	0.275	0.214	0.061
3.	0.237	0.262	-0.025
4.	0.291	0.263	0.028
5.	0.259	0.259	0.000
6.	0.362	0.342	0.020
7.	0.356	0.363	-0.007
8.	0.300	0.299	0.001
9.	0.214	0.170	0.044
10.	0.284	0.262	0.022
11.	0.151	0.197	-0.046
12.	0.185	0.155	0.030
13.	0.186	0.220	-0.034
14.	0.170	0.136	0.034
15.	0.074	0.171	-0.097
16.	0.139	0.125	0.014
17.	0.181	0.192	-0.011
18.	0.319	0.295	0.024
19.	0.226	0.298	-0.072
20.	0.168	0.154	0.014
21.	0.160	0.170	-0.010
22.	0.233	0.177	0.056
23.	0.351	0.324	0.027
24.	0.271	0.272	-0.001
25.	0.086	0.113	-0.027

TABLE 22

THE OBSERVED HIGHWAY STRUCTURE INDEX AND THE INDEX ESTIMATED BY THE NINTH REGRESSION

	Nation	Observed	Estimated	Residual (Observed minus Estimated)
1.	Tunisia	1.019	0.987	0.032
2.	Ceylon	0.846	0.885	-0.039
3.	Ghana	1.099	1.002	0.097
4.	Bolivia	1.066	1.134	-0.068
5.	Iraq	0.909	1.029	-0.120
6.	Nigeria	1.181	1.115	0.066
7.	Sudan	1.120	1.203	-0.083
8.	Thailand	1.162	1.111	0.051
9.	France	0.676	0.724	-0.048
10.	Mexico	1.092	0.928	0.164
11.	Yugoslavia	0.807	0.853	-0.046
12.	Sweden	0.812	0.768	0.044
13.	Poland	0.741	0.754	-0.013
14.	Czechoslovakia	0.711	0.700	0.011
15.	Hungary	0.745	0.762	-0.017
16.	Bulgaria	0.835	0.897	-0.062
17.	Finland	0.741	0.809	-0.068
18.	Angola	1.293	1.238	0.055
19.	Algeria	1.141	1.087	0.054
20.	Cuba	0.946	0.845	0.101
21.	Rumania	0.830	0.890	-0.060
22.	Malaya	0.949	0.842	0.107
23.	Iran	1.093	1.132	-0.039
24.	Turkey	0.838	0.961	-0.123
25.	Chile	0.980	0.972	0.008

TABLE 23

THE OBSERVED RAILROAD STRUCTURE INDEX AND THE INDEX ESTIMATED BY THE TENTH REGRESSION

Nation	Observed	Estimated	Residual (Observed minus Estimated
1.	0.941	1.176	-0.235
2.	1.352	1.148	0.204
3.	1.209	1.380	-0.171
4.	1.531	1.382	0.149
5.	1.322	1.375	-0.053
6.	1.766	1.656	0.110
7.	1.730	1.772	-0.042
8.	1.560	1.525	0.035
9.	0.875	0.767	0.108
10.	1.429	1.295	0.134
11.	0.900	1.044	-0.144
12.	1.020	0.947	0.073
13.	1.034	1.049	-0.015
14.	1.078	0.799	0.279
15.	0.611	0.939	-0.328
16.	0.955	0.851	0.104
17.	0.970	1.088	-0.118
18.	1.828	1.547	0.281
19.	1.031	1.392	-0.361
20.	1.019	0.983	0.036
21.	0.837	0.923	-0.086
22.	1.060	0.928	0.132
23.	1.634	1.639	-0.005
24.	1.420	1.377	0.043
25.	0.713	0.837	-0.124

actual transportation network, given the characteristics of the area that contains the network. The first step in obtaining the ultimate answer to this question is that of generating the characteristics of network structure, so our ability to answer the question with a yes in the case of the characteristics of the network is an essential part of the objective of the over-all objective of the research.

SIMPLICITY AND PRECISENESS

Simplicity and preciseness may always be taken to be desirable attributes of models. It should be stressed that the regression equations used in this study have these attributes. Measurement requirements for calculating values of the dependent variables are not at all demanding. Computation of size, shape, and relief measures is simple and straightforward. While development of the technological development and demographic level scales required a great deal of statistical work, values on these scales may be estimated from very simple information. The values of the independent variables are also very simple to measure, although they too depend upon certain mathematical considerations.

As was mentioned before, the fit of the regression model to the data may be regarded as quite good. This is a relative matter, of course. Experience with studies of this type reveals that cross-section models rarely display such strong relationships as those reported here. Also, it was noted that the residuals or errors were such that introduction of new data into the regressions and limited projections may be made with confidence.

RELATIONSHIPS

It was remarked earlier that these regressions are but one stage of an effort to reproduce actual transportation networks from data on the characteristics of areas and that the question of the over-all pertinence of the regressions will be left open until that portion of the research is discussed. However, it is possible to make certain summary remarks on the relationships within the regressions and those remarks will be made here. Tables given earlier in this discussion give the units within which the data were measured, the regression coefficients, and the variance ratios.

The first point to be made is that the relative importance of regression coefficients cannot be determined by comparing their magnitudes. The regression coefficient for the contribution of a unit of technological development to the number of vertices, for instance, was a -.008, while the contribution of a unit of size as to the number of vertices was .201. However, examination of the observed values of the independent variables (Table 1) reveals differences in the magnitudes and dispersions of the technological development and size variables, and this variability in measurement units accounts for part of the variation in the size of regression coefficients. The regression coefficients differ from regression to regression. The regression coefficients associated with technological development range from -.0918 to .0016, for instance. These regression-to-regression size differences also are related to differences in the sizes of units used. Table 5 showing the means of the dependent variables indicates some of the differences in the magnitudes involved, as do the various tables showing the observed values of the dependent variables, their estimated values and the residuals (Tables 14 through 23). Because of variations in units used, the regression coefficients do not lend themselves readily to comparative statements.

Notwithstanding the remarks just made, the regression coefficients are subject to certain general interpretations. The table showing variance ratios (Table 12) and Table 10 showing coefficients of determination permit identification of those regression coefficients associated with the stronger relations between independent and dependent variables. It may be seen from these tables that technological development is always a major determinate of structure, with other of the independent variables important only in cases. Referring to the technological development column of regression coefficients in the summary of regression coefficients (Table 8), it may be seen that for the first six regressions the more developed the country (and smaller the measure of technological development) the higher the value of the dependent variab'e. Just the reverse is true for the last four regressions. The more highly developed the country the smaller the average edge length and the smaller the structure index. The demographic measure is of little importance and size is of importance only in the case of the last four measures. The greater the size, the greater the average edge length and structure index, this being more true for highways than for railroads. The shape measure is significant in relation to diameter and the gamma index, while relief has its greatest effect on the edge length and structure indices for railroads.

The above statements are a mixture of observations based on the magnitudes of the regression coefficients and their relative significance in the variance ratio tests. Additional findings may be made by writing out the individual regression equations and studying their sensitivity to variations in the independent variables.

SUMMARY

This chapter reported the measurements used, computations, and results from ten regression studies. These regression studies established relationships between certain measurements of transportation network structure and measurements of the characteristics of the areas within which these networks lie. It was found that measures of network structure could be related rather closely to the characteristics of the areas containing the networks and that the technological development was the more important factor conditioning the character of transportation systems.

CHAPTER 4. EMPIRICAL STUDIES OF NETWORK STRUCTURE WITHIN NATIONS

The preceding chapter reviewed empirical studies of nation-to-nation variations in the structure or layout of transportation systems and their relation to varying environmental conditions. The present chapter reports on two studies of the structural details of transportation systems within nations.

As was noted in Chapter 2, the transportation system may be represented by its connection matrix. Connection matrices were constructed for the surface transportation networks of all the study nations mentioned in Chapter 3, as well as for a number of others. Connection matrices for local air transport services within a number of Central and Latin American nations were also constructed to see if any basic differences existed in the patterns developed by service and air transport systems. An examination of the connection matrices for the various surface transportation networks revealed one common factor: a heavy concentration of entries near the main diagonal of the matrix. That is, one dominant characteristic of the systems was that the routes that made up the system tended to link urban centers mainly to their nearby neighbors. Aside from this "neighborhood effect," the connection matrices also revealed certain minor groupings of off-diagonal elements. This tendency toward off-diagonal groupings was accentuated when connection matrices of the local service airlines routes were examined. Here some organizing effect, over and above the aforementioned neighborhood effect, appears to exist with certain nodes or groups of nodes exerting an organizing effect on the pattern of routes.

Apart from these very obvious and general conclusions, it is quite difficult to make any further statements about underlying structural patterns from visual examination of the connection matrices. Analysis in greater depth must utilize more explicit and powerful tools. The most appropriate appears to be a tool of statistical analysis known as component analysis, which has been used quite successfully by others in cases of a similar nature. (See MacRae (Reference 16).)

COMPONENT ANALYSIS

The following discussion is intended only to outline the basic operations undertaken in the performance of a component analysis. For a more detailed and rigorous discussion the reader is referred to the recent work by Harman (Reference 11). The regression studies described in the previous chapter were studies of dependence in that they examined the relationship between a given variable and a set of so-called independent variables in order to determine the precise characteristics of the linear relationship. Component analysis, on the other hand, is a study of interdependence; that is, a study of the relationships among a series of variables without specification to which is "dependent" and which are "independent." The data are examined to determine the basic dimensions of variation and once these basic dimensions are determined, the analyst seeks to describe them from both theoretical and empirical standpoints.

Component analysis begins with a given data matrix X, of dimensions n x m. The data matrix consists of n observations on each of m variables, and within the matrix the cell entry, x_{ij}, is the score or value associated with observation i on variable j. The basic interrelationships between the m variables are represented by the entries in a matrix of simple correlations, R. The basic object of the component analysis is to create an n x r matrix, V, which will have n observations and r new orthogonal variables, each one of which will express an independent pattern of variation or intercorrelation as contained in the matrix R. Each of the column vectors of the matrix V contains the scores of the n observations on one of the new dimensions. These new dimensions differ from the original variables in that they are orthogonal, that is, not mutually intercorrelated. In a certain sense, they represent "adjustments" of the old variables to eliminate the intercorrelations as expressed in the matrix R.

Now V is equivalent to the product of two matrices X, noted previously, and L; where L is the matrix of "factor-loadings" which expresses the relationship of each of the m variables to each of the new r dimensions. The solution to the problem of finding the new dimensions is achieved by processing the matrix R until $R = L\,L^{T}+S+E$ where S represents a specific variance term and E represents an error variance term. In practice, it is not normally possible to distinguish analytically between S and E; hence R is presumed to equal $L\,L^{T}+Z$, where $Z = S+E$.

In general the pattern of analysis is as follows: The variation accounted for by the first new dimension is estimated and then subtracted from the original correlation matrix R. Using the residual correlation matrix, R_1, as a basis, the analysis is then repeated to estimate the variation associated with the second new dimension. This is then subtracted from the first residual correlation matrix giving a new residual matrix, R_2, which forms the basis of analysis for the third new dimension. This processing is repeated again and again with each succeeding factor after first normally extracting less and less of the observed variation. The process of analysis continues until either (a) the analyst feels it would be unprofitable (in terms of the explanatory content of the new variables) to continue further, or (b) all of the observed variation has been explained in terms of a finite number of new dimensions where $r < m$.

THE STRUCTURE OF LOCAL SERVICE AIRLINE ROUTES

Figures 1 and 2 presented in Chapter 1 displayed the route structure of local service airlines in Guatemala and Honduras. It was pointed out in Chapter 1 that these systems appeared to be basically different since the system in Guatemala seemed to focus upon Guatemala City, whereas the Honduran network appeared to have no such over-all focal point. Table 24 displays various indices pertaining to network

TABLE 24

GUATEMALA AND HONDURAS
INDICES OF LOCAL SERVICE AIRLINE ROUTE STRUCTURE

	Guatemala	Honduras
Number of nodes	20	32
Number of routes	20	45
Mean number of connections/node	2.0	2.8
Cyclomatic number	1	14
Gamma index*	10.5	9.1
Alpha index*	0.6	3.0

*Nonplanar basis.

structure for these two nations. It seems that while the Honduran network is larger in size and displays a higher average number of connections per node, it is less directly connected than the Guatemalan network, but on the other hand displays a higher percentage of redundant routes. These indices certainly provide information about the basic network structure in the two countries; however, they do not provide information on the visual differences in structure which were remarked upon earlier.

Pursuing this line of analysis further, two nations were selected, Argentina and Venezuela, and component analyses were applied to the connection matrices developed from maps of their local service airline networks. Indices for these two networks are displayed in Table 25 and indicate that the networks might be generally considered to be similar. The actual networks in these two cases, however, are larger and considerably more complex than those displayed in Figures 1 and 2, so that it was impossible to develop any well-defined notions about structure from visual examination of the network maps. The results of the component analyses are presented in the following sections.

TABLE 25

ARGENTINA AND VENEZUELA
INDICES OF LOCAL SERVICE AIRLINE ROUTE STRUCTURE

	Argentina	Venezuela
Number of nodes	50	59
Number of routes	91	104
Mean number of connections/node	3.8	3.5
Cyclomatic number	42	46
Gamma index*	7.4	6.1
Alpha index*	3.6	2.8

*Nonplanar basis.

