Comprehensive Study of Provincial Commodity Transportation

A. S. Narasimha Murthy

A Thesis

in

The Department

of

Civil Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montréal, Quebéc, Canada

April 1988

ISBN 0-315-41647-5

ABSTRACT

Comprehensive Study of Provincial Commodity Transportation

A. S .Narasimha Murthy, Ph.D.
Concordia University, 1987

This thesis makes an attempt to develop an objective, comprehensive and statistically credible method in the development of provincial commodity distribution and modal split models. The first stage of the thesis deals with the commodity distribution models based on the commodity flow survey data. The commodity classifications are grouped into manageable commodity categories. A set of Origin-Destination (O-D) tables, and an optimized gravity model for each commodity category is developed. The optimized gravity model for each commodity category has a unique spatial separation factor. This factor is determined by using a power function and least square methods. The impedance or friction factor is based on trip distance between origin and destinations taking into account commodity category. The results of the calibration are tested using standard statistical tests such as coefficient of determination, coefficient of correlation, and using commodity haul frequency diagram. The results are compared with previous calibration efforts of this nature. The results of this modeling show that the developed models are acceptable. The calibration of the models using the optimization technique has potential for further applications in commodity transportation planning.

The second stage of this thesis deals with modal split analysis and modeling across the province. First efforts are made to recognize the pattern of commodity movements across Alberta. Then the modal split models are constructed using

log-linear and logit models. The main factors used in modal split analysis and modeling are the type of commodity, average shipment size, total shipment size, type of load (full load or less than full load), type of hire (private or for-hire), type of control (yes or no) and, distance (kilometers) between origin and destination.

Under modal split modeling, log-linear models are constructed in order to uncover associations among the variables and to choose a suitable unsaturated model to build logit models. A secondary analysis is carried out by constructing a suitable logit model to obtain the table of log-odds. A logit model is constructed so as to understand the effect of various factors on a designated response variable. In the analysis, the designated response variable is the mode of transport. The method suggested is a useful tool in analysing the commodity data required in commodity transportation planning.

In the final stage of the thesis, a discussion on the practical applications of the developed distribution and modal split models is given. The proposed commodity distribution and modal split methodology has implications for improving commodity transportation planning across the province.

To My Family

Especially to my MOTHER and Late FATHER

ACKNOWLEDGEMENTS

I wish to express my gratitude and deep appreciation to Dr.B.Ashtakala for supervising and guiding me through all stages of this thesis. I am indebted to him for providing financial assistance for my doctroral studies at Concordia University, through his Natural Science and Engineering Research Council of Canada (NSERC) operating grants.

I am thankful to the Chairman Dr.M.M.Douglass and other thesis review committee members. I thank Mrs. Anne Barkman, computer analyst for her interest and assistance. This thesis could not have been completed without the help of computer facilities provided by the Faculty of Engineering and Computer Science.Also, I am thankful to the library services at Concordia University, special thanks to Ms.Gail Flicker, Borrowing Assistant at Inter Loan Library for her efforts to find books and research materials for this study.

I extend my thanks to the Government of Alberta Transportation Department for allowingthe use of its " Alberta Commodity Transport Survey, 1977" data for this research project.

Finally, I would like to convey my great appreciation and love to my family members to whom this thesis is dedicated. I thank my wife for proof reading the manuscripts.

TABLE OF CONTENTS

CHAPTER 3 : DATA BASE

CHAPTER 4 : COMMODITY DISTRIBUTION MODEL

CHAPTER 5 : MODAL SPLIT ANALYSIS AND MODELS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

Symbol		Particulars
A_i	=	Gravity model constant;
C	=	Control type (yes or no);
C_j	=	Consumption (tons) at destinations (sink),tons;
C_j^*	=	Consumption value at the end of iterations;
CFT	=	Commodity Flow Table;
CHFD	=	Commodity Haul Frequency Diagram;
d_{ij}	=	Distance between origin and destinations, km;
D_{ij}	=	Observed interchanges, tons (origin & destinations);
FL	=	Full load;
L	=	Load type (full load, or less than full load);
H	=	Hire type (private or for-hire);
k	=	Specified range (distribution model);
K	=	Commodity classifications (logit model);
km.	=	Kilometers;
ℓ	=	Optimum condition (distribution model);
G^2	=	Likelihood ratio statistic;
LTFL	=	Less than full load;
LSM	=	Least Square Method
m	=	Commodity category (gravity models);
$\hat{m}_{ij}$	=	Fitted cell counts (expected);

O-D	=	Origin and Destinations;
P_i	=	Productions (tons) at origin (source), tons;
P_{ij}	=	Proportion (%);
r	=	Correlation coefficient;
R^2	=	Coefficient of Determination;
RDC	=	Regional Distribution Centre;
S	=	Average shipment size;
T_{ij}	=	Model interchanges in tons;
W	=	Constant (logit model);
x_{ij}	=	observed cell counts;
X_{++}	=	Total observations;
Z	=	Arbitrary function;
α	=	Level of significance;
λ	=	Power parameter;
μ	=	μ, terms (log-linear model);
Σ	=	Summation;
χ^2	=	Chi Square;
∞	=	Infinity; and
Θ	=	Odds Ratio.

CHAPTER 1 : INTRODUCTION

CHAPTER 1

INTRODUCTION

1.1. COMMODITY TRANSPORTATION: A GENERAL FRAME WORK

Commodity transportation represents an important sector within any regional or national economy. A transport model which forecasts the region to region commodity flow pattern given the regional levels of production and consumption is very important for long range planning and in improving transportation efficiency. The transportation network at any given place has implications for both private and public agencies. The cost at which a product is available to the consumer may very well depend on the existing transportation network. The knowledge of commodity transportation at the regional and national level will assist in improving the network for future demands and making suitable policy decisions which relate to the system. The principal components of commodity transportation are, the source (point of production), the sink (point of consumption) and the system (modes used). The inter-relationships are studied in order to develop a descriptive or predictive modeling technique. The appropriate relationships are entirely dependent on the size of sources and sinks including locations, the available modes of transportation between them and the specific commodity under consideration. Fig.1a shows the basic commodity transportation system.

The transportation of raw materials to the plant and later transporting the finished goods to the market place at an economical cost in a timely, undamaged condition requires an efficient and reliable transportation network(43,44). It is currently noted that a significant percentage of current traffic on highways is due to commodity movements between cities. The travel time value of commodity delayed is tied to the market value of the commodity(2). Compared to urban transportation planning, commodity transportation stands in sharp contrast in many aspects. For instance, in urban transportation there is only one client (government) and one commodity (people), however in case of commodity transportation planning, the complexities are enormous when we consider the different types of commodities(35).

In provincial commodity transportation planning, the two most important questions that frequently arise are the following: (1). What is the commodity distribution pattern across the province? (2). What is the type of mode used in distributing these commodities? Given the largely private character of commodity transport, it is necessary to obtain satisfactory answers to the above questions so as to have an efficient transportation network. In commodity transportation planning the emphasis should be on early identification of the problem followed by a feasible solution. For example, it is necessary to know the present level of commodity movements in order to make decisions on investment in facilities or equipment, taxation policies and regulations governing future demands by the transportation industry.

The study of commodity movements within the province is a study of a closed economy with greater emphasis on regional economy. It involves the identification of regional economic sectors which are either well served or being neglected; and in understanding the type and cost of services being offered. The provincial transportation departments and other agencies have come to understand that their responsibilities are not solely confined to the provincial highway system, but include the broad spectrum of commodity transport services being provided by numerous private agencies and the different modes used for this purpose. The attention to this area is mainly led by the abandoning of light density railway lines and a surge in the cost of highway maintenance by the use of heavier trucks.

There is a need for developing systematic and statistically credible methods by which one can analyze commodity flow across the province. In most provinces, future truck volumes are forecast as a percentage of aggregate traffic volume for both existing and proposed facilities. The major problem stems from the lack of data base rather than from an inability to devise suitable distribution models for forecasting. In case of modal split analysis, there is a need to identify the modes that dominate in transporting commodities across the province. The important question to be answered under transportation policy by provincial government could be the nature of subsidy which should be distributed to retain a balanced service. Urban transportation planning experience is helpful, but cannot be considered as a direct substitute in solving commodity transportation planning problems(35).

There are a number of factors that restrict the development and use of a suitable mathematical model to understand the movements of commodities across a province. The first major issue are the data limitations in the commodity sector. The lack of data containing commodity movements is crucial to developing a demand forecasting models. The cost of collecting necessary commodity data is enormous and requires a large staff. Commodity transportation planning is an extremely broad field and it is impossible to have all required data collected. Thus, investigators are forced to make several assumptions and adjustments in using available data; in many cases, they are constrained to put toghether useful aggregate data while developing models. The shipper or consignee survey required to validate the data requires a large costs and manpower. One of the greatest needs at present is to improve collection of data on a regular basis, particularly the commodity flow information on medium and long-distance.

The second limitation arises from the inability to forecast economic activity. Many planning decisions require demand forecasts that rather involve comparisons between proposed alternatives and existing conditions over a short period due to the unpredictable economic activity over longer periods. One prime example of unpredictable economic activity is the recent introduction of deregulation in some areas of ground transportation. Deregulation has brought both advantages and disadvantages to both consumers and industry.

The third limitation is to deal with the problem to users of these models. Models have been developed by researchers from many different fields with many different objectives. This wide range of approach leaves the user of these models with a problem; the user must be careful not to try and apply the model for situations where it is not suited and has specific limitations. Thus a well structured model is required with complete description of data source in order to understand commodity transportation planning. It is also required to develop comprehensive methodology that is applicable to the needs of commodity transportation planning.

1.3. OBJECTIVES AND SCOPE OF STUDY

1.3.1. Introduction

This thesis deals with transportation demand modeling in the commodity movement sector. Recently the field work involved in collecting commodity flow data has been intensified and provinces have become increasingly aware of the issues relating to commodity movements on the provincial highway systems. In commodity transportation planning, the objective is for early detection and resolution of problems with respect to highway network.

In transportation demand modeling, there are four major stages as shown in Fig. 1b are usually used. They are trip generation, trip distribution, modal split, and trip assignment. The purpose of each stage is defined briefly as follows:

1. **Trip Generation**: involves the determination of the quantity of trip ends associated within a zone or zones.
2. **Trip Distribution**: is the determination of the trip interchanges in the base year of a given number of trips between zones. This stage involves the use of mathematical formulation and calibration process.
3. **Modal Split**: the division of trips between competing modes used for transportation.
4. **Traffic Assignment**: the allocation of interchanges among available routes between the zones.

This study is concerned with the trip distribution and modal split stages of commodity transportation planning. The data collection done by Alberta Transportation comes under trip generation, and hence that stage was not considered, except for checking the data and using it for further analysis.

1.3.2. OBJECTIVES

The following are the main objectives of the present study:

1. To develop comprehensive and statistically credible methods and models required for commodity transportation planning.

2. To develop an intra-provincial commodity distribution model, by suitably modifying the standard gravity model.

3. To Formulate a methodology to analyze the vast commodity flow data and develop modal split models that distribute the commodity flows among competing modes.

4. To review of the data requirements and application of models for commodity transportation planning.

1.3.3. SCOPE OF STUDY

In province wide commodity transportation planning, the two questions that frequently arise is the type of commodity distribution and the mode(s) used for distribution. It is essential in commodity transportation planning to understand existing conditions before making any decisions. The knowledge of commodity transportation is essential because it contributes considerably to the gross domestic product of a place from the service sector both at a national and provincial level. The cost at which a product is available to the consumer depends on efficient and reliable transportation planning and the operation of the network.

The present study deals with the distribution and modal split analysis of the commodity movements across the Province of Alberta. It takes into consideration only the intra-provincial commodity flow survey data collected by Alberta Transportation in 1977. The actual tonnage that is distributed between certain selected communities across the province is only considered as these are expected to represent the movement of commodities across the entire province. The details of the commodity flow survey and the selected communities is discussed later in the text. The modal split analysis is studied with respect to the two major modes of transportation in the province, namely truck and rail. The modal split analysis and modeling is based on several factors, some of which have not been used in earlier studies of this nature. Some of the important factors considered are the commodity type, type of loads (full load or less than full load), type of hire (private or for-hire), control (yes or no), average shipement size, total tonnage, types of modes, etc. The logit models are used to build modal-split models. This study is invariably limited by the non-availibility of certain commodity flow data.