THE ARGENTINE CASE

Argentina is a nation which lies about two-thirds of the way along Berry's scales, discussed in Chapter 3, of economic-demographic development. Its system of local service airlines serves 50 cities via 91 routes. The nation itself is over 2200 statute miles from north to south and nearly 800 statute miles from east to west at its widest point.

Civil aviation at present is somewhat less important in Argentina than in a number of other Latin American countries whose surface transportation facilities are less well-developed than those of Argentina. In comparison with the rest of Latin America, Argentina has a relatively good system of railways, highways, and coastal and inland waterways. As of 1957, there were seven airlines operating domestic air route services in Argentina (see the report by the U.S. Department of Commerce (Reference 23)). In addition there were 34 other companies engaged in non-scheduled and related services. The routes analyzed in the present study are those of the regularly scheduled passenger carriers. The 1957 report on civil aviation in Argentina lists 235 airports, only a few of which were of an improved nature. Outside of Buenos Aires, only the airports at Mendoza and San Carlos de Bariloche are reported to have concrete runways while Clorinda, Cordova, Resistencia, Rio Gallegos, and Salta were reported to have asphalt runways.

THE ROUTE STRUCTURE

The basic pattern of local service air routes, as of late 1959, is displayed in Figure 14. (The figure is based on a map published by Official Airline Guide.) The connection matrix which is equivalent to this route map is displayed as Table 26. (The zeros are not shown in this matrix and the ones are shown as "x's.") An examination of this table reveals a strong tendency toward the neighborhood effect mentioned earlier, together with a moderate development of clusters of entries which are located off the main diagonal. As of 1959-60, the network was slightly over seven per cent connected and less than four per cent redundant.

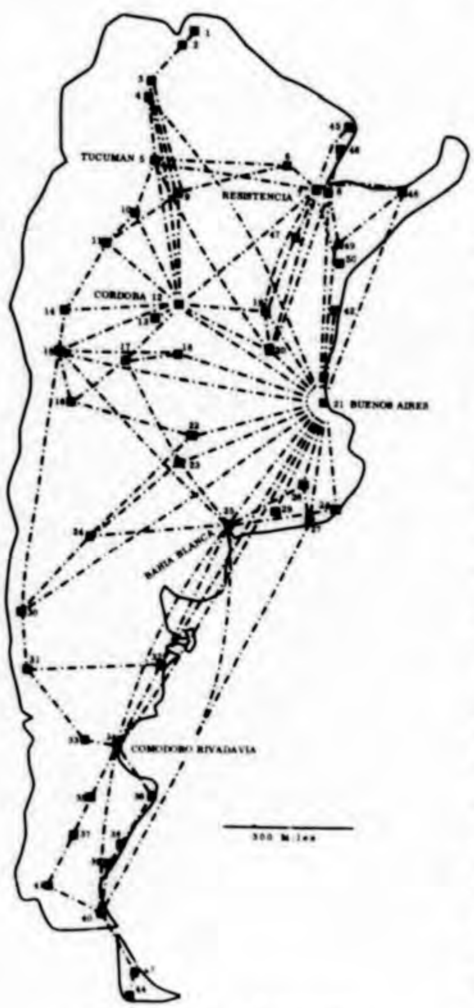


Figure 14. Argentina: Local Service Airline Routes.

THE COMPONENTS ANALYSIS

A component analysis was undertaken on the connection matrix described in Table 26 using the component analysis program (FA 108) of the University of Chicago's Computing Center. The results of the analysis are presented in Table 27.

The analysis produced a set of five basic patterns of variation which were extracted from the data expressed in the connection matrix. The five basic patterns explained over 52 per cent of the total observed variation, or structure of the network. (See Table 28). No one of these basic patterns contains the neighborhood effect alluded to several times previously in the discussion. It is apparent that the analysis concentrated on the off-diagonal clusters leaving the neighborhood effect as part of the unexplained variation. The fact that these five basic patterns accounted for over 52 per cent of the observed variation is indeed encouraging since this is over and above that explained by the neighborhood effect.

TABLE 26

CONNECTION MATRIX FOR LOCAL SERVICE AIRLINE ROUTES OF ARGENTINA

City	Connections																																																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
1. Tartagal		x																																																
2. Orán	x		x																																															
3. Jujuy		x		x					x																																									
4. Salta			x		x				x			x									x																													
5. Tucumán				x		x	x			x		x							x																															
6. Pres. Rogue Sáens Pena					x		x		x																																									
7. Resistencia					x	x						x							x	x																										x		x	x		x
8. Corrientes						x														x	x																										x			x	
9. Santiago del Estero			x	x							x	x																																						
10. Catamarca					x						x	x																																						
11. La Rioja									x	x		x		x																																				
12. Cordoba				x	x		x		x	x	x		x	x	x		x		x	x	x																													
13. Villa Dolores												x																																						
14. San Juan											x	x			x																x																			
15. Mendoza												x		x		x		x			x				x																									
16. San Rafael															x		x					x																												
17. San Luis												x				x		x					x																											
18. Rio Cuarto												x			x		x				x																													
19. Santa Fe							x													x	x																										x			
20. Rosario				x			x	x				x							x		x																													
21. BUENOS AIRES				x				x				x			x			x	x	x		x	x		x	x	x	x	x	x		x		x								x						x	x	
22. General Pico																x					x			x																									
23. Santa Rosa																	x				x			x	x																								
24. Neuquén																						x	x		x					x																				
25. Bahia Blanca															x						x		x	x		x		x				x		x						x										
26. Tandil																					x				x		x																							
27. Necochea																					x					x														x										
28. Mar del Plata																					x				x																									
29. Trés Arryos																					x																													
30. San Carlos de Bariloche															x						x			x							x																			
31. Esquel																														x		x	x																	
32. Trelew																					x				x						x			x																
33. Colonia Sarmiento																															x			x																
34. Comodoro Rivadavia																					x				x							x	x		x	x				x										
35. Periot Moreno																																		x			x													
36. Puerto Deseado																																		x				x												
37. Governador Gregores																																			x						x									
38. San Julian																																				x			x											
39. Santa Crus																																						x		x										
40. Rio Gallegos																									x		x							x					x		x		x							
41. Lago Argentina																																					x			x										
42. Concordia																					x																													
43. Rio Grande																																								x				x						
44. Ushuaia																																											x							
45. Clorinda							x																																											
46. Formosa								x																																										
47. Reconquista							x												x																															
48. Rosadas							x														x																												x	
49. Mercedes								x													x																											x		
50. Curusa Cuatia							x																																											

TABLE 27

VALUES FOR 50 ARGENTINE CITIES ON FIVE BASIC PATTERNS

City	Basic Pattern Values				
	1	2	3	4	5
1. Tartagal	0.007	0.018	0.005	-0.029	-0.010
2. Oran	0.049	0.087	-0.012	0.141	-0.115
3. Jujuy	0.311	0.361	0.095	-0.303	-0.121
4. Salta	1.192	0.720	-0.626	0.660	-0.335
5. Tucuman	1.162	1.284	0.440	-0.020	0.453
6. Pres. Rogue Saenz Pena	0.503	0.805	-0.282	-0.328	0.378
7. Resistencia	1.318	1.318	0.743	0.479	0.682
8. Corrientes	0.860	-0.077	-0.496	-0.094	0.342
9. Santiago del Estero	0.775	0.890	0.461	0.534	-0.461
10. Catamarca	0.646	0.793	0.150	0.481	-0.361
11. La Rioja	0.687	0.804	0.072	0.325	-0.455
12. Cordoba	2.334	1.541	-0.711	-1.604	-0.320
13. Villa Dolores	0.360	0.323	0.207	0.521	-0.199
14. San Juan	0.700	0.419	0.255	0.076	-0.822
15. Mendoza	1.524	-0.325	-0.190	1.511	-0.572
16. San Rafael	0.442	-0.159	0.104	-0.745	-0.715
17. San Luis	0.698	0.056	0.555	0.944	-0.692
18. Rio Cuarto	0.826	-0.241	-0.763	-0.541	-0.625
19. Santa Fe	1.353	0.715	-0.470	0.365	0.548
20. Rosario	1.569	0.848	-0.431	0.233	0.505
21. BUENOS AIRES	3.148	-1.219	2.752	-0.457	0.067
22. General Pico	0.647	-0.357	-0.862	0.296	-0.410
23. Santa Rosa	0.931	-0.635	-0.962	0.076	-0.230
24. Neuquen	0.619	-0.731	0.581	0.318	-0.111
25. Bahia Blanca	1.587	-1.511	-0.171	-0.879	-0.519
26. Tandil	0.836	-0.649	-0.540	0.419	0.351
27. Necochea	0.697	-0.521	-0.674	-0.076	0.052
28. Mar del Plata	0.728	-0.516	-0.681	0.404	0.191
29. Tres Arryos	0.484	-0.185	-0.677	0.118	-0.003
30. San Carlos de Bariloche	0.861	-0.490	-0.869	-0.383	-0.622
31. Esquel	0.308	-0.407	0.276	0.092	0.067
32. Trelew	0.935	-0.858	-0.604	0.274	0.377
33. Colonia Sarmiento	0.207	-0.341	0.077	-0.129	0.185
34. Comodoro Rivadavia	1.040	-1.085	-0.786	0.389	0.654
35. Perito Moreno	0.166	-0.284	0.156	-0.127	0.243
36. Puertc Deseado	0.166	-0.284	0.156	-0.127	0.243
37. Governador Gregores	0.039	-0.103	-0.026	0.073	0.262
38. San Julian	0.039	-0.103	-0.026	0.073	0.262
39. Santa Cruz	0.090	-0.202	-0.101	-0.083	0.148

Continued

TABLE 27 -- Continued

City		Basic Pattern Values				
		1	2	3	4	5
40.	Rio Gallegos	0.552	-0.862	0.346	0.277	0.973
41.	Lago Argentina	0.090	-0.202	-0.101	-0.083	0.148
42.	Concordia	0.484	-0.185	-0.677	0.118	-0.003
43.	Rio Grande	0.086	-0.187	-0.110	-0.070	0.113
44.	Ushuaia	0.013	-0.043	0.020	0.025	0.131
45.	Clorinda	0.203	0.306	-0.136	-0.220	0.409
46.	Formosa	0.132	-0.030	0.111	-0.019	0.077
47.	Reconquista	0.412	0.447	-0.007	-0.384	0.574
48.	Rosadas	0.801	0.060	-0.729	-0.175	0.507
49.	Mercedes	0.740	-0.222	-0.395	0.103	0.165
50.	Curuza Cuatia	0.203	0.306	-0.136	-0.203	0.409

TABLE 28

ARGENTINA
PROPORTION OF TOTAL VARIATION EXPLAINED
BY EACH BASIC PATTERN

Total variance	192.00	Per Cent 100.00
Explained by:		
pattern one	42.14	21.94
pattern two	20.65	10.76
pattern three	18.19	9.47
pattern four	10.96	5.71
pattern five	8.61	4.48
Total	100.55	52.36

The first factor or basic pattern scaled the nodes in terms of their total number of direct connections. That is, the first factor was a nearly linear scaling of the nodes by size. This accounted for nearly 22 per cent of the observed variation.

The second factor began the extraction of the basic structure of the network by breaking out two major regional systems. The first was a triangular grouping of directly connected cities to the northwest of Buenos Aires. The basic cities or nodal points in this major air transport region were Cordoba, Resistencia, and Tucman. The second major region was a nearly linear one lying to the south of Buenos Aires and centering around the urban centers of Bahia Blanca, Buenos

Aires and Commodoro Rivadavia. This regional structure accounted for nearly 11 per cent of the observed variation. It should be pointed out, of course, that factor two represents a major regionalization effect <u>after</u> the interrelationships have been adjusted for the size effect extracted by factor one. Because of this element, it is difficult to visually depict the exact pattern displayed by factor two. Figure 15 represents an attempt to illustrate the general notion involved rather than to display accurately all elements of the major regionalization effect extracted by the component analysis. <u>This figure, as well as the other figures dealing with the results of the component analysis, should be viewed as merely suggestive.</u>

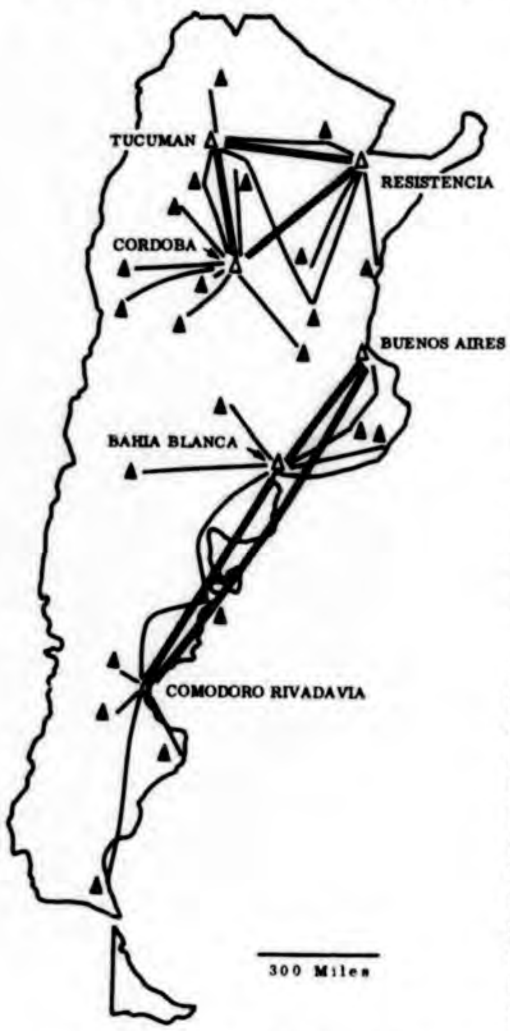


Figure 15. Argentina: Major Regionalization Effect

Pattern three, which represents an attempt to explain the residual variation left after patterns one and two have been extracted, accounted for over 9 per cent of the total variation and can best be interpreted as a field effect centering on Buenos Aires. It should be noted that the regionalization effect in Argentina appears to be slightly stronger than the field effect generated by Buenos Aires. Figure 16 represents an attempt to suggest the basic notion of the field effect extracted by factor three.