1.4 ORGANIZATION

Chapter 2 presents an overview of the commodity distribution and modal split models in existence. It provides a background on the commodity distribution and modal split models considered as important for the present study during the literature survey. This section examines the differences between various models used for commodity distribution and modal split by previous researchers. Under commodity distribution models, some of the major commodity flow studies are discussed (9,11,20,21,26,45). In the brief review of modal split models; the regression model, abstract mode model, discriminant analysis and, probit and logit models are considered(39,40,42).

The data base is discussed in Chapter 3. It provides a comprehensive view of the way in which the data were developed for the present study. It describes the commodity flow survey conducted by Alberta Transportation in 1977 to obtain commodity flow data across the province. The sub-sections in this chapter deal with the community selection, shipper or consignee interview and classification of commodities. A description of the community characteristics is also included in this section.

The theoretical framework of the commodity distribution model developed by this study is given in Chapter 4. It provides a consistent linkage between data structure and the model building process. The standard gravity model is modifyed to distribute the commodities across the province. This chapter also deals with the calibration and analysis of the developed model.

Chapter 5 examines into the modal split across the province and develops modal-split models for the commodity movements. First a detailed analysis of the commodity movements (pattern recognition) is carried out based on some important factors such as average shipment size, type of loads, etc. Secondly, log-linear and logit models are used to construct modal split models.

Chapter 6 gives an example of the practical application of the developed models in commodity transportation planning. The final chapter presents the summary and conclusions of the present study. Also, suggestions are made for future research in this area. Under conclusions some of the important findings of this study are indicated. The appendix contains coding instructions to shippers and consignees, sample survey data output format, instructions to computer programs, and program listings.

CHAPTER 2: COMMODITY DISTRIBUTION AND MODAL SPLIT MODELS

CHAPTER 2

COMMODITY DISTRIBUTION AND MODAL SPLIT MODELS

2.1. INTRODUCTION

Commodity transportation has been growing at a rapid pace over the years. It is worthwhile to record the importance and impact that the transportation sector has on the economy. Research regarding commodity movements began to develop rapidly due to the large movement of commodities crossing borders, particularly the study of national and international movements as shown in Fig.2. The one important factor that contributes to the movement, is the disparity in resources that exists between regions. People of different regions need different products that suit their living conditions, and are required to exchange goods for mutual advantage. Each place will have some kind of economic relationship with the other place by the flow of goods between them. This commodity flow could be achieved by any available mode of transportation. In order to understand this complex nature of commodity flows, various theories and models related to commodity transportation planning were developed.

An extensive search of pertinent literature on commodity distribution and modal split analysis uncovered information on analysis and thetheoretical framework for modeling. The first and most comprehensive commodity flow analysis was conducted by E.L.Ullman(45), using the 1950 commodity data for United States of America. Recent developments on commodity flow patterns have endeavoured to integrate the theoretical developments of the gravity formula into .ecasting models. In the following sections a few important commodity distribution models based on the gravity model and some modal split models are discussed. The advantages and disadvantages are also assessed where possible.

The commodity distribution models reviewed under section 2.2 are; the gravity model and commodity flow studies by E.L.Ullman, W.R.Black, B.G.Hutchinson, and W.Leontief. The modal split models reviewed under section 2.3 are: regression analysis, abstract mode model, discriminant analysis and, logit and probit models.

2.2. COMMODITY DISTRIBUTION MODELS

To enable transportation planners to meet the anticipated capacity requirements, it is necessary to have some explanation of the demand for transport and estimate the growth in the consumption for commodities over long-term. In developing models it is assumed that the models can provide reasonable estimates of the movement of commodities and future consumption. Though the success of urban transportation planning and forecasting is well-known, its success when applied to commodity distribution is rather limited(21).

The most widely used distribution model is the gravity model. This model was originally derived by analogy from the famous Newton's Law of Gravitation. The law of gravitation states that the attraction between two bodies is directly proportional to the mass of the bodies and inversely proportional to the square of the distance between the two bodies. However, this law is adopted for transportation by substituting trip production and trip attraction for the mass, and the distance is replaced by travel time(21). The proponents of this model pointed to the apparent analogy between movements of people from a lower economic zone to a higher economic zone. Earlier researchers believed that just as the interaction of two bodies is governed by Newton's law either due to force, energy or potential, so must there exist similar relationship for the movement of people or commodities between zones. Hence this emperical formula for trip distribution has become a corner stone and was appropriately named gravity model. The gravity model has been extensively used by both urban transportation planners and commodity transportation planners through the years to estimate the interchanges between zones(9,11,20,21,26,45,47,48).

The Gravity Model is expressed mathematically as follows;

$$T_{ij} = P_i \frac{C_j F_{ij}}{\Sigma_{j=1}^{n} C_j F_{ij}} \qquad (2.1)$$

Where,

T_{ij}	=	trip production from zone i to zone j
P_i	=	production in zone i
C_j	=	attraction in zone j
F_{ij}	=	$f(t_{ij})$
t_{ij}	=	impedance, which is usually the distance, time or cost.

Rigorous theoretical development was necessary in order to supress apprehension about the use of gravity model in transportation planning. Particularly it was appealing to study the migration of population and passenger movements. The only emparical support for gravity model comes from the observation that the interchanges between two regions is clearly a decending order of the distance between the origin and destination. Several studies dealing with intra-urban movement patterns made use of a gravity type distribution model(21).

Theoretical basis for the gravity has received some attention and this has increased the confidence in its use as a trip distribution model. In recent years there has been two important studies relating to the theoretical development of the gravity model for transportation studies. The first is the entropy maximizing methodology, based on an analogy with micro-canonical ensemble technique in statistical mechanics by Wilson(48). Wilson in the above study has proved that the constrained gravity model can be derived direclty from probability or entropy maximizing priciples and the use of statistical analogy. The results of Wilson study improves the justification in using this model in transportation studies.

The second theoretical support is put forward by Niedercorn and Bechdolt Jr. based on economic derivation(utility theory) of the gravity model(36). Their contribution has opened ways to link economic theory to the emperically formulated gravity model. There has been some criticism of their derivation and it is to do with the basic hypothesis made in the study. The hypothesis states tht the net utility is a function of the number of trips from a single origin to numerous destinations. Apart form the criticism and doubts the Niedercorn and Bechdolt derivation, their work ties the gravity model to the economic frame work of utility theory is most significant and proivedes necessary theoretical support to the gravity model.

Probability approach by Walter Isard(22) was the first attempt in the development of gravity model without relying on Newton's law. His model development is based on probability concepts. He divided an metropolitan area population P into several subareas depending on the population characteritsics based on age, occupation, sex, etc. If T represents the total number of internal trips in the metropolitan area, he tried to estimate trips T_{ij} beteen subarea i and subarea j assuming that neither cost or time are involved in undertaking the trip(i.e the friction factor is zero). Next, he postulated that for an individual in subarea i, the fraction of his journeys terminating in subarea j will be equal to P_j divided by P. Where P_j is the subarea j's population. Also, since the individual in subarea i is identical to an individual in any other subarea, and since transportation costs are zero, he estimated the total number of trips that the individual undertakes as the average number of trips per capita for the entire metropolitan area, or T/P. He replaced this quotient by the letter K. thus the absolute number of trips that an individual from subarea i makes to subarea j, is $K(P_j/P)$.Since there are P_i individuals in subarea i, then :

$$T_{ij} = K \frac{P_i . P_j}{P} \qquad \text{.................................} \quad (2.2)$$

He proposed that in a similar manner, the total expected number of trips for every possible combination of i's and j's can be estimated. Eq. 2.2 resembles the formulation of the gravity model.

The first and most comprehensive commodity flow analysis to date was conducted by E.L.Ullman(45). He used the 1950 commodity flow survey data of United States of America to draw flow diagrams depicting origins and destinations of several commodity groupings. His work was the first break-through for researchers. The data base he produced was useful for other researchers to test their theories and models.

Black(9) also attempted to study the commodity flow within United States of America. He used 24 commodity groupings and gravity model for distribution. He used consumption costrained model for the commodity distribution. The model formulation of Black was:

$$T_{ij} = \frac{O_i^k D_j^k F_{ij}^k}{\Sigma_{j=1}^n D_j^k F_{ij}^k} \quad \text{.............................. (2.3)}$$

Where,

T_{ij}^k - total tons of commodity k produced in region i and shipped to region j,

O_i^k - total shipments of commodity k from region i,

D_j^k - total demand for commodity k in region j, and

F_{ij}^k - friction factor equal to $1/d_{ij}^{\alpha}$, where d_{ij} ~~is~~ the straight line distance between region i and region j and α is an emperically derived exponent which may vary depending on the commodity group under consideration.

In estimating the value of the exponent, he started with a zero exponent with incerements of 0.025 at each iteration until the correlation between the actual and the estimated flows fail to increase. Also, he assumed unimodal over the entire region of search. He hypothesized the follwing for his study:

a. The greater the specialization for a partiuclar commodity, the lower the exponent value.

b. The greater the intra-regional consumption, the greater the exponent value.

c. The greater the per unit value of a commodity, the larger the exponent value.

The exponent values obtained by Black are discussed in distibution model Chapter 5. The model does perform slightly better for highly aggregated data, which substantiates the view of many researchers about the reliability of gravity model with aggregate data. Finaly, he states that; the efficency of the gravity model is quite stable and is not seriously affected by disaggregation of flows. The model does perform better with highly aggregated data, which tends to be consistent with the appication of the model in other research areas.

Hutchinson(20) attempted to forecast commodity flow in 10 regions of Ontario province using published data. To estimate the regional production he utilized the annual census of manufacturers. Where data was not available, they were estimated from labor force statistics and the average productivity rates. The consumption was obtained from the product of the regional average per capita consumption of each commodity type and regional population. The distribution model expressed the importance of transportation costs in distribution. Hutchinson's proposed model was ;

$$T_{ij}^{k} = P_{i}^{k} \, . \, D_{ij}^{k} \qquad \text{..................} \quad (2.4)$$

where, T_{i}^{k} - a square matrix of the inter-regional commodity flows by commodity type.

P_{i}^{k} - a diagonal matrix of the production of commodity k in region i, and

D_{ij}^{k} - a square matrix, equal to $\dfrac{A_{j}^{k}\ F_{ij}^{k}}{\Sigma_{j}\ A_{j}^{k}\ F_{ij}^{k}}$

A_{j}^{k} - cash value of the annual cosnumption of commodity k in region j, and

F_{ij}^{k} - a function of the transport costs involved in moving commodity k between regions i and j.

The proposed gravity model for the inter-regional commodity flow above is similar to one in Eq.2.1with impedence related to transport costs.

Leontief(26) proposed that the flow of any particular good 'k' from region 'i' to any other region 'j' is assumed to be directly proportional to its output in region 'i', to its total input to the aggregate amount of commodity 'k' produced and consumed in all the regions of the economy as a whole. The model is written as follows:

$$T_{ij}^{k} = \frac{O_i^k D_j^k}{T^k} K_{ij}^{k} \qquad \text{.......... (2.5)}$$

where,

O_i^k - total output of commodity k in region i

D_j^k - total input of commodity k in region j

T^k - total output and input of commodity k in all regions.

K_{ij}^k - emparically derived constant of proportionality.

The model resembles the gravity model, with impedance function being a relationship between total input and output instead of distance. This model resembles the concept of supply and demand in the inter-regional movements. The developed distribution model is discussed under Chapter 4.

2.3. MODAL SPLIT MODELS

The objective of modal split analysis and modeling is to allocate commodity movements among competing modes. The identification of mode choice is an important step in any commodity transportation planning process. The decision process in commodity transportation are complicated. For example, a firm should consider long-range decisions of plant location, technology of production etc., to the more short-run decisions involving available modes of transport and shipment size to be considered for raw materials and finished products.