Factor four extracted less than 6 per cent of the observed variation and appeared to define a minor regionalization centered around Cordoba and Mendoza and linked weakly to Buenos Aires. This region appears to be a generally triangular-shaped area acting as a feeder region for the major northern region. The fifth factor extracted just under 5 per cent of the observed variation and appeared to define still another minor region centered at Rio Gallegos and weakly linked to Buenos Aires, as well as to the second major region lying to the south of that city. The structure of these minor regions is suggested in Figure 17.

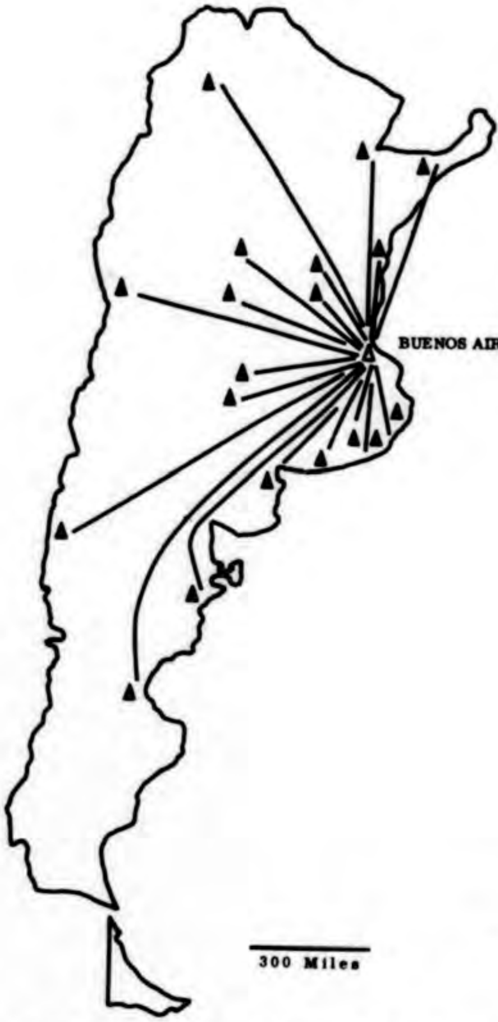


Figure 16. Argentina: Field Effect.

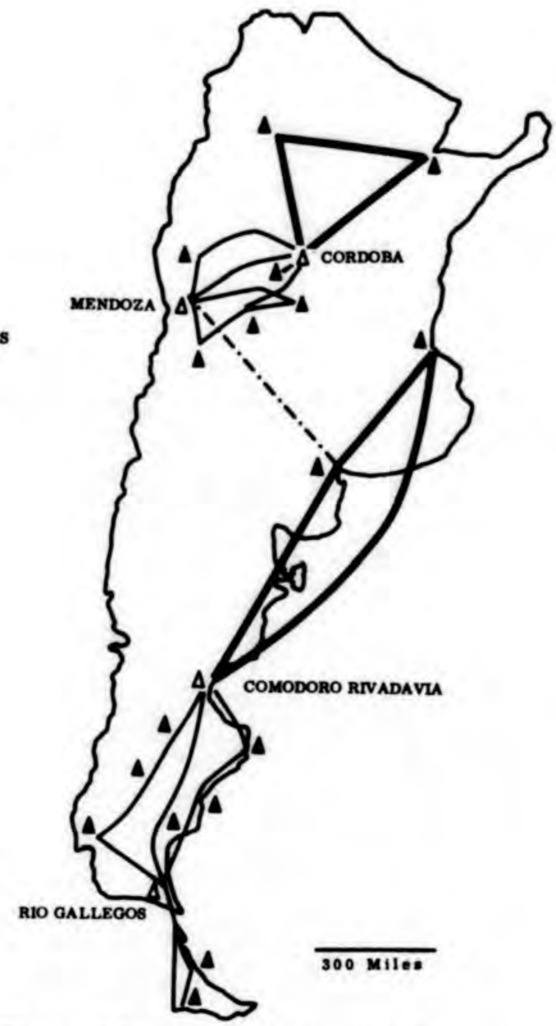


Figure 17. Argentina: Minor Regionalization Effects.

SUMMARY

The component analysis conducted on the connection matrix of Argentine local service airline routes yielded several basic patterns of variation. The first was an adjustment for size of node measured in terms of number of connections; the second was a major regionalization effect which picked out two major airline route regions within the nation; the third was a field effect centered on Buenos Aires; and the last was a two-member set of minor subregions which were tributary to the major regions derived earlier in the analysis and to the central node, Buenos Aires.

THE VENEZUELAN CASE

Venezuela lies very close to Argentina on the scales of economic-demographic development. Its local service airline routes constitute a substantial part of the nation's transportation system, serving 59 cities via 104 routes. The construction and operation of surface transport facilities are rendered difficult by mountain ranges and swamp lands and approximately four-fifths of the nation's population is concentrated in the northern highlands and coastal regions.

As of 1957 there are only three airlines operating domestic air route service in Venezuela (see the report by the U.S. Department of Commerce (Reference 24)). In addition nine non-scheduled airlines were recorded. The routes analyzed in the present study, as were the routes in Argentina, are those of the regularly scheduled passenger carriers. The 1957 Department of Commerce report listed 109 airports within Venezuela. Of these 109 airports only five were equipped for night landings; no information was available on runway types.

THE ROUTE STRUCTURE

The basic pattern of local service routes, as of late 1957, is displayed in Figure 18. This route map was taken from one prepared by

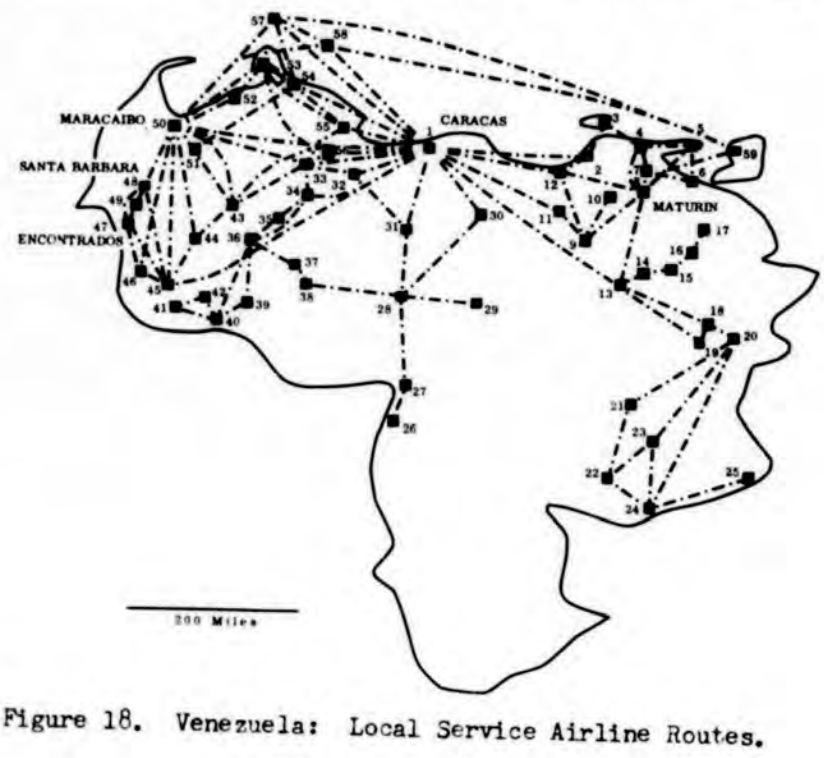


Figure 18. Venezuela: Local Service Airline Routes.

TABLE 29

CONNECTION MATRIX FOR LOCAL SERVICE AIRLINE ROUTES OF VENEZUELA

City	Connections																																																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
1. Caracas		x									x	x	x																	x	x	x	x	x										x		x					x			x	x	x	x	x	x	
2. Cumana	x		x									x																																														
3. Porlamar		x		x																																																						
4. Carupano			x		x		x	x																																																		
5. Guiria				x		x		x																																																	
6. Pedernales					x		x																																																		
7. Cachipo				x		x		x																																																	
8. Maturin				x	x		x		x			x	x																																												x
9. San Tome								x		x	x	x																																													
10. Santa Barbara									x																																																
11. Anaco	x								x																																																	
12. Barcelona	x	x						x	x																																																	
13. Ciudad Bolivar	x							x						x				x	x																																						
14. Puerto Ordaz													x		x																																									
15. San Felix														x		x																																								
16. Barrancas															x		x																																							
17. Tucupita																x																																									
18. Guasipati													x						x	x																																						
19. Upata													x					x																																								
20. Tumeremo																		x			x		x	x																																		
21. Kavanayen																				x		x																																				
22. Uriman																					x		x	x																																		
23. El Dorado																				x		x		x																																		
24. Icabaru																				x		x	x		x																																	
25. Santa Elena																								x																																		
26. Puerto Ayacucho																										x																																
27. Puerto Paez																									x		x																															
28. San Fernando																										x		x	x	x							x																					
29. Caicara																											x																															
30. Valle de la Pascua	x																											x																															
31. Calabozo	x																											x				x																											
32. Valencia	x																													x		x																										
33. Barquisimeto	x																														x		x									x							x					x	x			
34. Acarigua	x																															x		x	x																							
35. Guanare																																	x		x																							
36. Barinas																																	x	x		x		x	x																			
37. La Libertad																																			x		x																					
38. Brusual																											x									x																						
39. Palmarito																																			x				x																			
40. Guasdualito																																			x			x		x	x																	
41. Santo Domingo																																							x		x																	
42. Santa Barbara																																							x	x																		
43. Valera	x																															x											x						x	x								
44. Merida																																										x		x					x									
45. San Antonio	x																																										x		x	x	x		x									
46. La Fria																																											x		x			x									
47. Encontrados																																											x	x			x	x									
48. Santa Barbara																																											x		x		x	x									
49. Casigua																																													x	x											
50. Maracaibo	x																															x										x	x	x	x		x			x	x	x				x	x	
51. Lagunillas																																										x							x				x					
52. Dabajur																																																	x			x						
53. Las Piedras	x																																																x		x		x	x				
54. Coro	x																																																	x		x		x		x		
55. Puerto Cabello	x																															x																				x	x					
56. San Felipe	x																															x																										
57. Aruba	x																																																x				x				x	x
58. Curacao	x																																																x							x		x
59. Port of Spain (Trinidad)								x																																																x	x	

the Official Airline Guide, but for an earlier time period than the one for Argentina. The connection matrix which is equivalent to the route map is displayed as Table 29. An examination of Table 29 reveals a strong tendency toward the neighborhood effect mentioned earlier, together with a moderate development of clusters of entries which are located off the main diagonal. As of late 1957, the network was slightly over 6 per cent connected and less than 3 per cent redundant.

THE COMPONENT ANALYSIS

The component analysis was undertaken on the connection matrix described in Table 29 using the component analysis program (FA 108) of the University of Chicago's Computing Center. The results of the analysis are presented in Table 30.