From several studies(10,12,18,23,24) it has been shown that shipment weight and transportation cost are more significant factors in shipper selection of a mode than the length of haul. Some researchers have considered quality of service as an important factor in mode selection. In general the quality of service is measured by safety, reliability and transit time(22).

Meyer et al (22) in 1959 were the first to attempt to develop an analytical modal choice model. The model selected used the lowest cost criteria in selecting the mode depending on length of haul. The following gives a combined list of the various important factors considered by different researchers in explaining the mode selection and in model building(39).

1. Commodity Type,
2. Transit time,
3. Transit Distance,
4. Transportation Cost,
5. Weight of Shipment,
6. Size of Firm,
7. Frequency of Service,
8. Reliability of Service,
9. Measure of Safety,
10. Volume - demand limit time,
11. Handling cost, terminal cost, line haul cost,
12. Completion of Service,

13. Fragility, Perishability, Adaptability and Seasonality,
14. Value of goods, and
15. Density of goods.

The following are certain requirements and problems encountered during modal split analysis and modeling;

1) the availability of reilable data,
2) the proper categorizing of different commodities into manageable size, and
3) the difficulty of recognizing and including all parameters in analysis.

The following describes few of the modal split models selected for research. Each model has its own, general formulation, application, data requirements, and evaluation for use in province wide transportation system applications.

Simple and multiple regression methods have been employed extensively to examine modal choice. Due to the numerous variables that contribute in the selection of a mode multiple regression has been used more frequently. Surthi and Ebrahim(31) used the census of transportation data to develop a mode selection model. The general form of the multiple-regression equation, where there are n independent variables is written as;

$$P_k\,(x_{ij}) = a + b_1x_1 + b_2x_2 + b_3x_3 + \ldots\ldots + b_nx_n + e \qquad \ldots\ldots\ldots\ldots\ldots\ldots(2.6)$$

Where,

$P_k(x_{ij})$	=	estimated magnitude of the dependent variable; in this case the probability of selecting mode k for transporting goods between origin i and destination j;
x_n	=	independent variables;
b_n	=	parameters; and
e	=	error term.

Regression models are descriptive and attempt to explain the modal selection behaviour through mathematical relationships developed from a cross-sectional analysis of data. The methodology is mathematically based procedure utilizing statistical measures to determine the effects of independent variables have on modal shares of goods movement. Regression analysis seems ideally suited for determining estimates of modal choice probabilities, from which expected mode use can be calculated. The stepwise regression procedure is widely adopted in most research and the methodology has been computerized for an efficent and easy use(7).

The advantages of using the regression analysis is that it allows the modeler to understand the contribution of the selected independent variables in determining the modal share. The stepwise procedure allows the ranking of the independent variables. The methodology is totally computerized, simple and easy to interpret. The results are evaluated pricipally using standard statistical tests such as the F-test or t-value tests.

The disadvantage is one of unbouned probability by using this methodology, that is, the estimated conditional probability may produce values greater than 1 or less than 0 which is not acceptable. For this reason many researchers select logit or probit models(1). Also, this method requires large amount of data for reliable results.

The abstract mode model views demand for commodity along any link in a network and a particular (abstract) mode as dependent on the characteristics of that mode and on key commodity attributes. Abstract mode models are so called because the modal attribute variables used, can describe any particular mode. This approach handles the distribution and modal split elements as one task. Turner(42) noticed that the abstract mode technique was the method suited for explaining macro-level mode choice decisions under existing data availability limitations. He used the following variables for his model, origin and destination, distance shipped, transit time, size and commodity group of shipment and value. For Turner lack of information on reliability of service measures caused him some problems in the calibration of the model. The advantage of abstract mode model is its capacity to evaluate the sensitivity of the changes in mode, travel time or cost. This flexibility allows to evaluate the impacts of policies, such as revision of the commodity rate structure or provisions of capital investment subsidies for updating and improving a particular means of transportation.

The abstract mode concept addresses the problem of predicting the mode of transportation between specified origin and destination pairs by using disaggregate data. The study by Turner concluded the abstract mode technique is best suited for explaining macro-level mode choice decisions. Herendeen(19) proposed a modified version of this model as follows:

$$P_k(X_{ij}) = \alpha_0 R_k^{\alpha 1} C_k^{\alpha 2} T_k^{\alpha 3} F_k^{\alpha 4} \quad \text{..................} \quad (2.7)$$

with the constraint

$$P_k(X_{ij}) = 100 \quad \text{..................} (2.8)$$

$P_k(X_{ij})$ = percentage of X_{ijk} that will be shipped by mode k;

R_k = reliability of mode k;

C_k = relative cost of mode k = (relative cost of all available modes / rate of mode k)

T_k = relative transit time by mode =(lowest time by all modes / time by mode k)

F_k = relative frequency of service = (frequency by all modes / best frequency of service)

α_0 = coefficient to be determined by the regression analysis ; and

$\alpha_1, \alpha_2, \alpha_3, \alpha_4$ = exponents to be determined by the regression analysis.

Herendeen used this model for just two modes highway and rail freight, for five commodity groups all of whcih were manufactured goods, and three independent variables such as the relative costs, relative transit time, and reliability.

The disadvantage of this methodology is the lack of consideration given to the type of commodity type being transported. Turner has tried to add commodity characteristics to the model as dummy variables to help eliminate this problem. The effect of inter-modal are not taken into care since the model is formulated to consider each mode seperately.

Discriminant analysis methodology is a multivariate statistical methodology used to develop modal split models. The methodology is useful to classify a goods to one of several modes with a miniminum of error based on a set of specified characteristics of the modes and consignment. This is achieved by the use of a linear discriminant function that is comprised of a linear combination of explanatory variables(7,13).

Fig.3a shows graphically the intrinsic properties of the discriminant function method of estimating modal share. If the frequency distribution of the shipper's modal-choice characteristics are known for each mode of transportation, for a given consignment a disutility value, Z, can be computed and the probable mode of transportation can be predicted.