TABLE 30

VALUES FOR 59 VENEZUELAN CITIES ON FOUR BASIC PATTERNS

City	Basic Pattern Values			
	1	2	3	4
1. CARACAS	3.163	-2.176	1.407	-0.801
2. Cumana	0.631	0.697	0.271	-0.030
3. Porlamar	0.120	0.017	0.476	0.154
4. Carupano	0.104	0.471	0.990	0.798
5. Guiria	0.080	0.381	0.831	0.577
6. Pedernales	0.026	0.145	0.484	0.440
7. Cachipo	0.080	0.381	0.831	0.577
8. Maturin	0.377	0.295	1.722	0.878
9. San Tome	0.271	0.071	0.960	0.309
10. Santa Barbara	0.042	0.159	0.206	0.095
11. Anaco	0.539	0.830	0.062	-0.066
12. Barcelona	0.702	1.009	0.652	0.158
13. Ciudad Bolivar	0.610	1.153	0.446	0.115
14. Puerto Ordaz	0.104	-0.125	0.327	-0.001
15. San Felix	0.016	0.087	0.045	0.009
16. Barrancas	0.002	-0.007	0.027	-0.004
17. Tucupita	0.000	0.006	0.003	0.000
18. Guasipati	0.123	-0.043	0.413	-0.036
19. Upata	0.120	-0.036	0.375	-0.003
20. Tumeremo	0.021	0.109	0.117	-0.094
21. Kavanayen	0.004	0.001	0.056	-0.072
22. Uriman	0.002	0.027	0.043	-0.089
23. El Dorado	0.004	0.011	0.073	-0.108
24. Icabaru	0.004	0.011	0.077	-0.118
25. Santa Elena	0.000	0.009	0.016	-0.035

Continued

TABLE 30 -- Continued

	City	Basic Pattern Values			
		1	2	3	4
26.	Puerto Ayacucho	0.005	-0.021	0.019	-0.035
27.	Puerto Paez	0.033	0.127	-0.014	-0.107
28.	San Fernando	0.216	-0.328	0.272	-0.362
29.	Caicara	0.033	0.119	-0.013	-0.097
30.	Valle de la Pascula	0.530	0.790	-0.157	-0.259
31.	Calabozo	0.674	0.679	-0.064	-0.423
32.	Valencia	0.882	0.643	-0.092	-0.484
33.	Barquisimento	1.723	-0.305	0.066	-0.620
34.	Acarigua	0.825	0.864	-0.208	-0.531
35.	Guanare	0.163	-0.110	0.104	-0.295
36.	Barinas	0.178	-0.129	0.129	-0.415
37.	La Libertad	0.034	0.041	0.013	-0.159
38.	Bruzual	0.038	0.119	-0.005	-0.143
39.	Palmarito	0.033	0.054	0.003	-0.171
40.	Guasdualito	0.035	0.062	0.004	-0.205
41.	Santo Domingo	0.006	0.000	0.009	-0.078
42.	Santa Barbara	0.006	0.000	0.009	-0.078
43.	Valera	1.428	0.698	-0.648	0.067
44.	Merida	0.839	-0.727	-0.248	0.500
45.	San Antonio	1.316	0.467	-0.998	1.106
46.	La Fria	0.680	-0.551	-0.407	0.823
47.	Encontrados	0.465	-0.372	-0.532	1.077
48.	Santa Barbara	0.711	-0.621	-0.482	1.012
49.	Casigua	0.186	-0.069	-0.353	0.640
50.	Maracaibo	2.436	0.105	-0.921	0.754
51.	Lagunillas	0.829	-0.528	-0.102	0.054
52.	Dabajur	0.616	-0.287	-0.157	0.118
53.	Las Piedras	1.387	0.290	-0.263	-0.240
54.	Coro	1.266	0.474	-0.146	-0.381
55.	Puerto Cabello	1.200	0.575	-0.150	-0.552
56.	San Felipe	0.772	0.757	-0.198	-0.336
57.	Aruba	1.361	0.351	-0.063	0.015
58.	Curacao	1.184	0.481	-0.127	0.129
59.	Port of Spain (Trinidad)	0.470	0.065	0.524	0.300

The analysis produced a set of four basic patterns of variation which were extracted from the information contained in the connection matrix. The four basic patterns explained approximately 38 per cent of the total observed variation, or structure, in the network (see Table 31). Again none of these basic patterns contained the neighborhood effect discussed previously. Based upon this over-all figure, it would appear that the network was somewhat less structured than that of Argentina.

TABLE 31

VENEZUELA
PROPORTION OF TOTAL VARIATION EXPLAINED
BY EACH BASIC PATTERN

Total Variance	208.00	Per Cent 100.00
Explained by:		
factor one	38.88	18.69
factor two	15.50	7.45
factor three	13.70	6.59
factor four	10.90	5.24
Total	78.98	37.97

The first factor, or basic pattern, again scaled the nodes in terms of their total number of direct connections. This factor accounted for nearly 19 per cent of the total observed variation versus 22 per cent for the previously examined Argentine case.

The second factor began the extraction of the basic structure of the network by breaking out a factor that can best be interpreted as a field effect centered upon the city of Caracas. This basic pattern explained approximately seven and one-half per cent of the observed variation as compared with the nine and one-half per cent explained by a similar factor in the Argentine analysis.

The third factor accounted for slightly over six and one-half per cent of the observed variation and represented a major regionalization effect similar to that seen in the Argentine analysis. Again two regions were extracted, but their internal cohesion appeared to be somewhat weaker than those observed in Argentina. The two regions were one in the west centering on Maracaibo and Santa Barbara, and one in the east centering around Caracas and Maturin. It is interesting to note that while similar factors have been extracted in both Argentina and Venezuela, a field effect and a major regionalization effect, the importance of the effects is reversed with the field effect being stronger in Venezuela than in Argentina. Figures 19 and 20 suggest the structure of the second and third factors.

The fourth and last factor succeeded in extracting only slightly above five per cent of the observed variation. This appeared to be a minor regionalization effect but an extremely weak one. The strongest portion of this pattern was a triangular feeder region centered on Caracas and extending westward toward Maracaibo. Some weak developments also appeared which were associated with the major Maracaibo, Santa Barbara region mentioned earlier, but these were not really strong enough to be classed as a further regionalization.

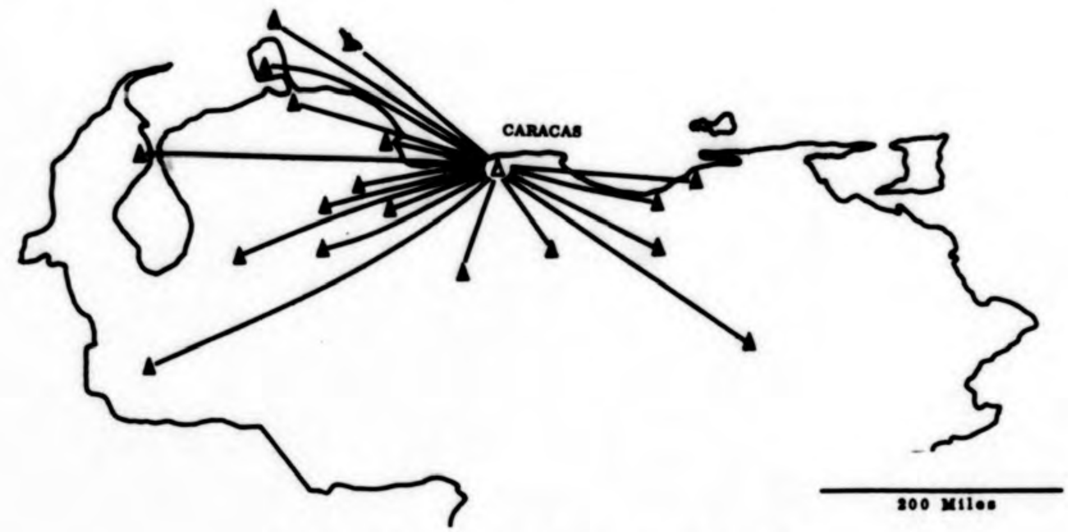


Figure 19. Venezuela: Field Effect.

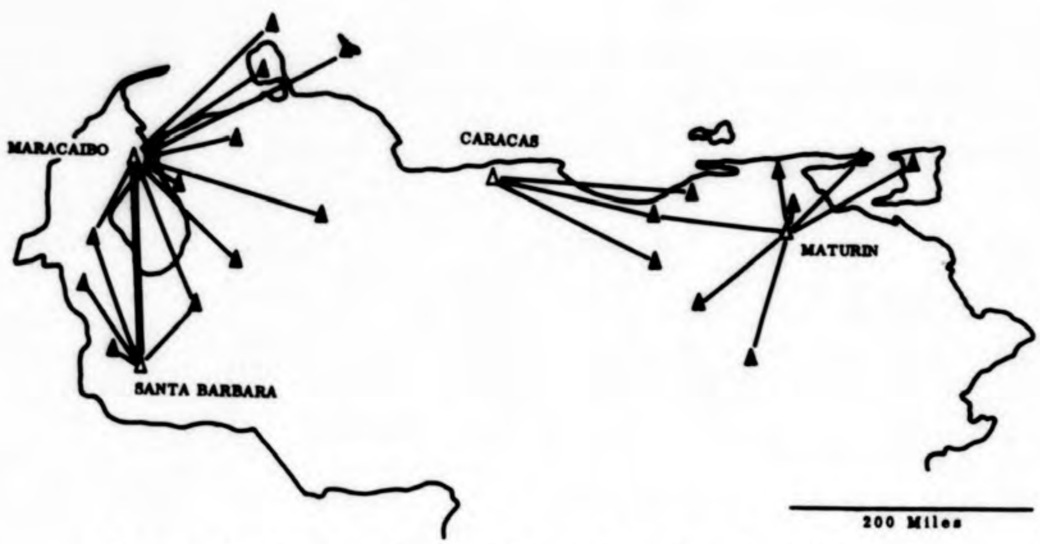


Figure 20. Venezuela : Major Regionalization Effect.

SUMMARY

The component analysis conducted on the connection matrix of Venezuelan local service airline routes yielded several basic patterns of variation. The first was an adjustment for size of node measured in terms of number of connections; the second was a field effect centered on the dominant urban center of the nation, Caracas; the third, a major regionalization effect which picked out two moderately well-defined airline route regions within the nation; and the fourth consisted of weak or minor regionalizations tributary to the major regionalizations mentioned earlier.

SUMMARY

The component analyses appear to have broken out several basic structural patterns within air transport networks. These basic effects appear to include: a field effect centered upon the major urban center of the country, a major regionalization effect consisting of a series of routes focusing upon a set of secondary urban centers, and a set of nominally triangular subregions or feeder areas which focus upon one or more of the secondary nodes as a gateway to the rest of the system.

Given these basic structural patterns it is felt that much of the remaining structure of the network may be accounted for in terms of two additional factors. First, there is a neighborhood effect, that is, a linking of nodes to nearby nodes. Second, there are system connections, that is, systems of chains and long simple circuits connecting regions.

CHAPTER 5. REPLICATING TRANSPORTATION NETWORKS

There is an article in the electrical engineering literature entitled, "How to Grow Your Own Trees from Given Cut-set and Tie-set Matrices." The subject of this chapter may be described by a somewhat similar phrase, namely: How to grow your own transportation networks from given graph-theoretic measures. This statement is to the point though somewhat flippant. Previous chapters have been used to report research on certain measures of transportation structure that have topological merit and that are empirically useful. This final chapter is concerned with actually replicating the geometry of transportation networks from measures of network structure.

The work reported in this chapter is part of current research, and the work is much less than complete. However, enough work has been done to indicate that the solution of this problem is practicable and to illustrate approaches to the solution. The materials presented in this chapter are sufficient to outline the problem and illustrate how it may be solved.

Previous chapters have identified information about transportation structure that may be estimated from knowledge of the characteristics of an area. This includes information on the number of vertices and nodes in the transportation system, knowledge of the redundancy or circuity of the system, and knowledge of edge length. In the task of replicating a transportation network this information may be assumed to be available because it may be generated from readily available knowledge about the characteristics of areas.

STUDY AREA

During the period of this research work has been done on the evolution of transportation networks of central Sweden, Sicily, Northern Michigan, and Ireland. For purposes of developing a formal model of the development process, it was decided to attempt to replicate the development of the railroad system of Northern Ireland. This area was judged small enough to be manageable for hand computation of models (though it turned out to be barely so). It contains a relatively complex system in comparison to some of the other small areas that were available for study (e.g., northern Michigan and Sicily); and a 130-year record of railroad development and abandonment was available. The area poses a case of considerable interest, too, because of population shifts and shifts in the structure of the economy during the period of railroad building and abandonment, and because of political developments during the period. An "easy" area was not desired. The section below presents certain background notes that amplify these characteristics of the area.

BACKGROUND NOTES

The term "Northern Ireland" is applied to the six northern countries of Ireland which retained close connections with Great Britain after the partition of Ireland in 1922. It is important to note that the highway and railroad systems of the six counties were constructed before the partition, and formed an integral part of the transportation network of the entire thirty-two counties of Ireland. The dichotomy which produced Northern Ireland also dislocated the transportation network with serious consequences to the railroad system and, to a lesser extent, the highways.

The change in population distribution between 1831 and the present day is a factor which cannot be overlooked in the study of the rail network since the changes that have occurred have for the most part provided an unfavorable background for the development of a rail system. During the one-hundred-and-thirty-year period under review continued emigration has resulted in considerable decreases in population in five of the six counties, the exception being County Antrim. In contrast to this decline in population, the city of Belfast has achieved a dominant position with an exceptionally high growth rate. In 1951 the population of the city was 443,671, compared to 53,287 in 1831. Elsewhere there has been a tendency for the larger country towns to at least maintain their 1831 population levels while those towns within a thirty-mile radius of Belfast have on a smaller scale shared a similar growth to that of Belfast. Notwithstanding this increased urbanization, three-fifths of the population live in rural areas in the western counties of Londonderry, Tyrone, and Fermanagh, and if the population of Belfast and suburbs were omitted from the totals for Counties Antrim, Down, and Armagh, a similar rural distribution would be revealed. The population distribution, therefore, has never been ideally suited to rail transportation. The bus, truck, and private automobile have provided the twentieth century answer to the problem of rural accessibility.