From Fig.3a it is clear that, if the computed value of Z is greater thatn Z', mode 2 is likely to be selected and vice versa. In the figure the two modes are seperated but with an overlapping region marked by shaded area. This shaded area indicates the possibility of misclassifications of consignments to either mode(34).

The general formulation of the discriminant function for a two mode choice can be written as(33,34):

$$Z_{ij} = [\, X - 1/2\, (\bar{X}_i - \bar{X}_j] \, S^{-1} \, (\bar{X}_i - \bar{X}_j) \quad \text{.........................(2.9)}$$

in which,

Z_{ij} = linear discriminant fucntion value;

X = the vector of characteristics of a given mode and/or commodity to be used to categorize a consignment as using mode i or mode j;

$\bar{X}_i$, $\bar{X}_j$ = the vector of mean characteristics of all shipments that use i and j, respectively;

S^{-1} = the inverse of the estimated variance co-variance matrix of X characteristics of a mode and/or commodity within each mode; and

[ ] = the transpose of the vector expansion within the brackets.

If there are only two modes i and j, the discriminant function can classify a consignment according to the value of $Z_{ij} > 0$, the consignment as a user of mode i, otherwise, the consignment is classified as a user of mode j. Two very important assumptions must be made in using linear discriminant method; 1. that the distributions of the variables are multi-variant normal and 2. the variance-covariance matrices of the two sets of variables are the same. The methodology can be tested for its ability to classify correctly using a 'confusion matrix' and an F-statistic can also be used to test if the sample size is adequate.

Milkus(29) used linear discriminant function to estimate the allocation of traffic between rail and truck mode. He used shippment records for frozen vegetables from the 1963 census of transportation. The data requirements were shipment weight, distance, and employment size of the originating plant. Milkus concluded that discriminant function appear to be better multi-purpose estimator function than any of the freight rate functions(simple liner regression models) that he could develop. Sasaki(40) used this method to allocate freight between rail and water. He found that weight and distance were significant. He concluded that discriminant function does have the ability to classify correctly individual shipments by mode of transportation. The method appears to be a relevant tool for analysis of commodity movements by different modes and discriminant formulation is quite sensitive to data changes. In using this method it is necessary to have quality and sufficient data to calibrate and test. This method has been widely used by many researchers in modal split modeling.

The logit model is chosen for review because of its reasonable explanatory powers to predict the probability that the shipper will select a specified means of transportation from a set of alternative modes . Mode choice models of the logit form have been used to predict the truck-rail modal split as a function of modal and commodity characteristics using aggregate data by Kullman(24), and using disaggreagate data by Hartwing and Linton(18). Logit analysis is another multivariate statistical tool used mainly in biometrics and has been used in the field of disaggregate modal choice modeling. The use of this model is mainly due to its two improtant characteristics; the first is that the estimated probability of mode choice always lies within 0 and 1, this eliminates the unbounded probability inherent in using linear probability functions such as the regression procedure. The sencond characteristics is that the sum of the probabilities of each alternative modes will be equal to 1.

Fig.3b shows the logistic formulation for binary choice model(33). In the figure it is clear that the logistic model is a non-linear model that is able to fit a linear probability function to a logistic growth curve. The general formulation of the logistic model, in its multinomial form can be written as;

$$P_k(x) = \frac{e^{u(x)}}{1 + e^{u(x)}} \qquad \text{.................................} \quad (2.10)$$

where,

$P_k(x)$ = the probability of a shipper choosing mode k out of the full set of mode alternatives

$u(x) = x_0 + x_1 X_1 + \ldots\ldots + x_n X_n$

X_1, X_n = independent variables, expressed as differences, and

x_n = coefficients of X_n

therefore if $P_k(x)$ is the probabilty of choosing mode k, the probability of not choosing mode k is $1-P_k(x)$, in which

$$1-P_k(x) = 1/(1+e^{u(x)})$$

The logit transformation of the linear probability function is written as

$$\log_e \left(\frac{P_k(x)}{1-P_k(x)} \right) = x_0 + x_1 X_1 \; \ldots\ldots \; + x_n X_n \; .$$

in which LHS is called the logit.

The best method for estimating the values of the coefficients (x_0, x_1 ...x_n) one can use the methods of maximum likelihood estimation method.

Probit analysis, like logit analysis, has been used extensively in the field of biometrics. Probit analysis provides a means of assigning a linear relationship to a probability function. Modal separation is accomplished through probability functions that can be represented by a cummulative normal- distribution frequency curve. Identification of the shipper's threshold response level is one of the key concepts of the probit analysis. The S-curve shown in Fig.3c is usually referred to as a sigmoidal response curve. Identificatioin of the shippers response level is one of the key concepts of probit analysis. It represents the indifference level where a shipper's preference is equalized in relation to selecting alternative modes. The value of X_1 in Fig. 3c depicts a shipper threshold level, where any value greater than X_1 represents a shipper's preference of choice to mode k over an alternative mode. The threshold response levels are assumed to be normally distributed(33). Hartwing and Linton(18) concluded that the probit analysis form is applicable to disaggregate mode choice modeling. Their analysis showed that binary mode choice truck or rail is more sensitive to changes in transportation costs than either changes in transit time or reliability.

Researchers using the logit and probit analysis has to deal with the problem of disaggregate data availibility. Logit and probit analysis are suited for relating observed shipper behaviour to factors for determining mode choice between two means of transportation. The draw back with respect to probit model is that it requires a more time-consuming calibration procedure than most other models (1,34). Modal split analysis and models developed are discussedin Chapter 5.

CHAPTER 3: DATA BASE

CHAPTER 3
DATA BASE

3.1. BACKGROUND

Alberta is one of the ten provinces in Canada and has a population over two million and a land area of 250,000 square miles. Agriculture directly accounts for nearly 17% of Alberta's net value of production and 25% of personal income. Edmonton and Calgary are the two major cities of Alberta. The principal primary industries are grain, coal, oil, gas, sulphur, forest products, etc(43,44). In trying to understand the vast commodity distribution and modal split across the province, the Alberta Transportation Department instigated a major study in 1977. The study was limited to selected communities across the province. Fig.4 shows the map of Alberta with the important communities across the province.