Before the construction of the railroads, land transportation in the area was slow and expensive. A limited network of horsedrawn car services had been established by Edward Bianconi and other entrepreneurs in the early years of the nineteenth century. The most important routes were those connecting Belfast with Armagh; passing through the towns of Lisburn, Lurgan and Portadown, and with the port of

Newry and thence to Dublin. There were in addition significant concentrations of traffic on the essentially local routes from Belfast to Holywood, Newtownards, Antrim, and Carrickfergus. Connections to the seaports of Counties Antrim and Londonderry were principally by coastal steamers before the advent of the railroads.

The opening of inland navigation in the eighteenth century made a significant contribution to the commercial and industrial development of many inland towns in Great Britain. In spite of a natural focus for an inland waterway system presented by Lough Neagh, canals in the north of Ireland were never a success. Lack of industrial development and a relatively sparse population were contributory factors to the failure of the inland waterways. In addition, the construction of railroads followed so soon after the canal era that canal companies scarcely had time to recover their initial costs before they were subjected to competition from a more efficient rival.

The physiography of the six counties presented substantial obstacles to railroad construction, the central area being occupied by the Sperrin Mountains and Lough Neagh, the eastern zone by the Antrim Plateau, and the south by the Mountains of Mourne and the Drumlin Drift belt. The rail routes which evolved tended to be circuitous, a fact which contributed little to their competitive position at a later date.

The Royal Commission on Irish Railways, in its report in 1837, expressed no great hopes that railways would ever become remunerative in Ireland, and recommended the construction of a few trunk lines only. The proposed route for the northern counties extended from Belfast up the Lagan valley to Portadown and Armagh and then southwards to Navan and Dublin. The object was to make the new rail route complementary to the existing canal network. The poor prospects for profits offered by investment in Irish railways deterred English financiers from risking their capital, and the network grew slowly and in a piece-meal manner. The most promising links were established first, e.g., Belfast to Portadown, Newry, Coleraine, Londonderry, Larne, and Holywood.

Those routes which from an investors point of view were distinctly marginal were added during the last two decades of the nineteenth century. During the twenty years following 1840 the rail network developed so slowly that there were frequent demands for government aid for railway construction. Public works, such as railway construction, were seen by many as a panacea to the destitution which resulted from the 1846 potato famine. However, such aid was not forthcoming until the passing of the Tramways and Public Companies Act of 1883 and its more notable successor, the Light Railways Act of 1889. The latter proved to be a great stimulus to rail construction because it provided for a degree of state aid for the construction and maintenance of new railways subject to certain conditions.

The act permitted the lowering of normal construction standards in order to reduce the initial costs of the lines built under its auspices.

The period between 1880 and 1900 saw the virtual completion of the network. Many of the routes constructed under the Light Railways Act were of narrow gauge (three feet as compared to the standard Irish broad guage of five feet and three inches), and these routes penetrated into thinly populated rural areas, particularly in Antrim and Tyrone, where traffic potential was low. Few of these lines ever operated at a profit and all are now abandoned. As a general rule it may be stated that those routes which were the last to be added to the network were the first to be abandoned. It should be noted that these narrow gauge routes functioned essentially as feeders to the broad gauge lines, and there was no tendency for a narrow gauge system to develop independently of the broad guage such as in Belgium.

Throughout the nineteenth century there was little unification of control, and a plethora of small and often financially unsound companies operated the component segments of the network. Under such conditions a tendency for amalgamation was a natural development, and by 1900 there emerged three companies which were to continue in existence until 1947: the Great Northern Railway (Ireland), the Belfast and Northern Counties Railway (which became the Northern Counties Committee after it had been taken over by the Midland Railway of England in 1903), and the Belfast and County Down Railway. Other operators were of little significance. Each company, until the advent of highway competition, enjoyed a monopoly of a fairly compact area. The Northern Counties Committee linked Belfast to principal centers in Antrim, Eastern Tyrone, and North County Londonderry. The Great Northern Railway provided service from Belfast to north and western parts of County Down, County Armagh, South and Southwest Tyrone, South County Londonderry, County Fermanagh, and to the northeastern region of what has since become the Republic of Ireland. The Belfast and County Down was, as its name suggests, a local operation linking Belfast to points in County Down. The only competing routes that existed were those linking Belfast with Londonderry, Cookstown, and Newcastle (County Down). In 1910-1920, when the network had reached its maximum stage of development, the route mileage had reached 773 miles and every town with a population exceeding 1,750 had rail service. The focal points on the six county networks from which four or more routes radiated in 1920 were Belfast, Lisburn, Portadown, Goraghwood, Armagh, Ballymena, Londonderry, Strabane, and Magherafelt.

In 1922 the partition of Ireland resulted in the imposition of a political frontier on an established rail network which crossed the new boundary at eleven places. Customs posts were established at each crossing point, and the regular inspection of all trains inevitably delayed movement, particularly that of freight. Following partition the economic policies adopted by the Irish Free State and Northern Ireland resulted in a decline in the flow of freight across the border. Belfast, already the focal point of the Northern Ireland railway system, now assumed additional importance as the capital city.

The Great Northern Railway, which operated in an area bounded by Belfast, Londonderry, Bundoran (Eire), and Dublin (Eire), encountered operating difficulties on many of its routes. In addition to the usual delays occasioned by customs examination, political unrest in the border areas persistently manifested itself in damage to rail installations for a further thirty-five years.

By 1930 private bus operators had extended their services to even the remotest areas. The bus operator possessed the advantage of being able to reach a predominantly scattered rural population, which was not clustered in urban areas such as in the United Kingdom. This new form of competition rapidly eroded the traffic of the secondary routes. Between 1930 and 1941 four narrow gauge sections were closed to all traffic as were four broad gauge lines. During this period the Great Northern Railway developed an extensive network of bus services which were intended to be complementary to its rail services.

World War II brought immense operating difficulties for the rail systems. All railways in Northern Ireland were brought under government direction and were taxed to capacity dealing with heavy military traffic to the various army camps and airfields established throughout Northern Ireland. The difficulties of the Great Northern Railway were increased by the fact that it was now operating partly in Northern Ireland which was at war, and partly in neutral Eire.

The legacy of the war was an accumulation of dilapidated equipment, much of which was beyond repair. The time had now come for a reassessment of the position of rail and road transportation in the national economy. The Transport Act of 1947, which nationalized the railways of Britain, prompted the Northern Ireland government to assume control of the Northern Counties Committee Railway and the Belfast and County Down Railway, together with all bus services in the six counties. The Great Northern Railway, because of its international status, continued to function separately although it was now subjected to competition from its former bus services in Northern Ireland, which had been taken over by the newly formed Ulster Transport Authority.

By 1956 the Great Northern Railway had reached a state of bankruptcy, and joint control was assumed by the governments of Northern Ireland and the Republic of Ireland. The refusal of the Northern Ireland government to continue its subsidization of Great Northern rail services after October 1957 resulted in the collapse of joint control, and the rail routes within Northern Ireland passed to the

Ulster Transport Authority. Immediately all but four of the former Great Northern routes were abandoned, which left all of County Fermanagh and much of County Tyrone without rail service. Thus from October 1957 all public passenger transportation in Northern Ireland became the monopoly of the Ulster Transport Authority.

Since 1947 the Northern Ireland government has pursued a policy based on the assumption that bus and truck transportation are best adapted to conditions within the six counties. Thus apart from the dismemberment of the former Great Northern Railway, the entire Belfast and County Down Railway, with the exception of twelve significant miles between Belfast and Bangor, was abandoned. Similar economy measures were applied to the network of the former Northern Counties Committee with the result that all remaining narrow gauge sections were closed and four broad gauge lines were also abandoned. In most cases where a section of line was abandoned it was not necessary to institute a new replacement bus service as such services had been in operation for several years.

Today there remains in Northern Ireland a severely truncated rail network which is composed principally of two types of routes -- trunk lines and commuter routes. The latter type has become particularly important in recent years. Heavy commuter traffic is now carried by the following routes: Belfast-Holywood-Bangor, Belfast-Lisburn-Lurgan-Portadown, Belfast-Carrickfergus-Larne, and to a limited extent between Ballymena and Belfast. The Larne route has an additional significance in that it connects with the steamer services to Stranrear, Scotland. The existing trunk routes consist of the former N.C.C. route from Belfast to Londonderry via Ballymena, Ballymoney, and Coleraine, and the former Great Northern route to Londonderry via Portadown, Dungannon, Omagh, and Strabane. In addition there is the main line from Belfast via Portadown to Dublin (Republic of Ireland). Apart from these routes, two branch lines continue to carry passenger traffic while two others deal with freight traffic only. Rail transportation is now relatively insignificant as far as the carriage of passengers is concerned and even more so in the sphere of freight traffic. A dense network of frequent bus services now covers the entire province and nearly all freight moves by truck.

MAPS

The extension of routes on the Irish transportation system and their subsequent abandonment is shown on the accompanying series of seven maps (Figure 21). All but one of the maps deals with the period of development because the present study attempts to replicate the development phase. It may be noted that the first loop on the system does not appear until the 1870 map, but that by the time that the system reached its maximum extension in the early decades of the 1900's the system was characterized by a number of loops. Also, the

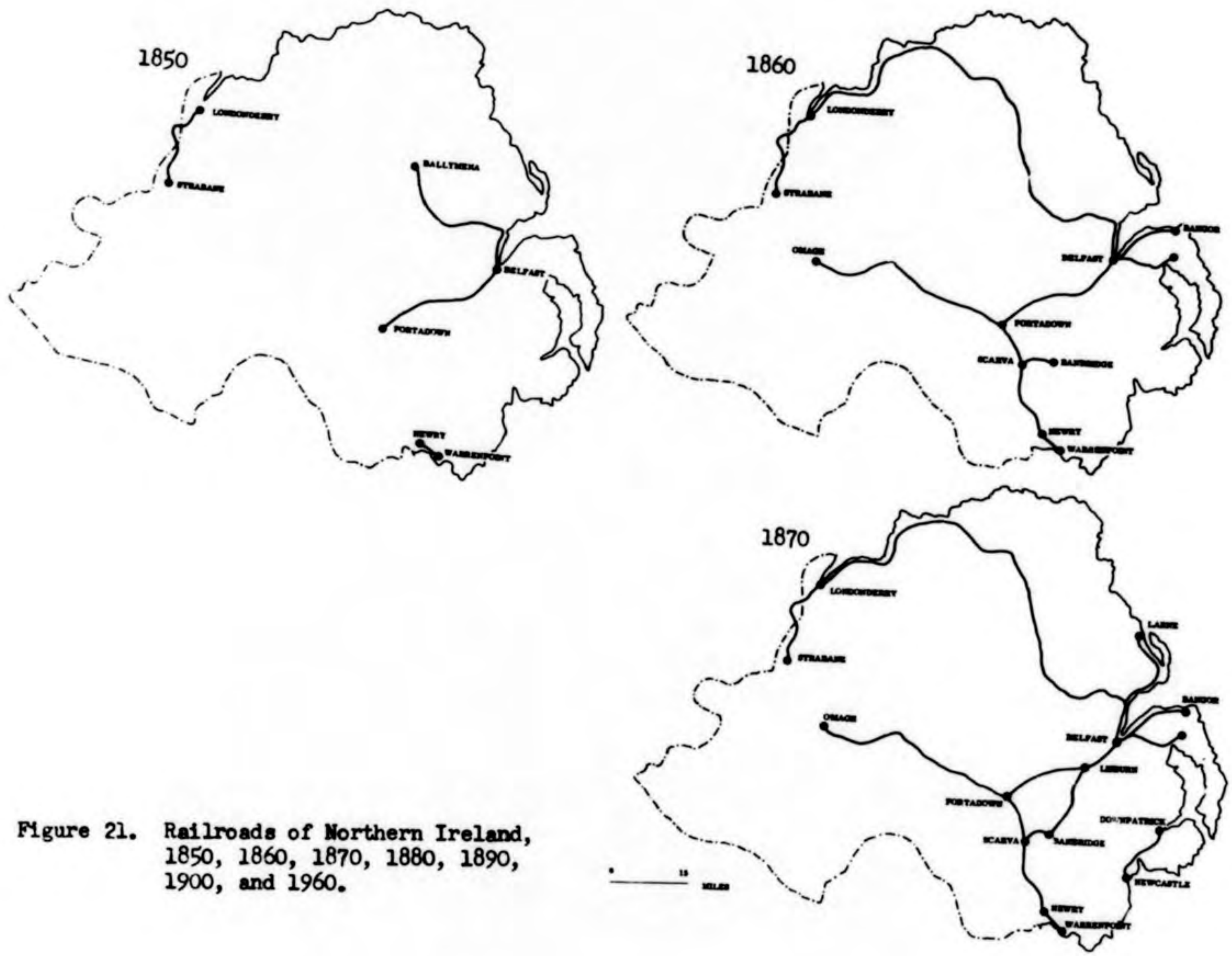


Figure 21. Railroads of Northern Ireland, 1850, 1860, 1870, 1880, 1890, 1900, and 1960.

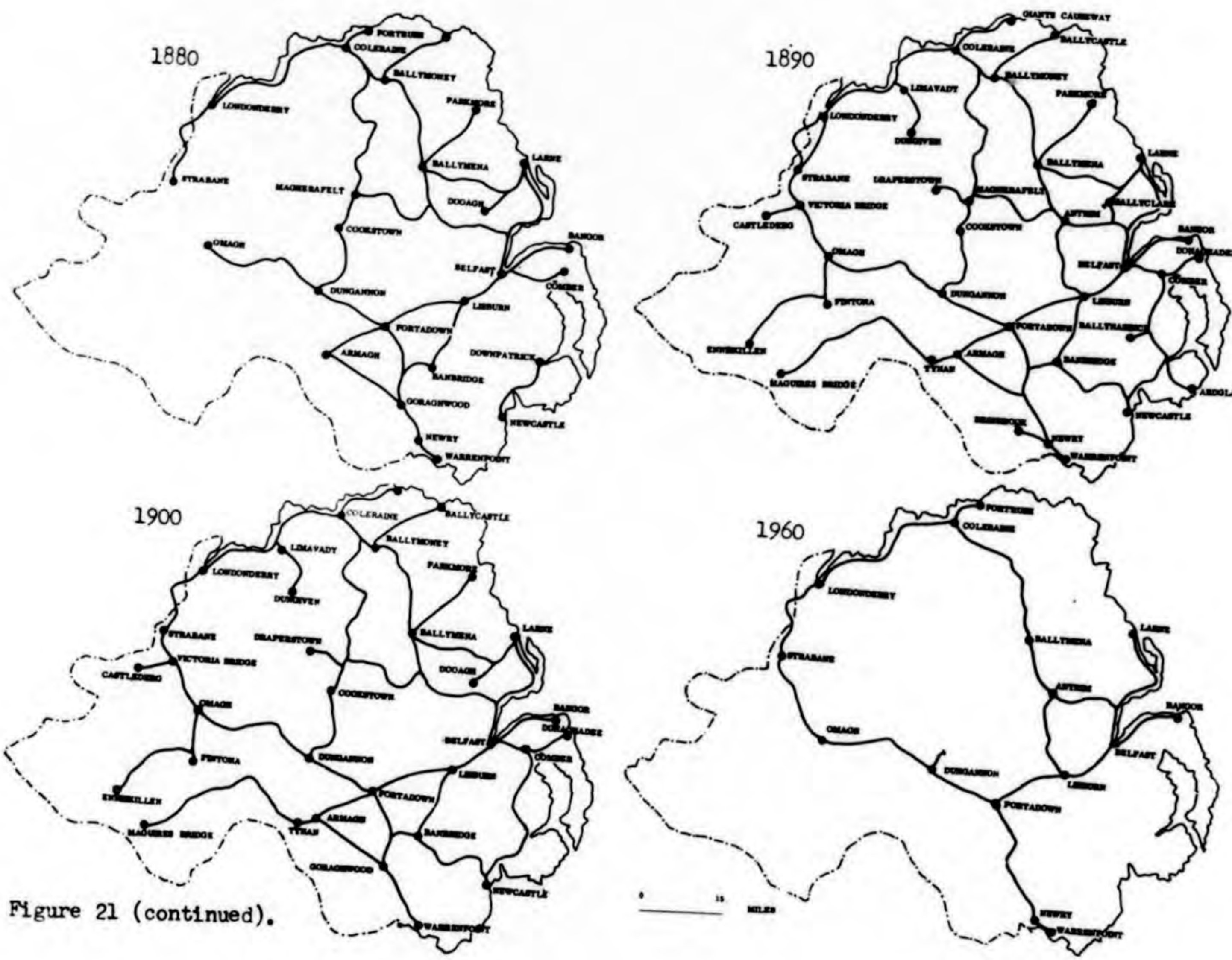


Figure 21 (continued).

present system consists for the most part of a loop around the central portion of the area. As was mentioned before, system mileage reached its greatest extent in 1910-1920 when there were some 773 miles of railroad in the area. By 1960 the mileage had been reduced to approximately 275 miles.

THE MODELS

Three models were used to replicate the Northern Ireland railroad network. In order to narrow the subject matter of the models, they were limited to the period of the development of the system -- 1830 through 1910-1920. The connections of Northern Ireland with what is now The Republic of Ireland were disregarded. Because it is known that the number of links, vertices, and the circuity of the network may be determined from characteristics of any area, the number of links, vertices, and the degree of circuity of the network were taken as known input quantities. The models consisted of decision rules which used input information to produce maps of the transportation system.

Certain other outside data were also used as inputs to the model. The places eligible to be vertices on the transportation system were taken to be all places with populations greater than 100 in 1900. A more precise operation of the model would use population data that varied from time to time during the course of the development of the network. An inspection of population data revealed that, with the exception of Belfast, urban populations were relatively stable during the period of study, and the use of a single set of population data simplified the data-handling problem considerably. The relative locations of these places were also used as inputs to the model.

There were forty places in Northern Ireland in 1900 with populations of 1,000 or greater and sixty-three with populations greater than 100. The matrix of distances between these places was triangular with sixty-two rows and sixty-two columns. This was simply a distance array similar to those on petroleum company touring maps and the distance arrays that appear in Atlases. Populations were multiplied together to form a second triangular matrix of the same order. The ij^{th} cell of this matrix contain the population of the i^{th} place times the population of the j^{th} place.

TWO STOCHASTIC MODELS

The term stochastic model refers to a model with probability properties. Models of this type are especially desired because the alternate to this type of model, a deterministic model, leaves no room for uncertainties and chance. Stochastic models also take account of some of the facts revealed in the previous two chapters of this monograph. For instance, a given set of conditions within an area would tend to give rise to a certain type of transportation network, but observations have revealed some variation about expected values. (From an operational point of view this corresponds to variation in the error term, or residuals, of the regressions, or unaccounted for variance in the components studies.)

The manner in which a stochastic model works may be described in a few sentences. The model has input data available and selects particular inputs from these data. The inputs are operated upon by decision rules within the model, generating outputs. Outputs in the present case are maps of the transportation network. The models tested in this research were operated for ten-year intervals. There were eight such intervals, 1830-1840, 1840-1850, etc. through 1900-1910, and thus there were eight inputs to and eight outputs from each model.

Each input was the number of links to be added during the period. This number was determined from a frequency distribution of the number of links actually added when the railroad transportation system was constructed. A table of random numbers was consulted and random numbers were matched against numbers arbitrarily assigned to the frequency distribution. For example, if the number 4 was obtained from the list of random numbers, then six links would be added during the period because six links were built in the fourth decade of the construction of the railway system.

Certain of the decision rules within the models also involved probability considerations. For each edge or link, three questions had to be answered:

1. Does this link make a loop?
2. What is the length of this link?
3. Where does this link go?

The first two of these questions were answered by consulting a frequency distribution of link lengths and a frequency distribution of links that form loops on the system. The manner in which these data were treated differed somewhat between the two stochastic models as did the answer to the third question, and further comments on decision rules will be made when these models are discussed.

EXPECTED FINDINGS

Because models of this type may not be familiar to the reader, it might be well to make certain remarks on what may be expected from the operation of a model of this type. A model of this type provides three kinds of information. For one, it provides a set of decision rules that "explain" development of transportation systems. These decision rules are unambiguous statements of the mechanics of system development. It is true that the development of a transportation system is actually based upon many, many decisions. On the other hand, it has been found that systems of this complexity can be replicated using a few decision rules, and there is every reason to believe that this would be true for transportation systems.

Second, the model provides a way of studying the sensitivity of the development process to changes in its parameters. It should be understood that a stochastic model will generally give a different result each time the solution is worked out. Working out a number of solutions under one set of parameters gives a notion of possible outcomes under that set of parameters. If it is desired to know how the situation might change if the parameters changed, then parameters can be changed and the model solved a number of additional times. By comparing the patterns generated before and after the shift in parameters, one can obtain certain notions of the sensitivity of the system to these parametric shifts.

FIRST STOCHASTIC MODEL

The first stochastic model did not give results at all like the development of the Northern Ireland transportation network. However, this model will be discussed here very briefly in order to make it a matter of record, and this may be helpful to others working with models of this type.

For purposes of exposition, the operation of the model will be discussed for one of the time periods after the operation has started. The input to the model was the number of routes to be built during the period. This number of routes was selected using random numbers from the frequency distribution of number of links. The decision rules in the model allocated these links one at a time. The first question asked was length of the first link to be added during the period. This was determined by entering the cumulative distribution of link lengths using random numbers.

The second question was where that route was to be placed. This question was answered by the single decision rule contained in the model, namely: inspect the matrix of population products (discussed earlier) for the maximum entry p_ip_j, where p_i is a place on the network and p_j is a place either <u>on or off</u> the network within the link length determined earlier, and connect p_i with p_j. In other words, the distance measure determined all places eligible to be connected by the link to be added. In some cases this rule would lead to the linking of a new place to the network, that is, p_ip_j = maximum for some j not on the network. In other cases p_ip_j was a maximum for some j already on the network and a loop was constructed.

The operation of the model using this rule quickly revealed that the larger places on the network were dominating the system, nearly all routes were leading directly to Belfast, and work with this model was stopped.

SECOND STOCHASTIC MODEL

A second stochastic model was operated using a second set of decision rules. Again, the inputs to the model were the number of routes to be built during the time periods. The central decision rule was: link to the system the largest place off the system but within the link length. (As was true of the first stochastic model but not mentioned there, a special rule had to be used for the first link during the first time period. This was, start the system by linking the largest pair.) Two supplemental considerations were made. The first of these was the establishment of the lengths of links, and this was determined by sampling from a frequency distribution of link lengths.

Another consideration in this model was whether or not a link to be added was part of a loop. This question was asked and answered for every link before that link was added to the system. For every time period it was known what proportion of the links in the system were excess to those on the tree -- i.e., the cyclomatic number. This proportion was sampled from random numbers. If the instruction to add a loop was received, the two largest places on the system (not previously connected by loop-making links) within the link length permitted were connected.

The actual rail network in 1910 is shown in stylized form in Figure 22 and one map achieved by operating the second stochastic model is shown in Figure 23. Correspondence between the two maps is not very great and this second model is not regarded as successful. Actually, a model should not be accepted or rejected on the basis of a single run. A number of trials should be made with the model and the model accepted or rejected on the basis of the general pattern derived. However, in this case the map achieved is so far from that actually existing in 1910 that additional operations of the model would seem to be of little value.

CRITICISM

Neither of the two stochastic models proved useful and useable. The reason for this is very simple -- not enough work has been done on unambiguous decision rules and not enough work has been done on model inputs. This type of model is part of current research on which work is just beginning and results of certain work on allocation rules are not yet available. In particular, a rule is needed relating to

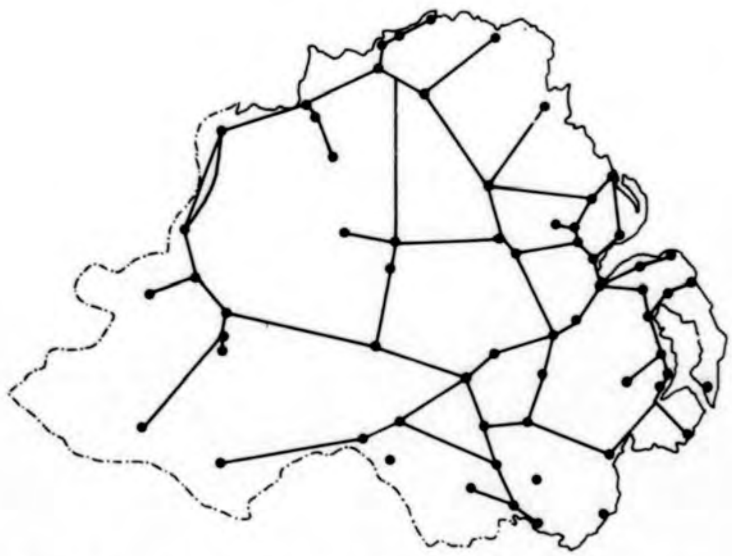


Figure 22. Railroads of Northern Ireland, 1910.

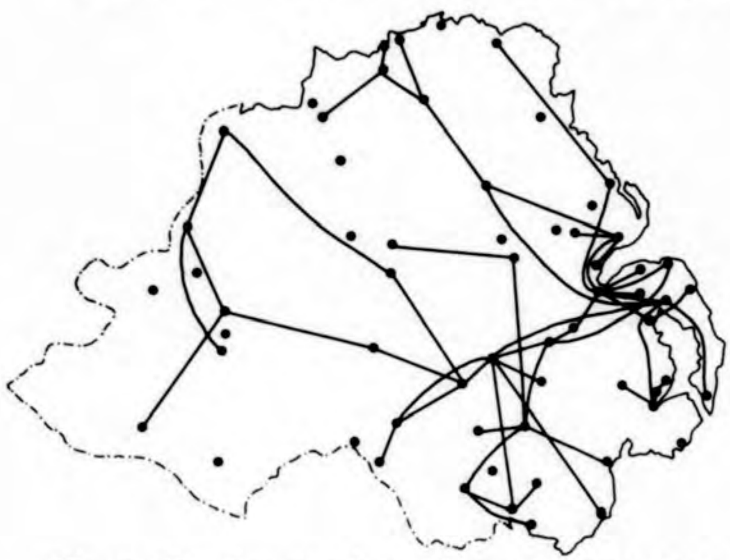


Figure 23. Output of the Second Stochastic Model.

the deviation of paths between major places in order to pass through intermediate places. This question of angles of deviation is under study. The amount of path deviation is a function of variable cost versus fixed cost in a system. If variable costs are high and fixed costs are low, then there should be little deviation. When variable costs are low relative to fixed cost, then deviation from straight line paths may be warranted in order to keep down total route mileage in the system.