The main objectives of the survey were; (1). to develop a first level understanding of commodity flows in Alberta. (2). to determine, the best method for a major comprehensive study of commodity transportation in Alberta.(3). to examine the possibilities of developing mathematical relationships describing commodity transportation, and (4). to develop appropriate generalizations for expansion or forecasting of commodity activity beyond the sample universe. The study was limited to (1). Principal primary industries such as grain, coal, oil, gas sulphur, forest products etc., with emphasis on resource outflows, and (2). the primary highway and railway network interconnecting major population zones.

3.2. COMMODITY FLOW SURVEY

3.2.1. Introduction

The commodity flow survey was assigned to a private consulting firm and one of the principal objectives of the survey was to develop commodity volume estimates. The commodity related information was obtained through informal interview techniques with both shippers and consignees using pre-forwarded survey forms as shown in Fig. 5. The survey questionnaire was developed by consulting the clients. To the extent possible, lines 1 and 2 of the survey form were pre-coded in the office prior to the surveyor visiting the community. The information for the first 2 lines was taken from existing telephone directories. Once in the field the surveyor was able to confirm the firm type and proceed to gather the data. The community file could be revised on these visits by deleting firms which might have gone out of business or changed owners or names and new businesses could be added to the file(43,44).

The survey took the form of interviewing consignees and shippers of goods in any significant quantities across the province. A significant amount of the interviews were conducted in person and the rest by telephone. The initial list os interviewees were compliled from current telephone directories and the list was updated by the sruveyors on visits to the communities. The basic information to be collected during the commodity flow survey was to develop commodity movement estimates from all interviews. Later the data was transfered on to a computer in a pre-coded format. It should be noted that place information and commodity codes had the same structure as the codes currently used by Alberta Transportation for weigh scale interviews - with certain modifications as deemed necessary to overcome specific deficiencies when they were encountered-were made. The basic information on commodity flow data collected were, origin-destination, type of commodity,type of firm, type of goods, annual tonnage, average shipment size, type of load (full load or less than full load), type of hire (private or for-hire), control (yes or no) , market share etc. In addition to commodity flow data, demographic data, such as population, employment, income, age distribution, etc. were also collected from various sources. The details of coding instructions and sample data fromat used are shown in Appendix I.

3.2.2. Communities Selection

The initial survey consisted of classifying the communities across the province as sources, sinks, or as both. The identification of communities is an very important step in conducting a commodity flow survey of this dimension. Production centres can be considered as sources, that is a place from which supply is originated. Consumption centres can be considered as sinks, that is places of demands for commmodity. A few communities can be considered as sources and sinks, because they behave as re-supply centers along major corridors of commodity movements.

The identification of communities helps to decide whether to collect data from sources (points of production) or from sinks (points of consumption) or from both depending upon what the community is classified. Few of the major sources and sinks in the province of Alberta are shown in shown in Fig.4. In Alberta, majority of the goods are distributed from Edmonton, Calgary, RedDeer, LethBridge and

Medicine Hat. These places are used for principal distribution of commodities across the province by provincial, national, and international manufacturers and suppliers. It is to be noted that some of the source centers may not be the origin of production for all commodities, but they are considered as main sources for the rest of the province. Most of the provincial manufacturing sectors are also located at the source centres.

In addition to these major centres which act as sources, there are many other small, medium and large communities which behave as sources, or re-supply centres either on a regional or corridor basis for the distribution of commodities. For example, Peace River acts as re-supply centre. The other interesting categories are the oil and mines industries that are spread across the province. These places serve both as sources and sinks for commodities that usually amount to large tonnage movement even though the population at these places is very small.

The majority of the communities across the province are identified as sinks, that is places of demand for commodities. For example, Grand Prairie, Grand Cache, Stettler, Three Hills, and Altario behave as sinks. It is to be noted that these communities do produce some commodities, but they primarily depend on one or more major source centres, and thus behave predominantly as sinks. The selection of communities were conducted in consultation with Alberta Transportation on the basis of community size, population, industries and other important factors so as to obtain a good representative sample for the survey. Also, the selection is made such that the centres selected could be used for extrapolation for similar places excluded from the study and for any possible future strategic examination.

Table 1 shows the selected communities in terms of economic base, and commodity supply from competing sources. The type of community and other factors represent very significant indices of activity. For, example, communities which are closer to dominant supply sources others tend to has reduced proportionate retail activity, than do similar communities that are not dominated. This is reflected in reduced tonnage for consumers items, with the difference accounted for by increased shopping travel by consumers going to alternate shopping community. Another factor to be considered, which directly has impact on the consumption patterns is the presence or absence of a high proportion of transients (e.g tourists or oil field servicing crews, lumbering, etc.) within a community's market area. This factor also has been shown in Table 1.

3.2.3. Cosnsignee or Shipper Interviews

The survey took the form of interviewing consignees and shippers of commodities. An attempt was made to interview all of the firms which moved comodities in any significant quantities. Because of their size, Edmonton and Calgary were treated seperately. For Edmonton and Calgary, a 100 percent commodity flow sample was not fesible, and consequently heavy reliance was placedupon both consignee "source" information withinthe selected communities. Also, for this the Alberta Transportation's 1977 cordon survey was used. The data that were collected was used to make a special study entitled 'Mode choice Survey prepared in 1978(43,44). A significant portion of the interviews were conducted in person. The rest was done over the telephone. The initial list of interviewees were compiled from current telephone directories and the list updated by the survey or on visits to the various communities. With the exception of alcohol and grain movement, all businesses within the community were interviewed. The information on shipments of alcoholic beverages were developed with assistance of the Alberta Brewer's Agents. Grain shipments were developed on the basis of Canadian Grain Commission published statistics for each county point. The number of people who refused to co-operate was insignificant in terms interviews conducted. The nuber of people who refused to give data during the survey was not reported (43,44).