A DETERMINISTIC MODEL

Even though the stochastic models are not yet operational, previous research is sufficient to generate networks very similar to those that actually appear. Considerations in Chapter 3 are sufficient to determine the number of vertices, edges, and the amount of circuity in a network. Considerations in Chapter 4 illustrate the presence of field, neighborhood, and regional effects in the arrangement of links on networks. It was found that a set of three rules would allocate links to a network according to the field, regional, and neighborhood principles, and to the degree required by the estimates of vertices and links. (In this case, however, the numbers of vertices and links were not used as constraints on the system.) Figures 24 and 25 show the network structure for Ireland developed under these rules.

The rules were very simple. The neighborhood effect was taken account of by the rule that each place be connected to its nearest neighbor. The result was a series of subgraphs (Figure 24). The regional effect was taken care of by requiring that each subgraph be connected to the nearest subgraph. For purposes of this trial the field effect was restricted to Belfast. It was required that each subgraph be linked to other subgraphs along a Belfast axis, but with a twenty degree deviation allowed. For instance, the line between place 38 and place 39 represents the field effect which connects a subgraph (made up of places 35, 47, 46, 38, 48 and 49) and another subgraph (places 39, 43, 44 and 58). Table 32 provides a key to the place names indicated by these numbers. The subgraph of places located between 35 and 49 is linked to two other subgraphs via the rule that subgraphs be linked to their nearest subgraphs. This subgraph is closest to the subgraph of 42 and 59 and it is also linked to the subgraph consisting of places 1, 63, 2, and 60.

Figure 25 may be compared with the reference map, Figure 22. The two are not exactly the same, of course. One reason for this is the failure to consider field effects other than for Belfast. Also, because the number of links added depends upon the number of subgraphs identified in the first step, the number of edges on the system was determined by the configuration of the points to be served and not by exogenous information. The map could be improved by using a rule that would determine a number of edges on the system other than that determined by the configuration of the vertices. However, these improvements have not been made simply because further work with the deterministic model has been postponed until after further work has been done with stochastic models.

Figure 24. Subgraphs.

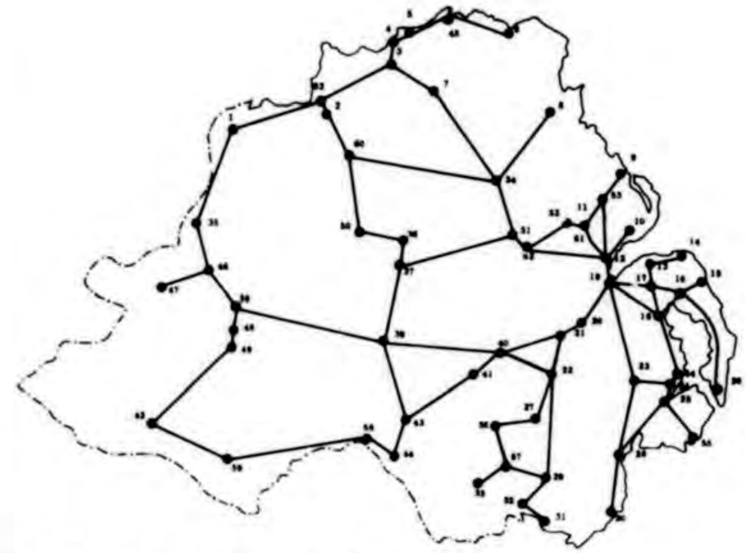


Figure 25. Output of the Deterministic Model.

TABLE 32

KEY TO CITY NUMBERS ON FIGURE 25

1.	Londonderry	23.	Ballynahinch	43.	Armagh
2.	Limavady	24.	Ballynahinch Junction	44.	Keady
3.	Coleraine	25.	Downpatrick	45.	Giants Causeway
4.	Portstewart	26.	Portaferry	46.	Victoria Bridge
5.	Portrush	27.	Banbridge	47.	Castlederg
6.	Ballycastle	28.	Newcastle	48.	Fintona Junction
7.	Ballymoney	29.	Rathfriland	49.	Fintona Town
8.	Parkmore	30.	Kilkeel	50.	Draperstown
9.	Larne	31.	Warrenpoint	51.	Cookstown Junction
10.	Carrickfergus	32.	Newry	52.	Dooagh
11.	Ballyclare	33.	Bessbrook	53.	Ballyboley Junction
12.	Whiteabbey	34.	Ballymena	54.	Killylaagh
13.	Holywood	35.	Strabane	55.	Ardglass
14.	Bangor	36.	Magherafelt	56.	Scarva
15.	Donaghadee	37.	Cookstown	57.	Goraghwood Junction
16.	Newtownards	38.	Omagh	58.	Tynan
17.	Dundonald	39.	Dungannon	59.	Maguires Bridge
18.	Comber	40.	Lurgan	60.	Dungiven
19.	Belfast	41.	Portadown	61.	Ballyclare Junction
20.	Dunmurry	42.	Enniskillen	62.	Antrim
21.	Lisburn			63.	Limavady Junction
22.	Dromore				

SUMMARY AND FURTHER EVALUATION

This chapter has presented some fragmentary work which uses information on network structure to generate actual maps of transportation networks. The case at point was Northern Ireland and attempts were made to replicate the map of its rail network. The attempt to replicate the map using a simple set of rules and stochastic information has not yet proved successful. However, the network can be very nearly replicated through an arbitrary set of rules based on neighborhood, regional, and field notions.

The discussion was based on preliminary work with models. The work that has been done is so slim that there is very little basis for the over-all evaluation of this phase of the research. However, experience by the researchers and others in other cases indicates that development of such models is practicable and that such models prove extremely useful. These models may be used to evaluate the variety of networks that might develop from a given set of conditions, and for the study of the effects of shifts of parameters on network development.

BIBLIOGRAPHY

1. Bartlett, M.S., An Introduction to Stochastic Processes, The University Press, Cambridge, Massachusettes, 1955.

2. Beckmann, M., "A Continuous Model of Transportation," Econometrica, XX, 1952, pp. 643-660.

3. Beckmann, M., "Principles of Optimum Location for a Transportation Network," paper given at NAS-NRS Symposium on Quantitative Problems in Geography, 1960.

4. Berge, C., Théorie des Graphes et ses Applications, Dunod, Paris, 1958.

5. Berry, Brian J. L., "Basic Patterns of Economic Development," Chapter in Atlas of Economic Development by N.S. Ginsburg, ed., University of Chicago Press, Chicago, 1961.

6. Berry, Brian J. L., "An Inductive Approach to the Regionalization of Economic Development," Chapter in Essays on Geography and Economic Development by N. S. Ginsburg, ed., University of Chicago Press, Chicago, 1960.

7. Bhagat, D. G., "Regional Concept for Highway Planning -- The Web and Lattice Pattern," Journal of the Indian Roads Congress, XXIV, 1960, pp. 629-666.

8. Cherry, E. Colin, "Generalized Concepts of Networks," Proceedings of the Symposium on Information Networks, Polytechnic Institute of Brooklyn, New York, 1954, pp. 175-184.

9. Garrison, William L., "Connectivity of the Interstate Highway System," Papers and Proceedings, Regional Science Association, VI, 1960, pp. 121-137.

10. Ginsburg, N. S., Atlas of Economic Development, University of Chicago Press, Chicago, 1961.

11. Harman, Harry H., Modern Factor Analysis, University of Chicago Press, Chicago, 1960.

12. Hohn, F. E., S. Seshu, and D. D. Aufenkamp, "The Theory of Nets," Transactions of the Institute of Radio Engineers, EC-VI, 1957, pp. 154-161.

13. König, D., Théorie der Endlichen und Unendlichen Graphen, Chelsea Publishing Co., New York, 1950.

14. Lalanne, Leon, "An Essay on the Theory of Railway Systems, Based on Observation of Facts and the Basic Laws Governing Population Distribution," Comptes rendus des seances de L'Academie des Sciences, LVII, 1863, pp. 206-210.

15. Lösch, August, The Economics of Location, Translated from the second revised edition by W. H. Woglom, Yale University Press, New Haven, Conn., 1954.

16. MacRae, Duncan, Jr., "Direct Factor Analysis of Sociometric Data," Sociometry, XXIII, 1960, pp. 360-371.

17. Mital, N. C., "Crystal Pattern Formula for Integrated Highway Planning in India," Journal of the Indian Roads Congress, XXIV, 1959, pp. 336-399.

18. Nystuen, John D., and Michael F. Dacey, "A Graph Theory Interpretation of Nodal Regions," Papers and Proceedings, Regional Science Association, VII, 1961.

19. Ross, Ian C., and Frank Harary, "On the Determination of Redundancies in Sociometric Chains," Psychometrika, XVII, 1952, pp. 195-208.

20. Seshu, Sundaram and Myril B. Reed, Linear Graphs and Electrical Networks, Addison-Wesley Publishing Co., Inc., Reading, 1961.

21. Shimbel, Alfonso, "Applications of Matrix Algebra to Communication Nets," Bulletin of Mathematical Biophysics, XIII, 1951, pp. 165-178.

22. Shimbel, Alfonso, "Structural Parameters of Communication Networks," Bulletin of Mathematical Biophysics, XV, 1953, pp. 501-507.

23. U. S. Dept. of Commerce, Civil Aviation in Argentina, World Trade Information Service, Part 4, No. 57-9, June, 1957.

24. U. S. Dept. of Commerce, Civil Aviation in Venezuela, World Trade Information Service, Part 4, No. 57-8, May, 1957.

25. Vadnal, A., "The Location of the Communication Network in a Circle," paper given at the European Economic Congress, 1958.

26. Veblen, Oswald, Analysis Situs, American Mathematical Society, Providence, 1931.

27. Whitney, Hassler, "Congruent Graphs and the Connectivity of Graphs," American Journal of Mathematics, LIV, 1932, pp. 150-168.

28. Whitney, Hassler, "On the Classification of Graphs," American Journal of Mathematics, LV, 1933, pp. 236-244.

29. Whitney, Hassler, "A Set of Topological Invariants for Graphs," American Journal of Mathematics, LV, 1933, pp. 231-235.

30. World Railways, Henry Sampson, ed., Sampson Low's, Ltd., London, 1960.

DISTRIBUTION

DEPARTMENT OF DEFENSE

Transport & Ships Division
OFMO, ODDR&E
Room 3D-1075, The Pentagon
Washington 25, D. C. (1)

Equipment & Supplies Division
OFMO, ODDR&E
Room 3D-1075, The Pentagon
Washington 25, D. C. (1)

UNITED STATES CONTINENTAL ARMY COMMAND

Commanding General
Second United States Army
ATTN: Transportation Officer
Fort George G. Meade, Maryland (1)

Commanding General
Third United States Army
ATTN: Transportation Officer
Fort McPherson, Georgia (1)

Commanding General
Fourth United States Army
ATTN: Transportation Officer
Fort Sam Houston, Texas (1)

Commanding General
Fifth United States Army
ATTN: Transportation Officer
1660 East Hyde Park Boulevard
Chicago 15, Illinois (1)

Commandant
Army War College
ATTN: Library
Carlisle Barracks, Pennsylvania (1)

Lt. Col. Oliver R. Dinsmore
Army Research Office
Box CM, Duke Station
Durham, North Carolina (1)

President
United States Army Armor Board
Fort Knox, Kentucky (1)

Headquarters
U. S. Army Aviation Test Office
ATTN: FTZAT
Edwards Air Force Base, California (1)

Deputy Chief of Staff for Logistics
ATTN: Requirements Division (1)
ATTN: Materiel Maintenance Division LOG/B4 (1)
Department of the Army
Washington 25, D. C.

Director
The Research Analysis Corporation
ATTN: Library
6935 Arlington Road
Bethesda 14, Maryland (1)

Army Research Office
Office of the Chief of Research and Development
ATTN: Research Support Division
Department of the Army
Washington 25, D. C. (1)

Chief
U. S. Army TMC Navy Coordinating Office
Bureau of Naval Weapons
Department of the Navy
ATTN: ATMC
Washington 25, D. C. (1)

TECHNICAL SERVICES

Corps of Engineers

Director
U. S. Army Engineer Research & Development Laboratories
ATTN: Technical Documents Center
Fort Belvoir, Virginia (1)

Ordnance Corps

Commanding General
Ordnance Tank Automotive Command
ATTN: ORDMC-REO
Detroit Arsenal
Center Line, Michigan (1)

Deputy President
The Ordnance Board
ATTN: Library
Aberdeen Proving Ground, Maryland (1)

Quartermaster Corps

President
Quartermaster Board, U. S. Army
Fort Lee, Virginia (2)

Commanding General
Quartermaster Research and Engineering Command, U. S. Army
ATTN: Technical Library
Natick, Massachusetts (1)

Commanding Officer
Quartermaster Field Evaluation Agency
ATTN: Technical Library
Quartermaster Research & Engineering Command
Fort Lee, Virginia (1)

Commandant
Quartermaster School, U. S. Army
ATTN: Chief, Quartermaster Library
Fort Lee, Virginia (1)

Signal Corps

President
U. S. Army Signal Board
Fort Monmouth, New Jersey (1)

Transportation Corps

Chief of Transportation
ATTN: TCDRD (2)
ATTN: TCMNT (1)
ATTN: TCTER (2)
ATTN: TCREG (1)
ATTN: TCSOP (1)
ATTN: TCCAD (1)
Department of the Army
Washington 25, D. C.

President
U. S. Army Transportation Board
Fort Eustis, Virginia (1)

Commanding Officer
U. S. Army Transportation Combat Development Group
Fort Eustis, Virginia (1)

Commanding General
U. S. Army Transportation Materiel Command
ATTN: TCMAC-APU (16)
ATTN: Deputy for Surface Engineering (1)
P. O. Box 209, Main Office
St. Louis 66, Missouri

Commandant
U. S. Army Transportation School
ATTN: Adjutant
Fort Eustis, Virginia (4)

Commanding Officer (1)
U. S. Army Transportation Research Command
ATTN: Executive for Programs (2)
ATTN: Research Reference Center (14)
ATTN: Transportation Engineering Directorate (1)
ATTN: Aviation Directorate (1)
ATTN: Research Directorate (5)
ATTN: Long Range Technical Forecast Office (1)
ATTN: Military Liaison & Advisory Office (4)
Fort Eustis, Virginia

Commanding Officer
USA Transportation Research Command Liaison Office
ATTN: MCLATS
Wright-Patterson AFB, Ohio (1)

Transportation Corps Liaison Officer
U. S. Army Engineer Research and Development Laboratories
Building 314, Room A-216
Fort Belvoir, Virginia (1)

U. S. Army Research and Development Liaison Group (9851 DU)
ATTN: USATRECOM Liaison Officer
APO 757, New York, New York (1)

U. S. Army Engineer Waterways Experiment Station
ATTN: U. S. Army Transportation Corps Liaison Officer
Vicksburg, Mississippi (1)

U. S. Army Transportation Research Command Liaison Officer
U. S. Army Airborne, Electronics, and Special Warfare Board
Fort Bragg, North Carolina (1)

U. S. Army Transportation Research Command Liaison Officer
Detroit Arsenal
ATTN: ORDMX-LT
Center Line, Michigan (1)

Commanding General
U. S. Army Transportation Terminal Command, Atlantic
ATTN: Plans Office
Brooklyn 50, New York (1)

Commanding General
U. S. Army Transportation Terminal Command, Pacific
Fort Mason, California (3)

Chief
U. S. Army Transportation Materiel Command
Directorate of Cataloging and Standardization
Standardization Division
Fort Eustis, Virginia (2)

Chief
U. S. Army Research and Development Liaison Group (9851 DU)
APO 757, New York, New York (1)

OVERSEAS COMMANDS

Commanding General
U. S. Army, Europe (Rear)/Communications Zone
ATTN: Transportation Officer
APO 58, New York, New York (2)

Commanding Officer
Hq USATDS
ATTN: Tech. Procedures Office
APO 28, New York, New York (1)

Commander in Chief
United States Army, Pacific
ATTN: Transportation Officer
APO 958, San Francisco, California (1)

Commanding General
U. S. Army, Alaska
ATTN: ARATC
APO 949, Seattle, Washington (3)

Commanding General
U. S. Army, Ryukyu Islands/IX Corps
ATTN: Transportation Officer
APO 331, San Francisco, California (1)

Commanding General
US Army Caribbean
ATTN: Transportation Officer
Fort Amador, Canal Zone (1)

Commander, Allied Land Forces Southeastern Durope
ATTN: Chief, Transportation Branch G4 Division
APO 224, New York, New York (2)

UNITED STATES AIR FORCE

U. S. Army Representative
Hq AFSC (SCS-3) Air Force Systems Command
Andrews Air Force Base
Washington 25, D. C. (1)

Commander
Aeronautical Research Laboratories
ATTN: RRLA (Library)
Wright-Patterson AFB, Ohio (1)

UNITED STATES NAVY

Chief of Naval Operations
(Op-343)
Department of the Navy
Washington 25, D. C. (1)

Chief of Naval Research
Code 407M, Col. R. J. Oddy
Washington 25, D. C. (1)

Chief, Bureau of Ships
Department of the Navy
R&D Current Program Planning Branch (Code 333)
Washington 25, D. C. (3)

Chief, Bureau of Naval Weapons (R-38) (1)
Department of the Navy
ATTN: RA-4 (1)
ATTN: RRSY-15 (1)
ATTN: RRSY-2 (1)
ATTN: RRSY-5 (1)
Washington 25, D. C.

Asst. Chief for Research & Development (OW)
Bureau of Supplies and Accounts
Navy Department
Washington 25, D. C. (1)

Chief, Bureau of Yards and Docks
(Code D-400)
Department of the Navy
Washington 25, D. C. (1)

Commanding Officer and Director
U. S. Naval Civil Engineering Laboratory
Port Hueneme, California (1)

Officer in Charge
U. S. Naval Supply Research and Development Facility
ATTN: Library
Naval Supply Depot
Bayonne, New Jersey (1)

Library
Technical Reports Section
U. S. Naval Postgraduate School
Monterey, California (1)

Commanding Officer and Director
David Taylor Model Basin
Aerodynamics Laboratory Library
Washington 7, D. C. (1)

UNITED STATES MARINE CORPS

Director
MC Educational Center
Marine Corps Schools
Quantico, Virginia (2)

U. S. GOVERNMENT INDEPENDENT OFFICES

National Aviation Facilities Experimental Center
ATTN: Library
Atlantic City, New Jersey (10)

National Aeronautics and Space Administration
ATTN: Bertram A. Mulcahy
Assistant Director for Technical information
1520 H Street, N. W.
Washington 25, D. C. (6)

Librarian
Langley Research Center
National Aeronautics & Space Administration
Langley Field, Virginia (3)

Ames Research Center
National Aeronautics and Space Agency
ATTN: Library
Moffett Field, California (1)

Mr. Julian S. Hamilton
M-LOD-DR, Launch Operations Directorate
National Aeronautics and Space Administration
George C. Marshall Space Flight Center
Huntsville, Alabama (4)

National Aeronautics & Space Administration
Lewis Research Center
ATTN: Library
21000 Brookpark Road
Cleveland 35, Ohio (1)

Exchange and Gift Division
Library of Congress
Washington 25, D. C. (2)

Major Thomas Benson
Office of Assistant Director (Army Reactors)
Division of Reactor Development, USAEC
Washington 25, D. C. (1)

MISCELLANEOUS

Commander
Armed Services Technical Information Agency
ATTN: TIPCR
Arlington Hall Station
Arlington 12, Virginia (10)

Head, Geography Branch
Office of Naval Research
Department of the Navy
Washington 25, D. C. (2)

UNCLASSIFIED

Transportation Center, Northwestern Univ., Evanston, Ill.
THE STRUCTURE OF TRANSPORTATION NETWORKS - W. L. Garrison and D. F. Marble

31 May 62, 100 pp - illus - tables
Contract DA 44-177-TC-685, Transportation Forecast and Prediction Study, Unclassified Progress Report

Progress is reported on basic research on forecast and prediction problems. Descriptive indices of

1. Transportation network structure
2. Contract DA 44-177-TC-685

UNCLASSIFIED

Transportation Center, Northwestern Univ., Evanston, Ill.
THE STRUCTURE OF TRANSPORTATION NETWORKS - W. L. Garrison and D. F. Marble

31 May 62, 100 pp - illus - tables
Contract DA 44-177-TC-685, Transportation Forecast and Prediction Study, Unclassified Progress Report

Progress is reported on basic research on forecast and prediction problems. Descriptive indices of

1. Transportation network structure
2. Contract DA 44-177-TC-685

UNCLASSIFIED

Transportation Center, Northwestern Univ., Evanston, Ill.
THE STRUCTURE OF TRANSPORTATION NETWORKS - W. L. Garrison and D. F. Marble

31 May 62, 100 pp - illus - tables
Contract DA 44-177-TC-685, Transportation Forecast and Prediction Study, Unclassified Progress Report

Progress is reported on basic research on forecast and prediction problems. Descriptive indices of

1. Transportation network structure
2. Contract DA 44-177-TC-685

UNCLASSIFIED

Transportation Center, Northwestern Univ., Evanston, Ill.
THE STRUCTURE OF TRANSPORTATION NETWORKS - W. L. Garrison and D. F. Marble

31 May 62, 100 pp - illus - tables
Contract DA 44-177-TC-685, Transportation Forecast and Prediction Study, Unclassified Progress Report

Progress is reported on basic research on forecast and prediction problems. Descriptive indices of

1. Transportation network structure
2. Contract DA 44-177-TC-685

network structure are presented using concepts from the theory of graphs. Studies discussed are: 1) regression analyses of nation-to-nation variations of indices in relation to levels of economic and technical development, 2) components analyses of properties of intra-nation network structures, and 3) simulation of network expansion. The indices provide effective ways to measure network properties. Other findings are related to the over-all research task, which is discussed in the foreword.

network structure are presented using concepts from the theory of graphs. Studies discussed are: 1) regression analyses of nation-to-nation variations of indices in relation to levels of economic and technical development, 2) components analyses of properties of intra-nation network structures, and 3) simulation of network expansion. The indices provide effective ways to measure network properties. Other findings are related to the over-all research task, which is discussed in the foreword.

network structure are presented using concepts from the theory of graphs. Studies discussed are: 1) regression analyses of nation-to-nation variations of indices in relation to levels of economic and technical development, 2) components analyses of properties of intra-nation network structures, and 3) simulation of network expansion. The indices provide effective ways to measure network properties. Other findings are related to the over-all research task, which is discussed in the foreword.

network structure are presented using concepts from the theory of graphs. Studies discussed are: 1) regression analyses of nation-to-nation variations of indices in relation to levels of economic and technical development, 2) components analyses of properties of intra-nation network structures, and 3) simulation of network expansion. The indices provide effective ways to measure network properties. Other findings are related to the over-all research task, which is discussed in the foreword.

UNCLASSIFIED

Transportation Center, Northwestern Univ., Evanston, Ill.
THE STRUCTURE OF TRANSPORTATION NETWORKS - W. L. Garrison and D. F. Marble

31 May 62, 100 pp - illus - tables
Contract DA 44-177-TC-685, Transportation Forecast and Prediction Study, Unclassified Progress Report

Progress is reported on basic research on forecast and prediction problems. Descriptive indices of

1. Transportation network structure
2. Contract DA 44-177-TC-685

UNCLASSIFIED

Transportation Center, Northwestern Univ., Evanston, Ill.
THE STRUCTURE OF TRANSPORTATION NETWORKS - W. L. Garrison and D. F. Marble

31 May 62, 100 pp - illus - tables
Contract DA 44-177-TC-685, Transportation Forecast and Prediction Study, Unclassified Progress Report

Progress is reported on basic research on forecast and prediction problems. Descriptive indices of

1. Transportation network structure
2. Contract DA 44-177-TC-685

UNCLASSIFIED

Transportation Center, Northwestern Univ., Evanston, Ill.
THE STRUCTURE OF TRANSPORTATION NETWORKS - W. L. Garrison and D. F. Marble

31 May 62, 100 pp - illus - tables
Contract DA 44-177-TC-685, Transportation Forecast and Prediction Study, Unclassified Progress Report

Progress is reported on basic research on forecast and prediction problems. Descriptive indices of

1. Transportation network structure
2. Contract DA 44-177-TC-685

UNCLASSIFIED

Transportation Center, Northwestern Univ., Evanston, Ill.
THE STRUCTURE OF TRANSPORTATION NETWORKS - W. L. Garrison and D. F. Marble

31 May 62, 100 pp - illus - tables
Contract DA 44-177-TC-685, Transportation Forecast and Prediction Study, Unclassified Progress Report

Progress is reported on basic research on forecast and prediction problems. Descriptive indices of

1. Transportation network structure
2. Contract DA 44-177-TC-685

network structure are presented using concepts from the theory of graphs. Studies discussed are: 1) regression analyses of nation-to-nation variations of indices in relation to levels of economic and technical development, 2) components analyses of properties of intra-nation network structures, and 3) simulation of network expansion. The indices provide effective ways to measure network properties. Other findings are related to the over-all research task, which is discussed in the foreword.

network structure are presented using concepts from the theory of graphs. Studies discussed are: 1) regression analyses of nation-to-nation variations of indices in relation to levels of economic and technical development, 2) components analyses of properties of intra-nation network structures, and 3) simulation of network expansion. The indices provide effective ways to measure network properties. Other findings are related to the over-all research task, which is discussed in the foreword.

network structure are presented using concepts from the theory of graphs. Studies discussed are: 1) regression analyses of nation-to-nation variations of indices in relation to levels of economic and technical development, 2) components analyses of properties of intra-nation network structures, and 3) simulation of network expansion. The indices provide effective ways to measure network properties. Other findings are related to the over-all research task, which is discussed in the foreword.

network structure are presented using concepts from the theory of graphs. Studies discussed are: 1) regression analyses of nation-to-nation variations of indices in relation to levels of economic and technical development, 2) components analyses of properties of intra-nation network structures, and 3) simulation of network expansion. The indices provide effective ways to measure network properties. Other findings are related to the over-all research task, which is discussed in the foreword.