A large portion of interviews was conducted without prior discussion with the subjects. In majority of the cases, most people were co-operative once the purpose of the study was explained and confidentiality was ensured. Many of the firms contacted were not able to talk in terms of truck loads or car loads except shippers and receivers of bulk commodities. Thus when, calculating volumes, the interviewer would have to ask questions like, " How many cases of milk per week do you receive?" or "How much does the average tractor weigh and how many tractors do you sell per year?". Once the basic unit was established, it was relatively easy to develop weekly, or yearly average movements. In certain cases, it was necessary to follow-up a visit to outline specifically the type of data required. Although this procedure was time consuming, it proved to be effective. In case of suspected responses, it was checked through suppliers and comparisons made against competitors' data and adjusted.

In dealing with certain commodity groups such as men's and ladies wear, drug and shoe stores, it proved expedient to extrapolate the data for some of the businesses from data collected from a representative sampling of similar firms. This was effective in that there was a high degree of similarity in soruces of supply, volumes and transportation modes. The validation of the collected data was accomplished in many ways. For example, interview was conducted with both consignees and shippers in order to check consignee estimates and to provide a comparative demonstration of the survey technique. In interviewing the consignee the emphasis was to collect commodity volume data along with an estimate by each consignee of his market share. Names of the competitors were solicited in order to ensure that no major consignee would be overlooked during the survey. Other questions that were asked pertained to nature of truck service and problems encountered by the consignee with a view toward further understanding the type of service for a particular commodity being transported. Where necessary, new data were developed with existing inventories and data base for maximum validation.

The survey was structured according to sampling of shipper or recepients of commodities based on selected communities as listed in Table 1. In order to develop the complete data, reference was made to various agencies, publications, telephone interviews, letters, and visits to respective agencies. The complete survey procedure and validation of data has been presented in two reports presented by the consulting firm to Alberta Transportation Department(43,44).

3.2.4. Classification of Commodities

The classification of commodities into distinct classification forms one of the major tasks in commodity transportation planning and this work is complicated by the fact that there are numerous products to be classified. In this survey, the commodities were classified as major and minor codes. There are 18 major commodity codes and each of the major commodity codes has several minor commodity codes. Table 2 shows the major and minor commodity codes adopted during the survey. This list was presented to the consignee/shipper in completing the survey forms. It was expected that the classification in Table 1 is adequate for the consignee or shipper to select a particular major and minor commodity code combination to describe the type of commodities being transported. There were 7,493 interview records available, of which 1,318 were of shippers and 6,175 of consignees, indicating that the communities across the province are mostly demand points. The principal restriction of the data base lies in the suitability of their use for developing a realistic sample of activity pertaining to smaller centres throughout the province. For larger centres such as Edmonton, and Calgary, the statistical validity of the data collected is greater because of the larger sample size involved. Overall the data collected are reliable and can be used for commodity transportation analysis.

CHAPTER 4 - COMMODITY DISTRIBUTION MODELS

CHAPTER 4

COMMODITY DISTRIBUTION MODELS

4.1. INTRODUCTION

It has been predicted in our modern industrial world, over twenty percent of the labour force is directly or indirectly employed in the operation, servicing, manufacturing and selling transportation facilities. From this, one can deduce the importance and impact of transportation sector has on economy. Commodity transport alone has been growing at a rapid pace over the years. This substantial growth has put a new strain on both national and provincial planners. The large geographical dispersion of places and the wide variety of commodities require innovative planning to effectively co-ordinate and integerate overall potential of our tranportation system(12).

In this environment it is necessary to develop models so as to make decisions and obtain some kind of optimal solution to the problems at hand. The models developed must be able to simulate the demand for regional commodity movements so that the overall efficency of the transportation system can be improved to meet any future demand perceived from the model.

4.2. O-D COMMODITY FLOW TABLES

The origin-destination (O-D) commodity flow tables indicate the volume of commodity in tons shipped between each origin (source) and destination (sink). The preparation of commodity flow tables is time consuming and a difficult task in the development of distribution models. The 'Commodity Flow Table' (CFT) is similar to the 'Trip Table' used in urban transportation planning(21). But preparing an CFT is more complex than preparing a trip table. The complexity arises from the thousands of types of commodities that are being transported by different modes of transportation. It is necessary to reduce this vast data into manageable two-dimensional array for the development of models.

The preparation of the CFT's was done in four stages. The data for this is obtained from the commodity flow survey as explained in chapter 3. In the first stage, the modes or mode combinations to be used were checked. For this purpose, the average shipment size and modes were considered. From Table 4 it is evident that truck and rail mode are the only significant modes of transportation across the province. Hence only truck and rail modes were considered in constructing the CFT's. A further examination into the actual modal share revealed that the commodity flow by rail was not much and hence there was no need to construct sepearate CFT's for rail. Thus the commodity flow by trucks and rail were combined.

The second stage involved the study of origin-destination(O-D) pairs across the province. In constructing the CFT's only the O-D pairs related to the calibration communitites listed in Table 1 were considered. All other available O-D pairs were eliminated from the record. After deciding the mode combination and O-D pairs, the next task was to allot the observed flows into the CFT's. For this purpose, the total annual tonnage of commodities transported between each O-D pairs is used. The preparation of the CFT's was accomplished by making it as simple and objective as possible by taking into consideration all important movements and eliminating relatively unimportant movements between the O-D pairs. That is, interchanges between origin and destinations, which contained large number of zero flows(tons) were considered as insignificant to be included in developing the CFT's.

The third stage consisted of categorizing commodity classifications shown in Table 2 into a manageable number for meaningful commodity categories. The first step in this stage was to identify individual commodities with similar characteristics and group them into a single category. This task was accomplished by carefully looking into the major classifications and the minor classifications given in Table 2 for each commodity. The commodities were grouped into six categories as listed in Table 3. From this point on in the text these six categories will be refered as General Freight, Food Stuffs, Forest, Petroleum, Bulk Liq. and dry, and Heavy Mach. (Heavy Machinery). Some of the commodities had to be left out because of their characteristics or due to lack of sufficient data in the records. The commodities eliminated at this stage were, live stock, trailer-mobile homes, household goods, grains, feed products, and mail.

The final stage of constructing the CFT's consisted in compiling the annual tonnage for each category listed in Table3. From the point of keeping the flow tables simple and objective as mentioned earlier, all insignificant flows between O-D pairs were omitted at this stage. CFT's were constructed for each of the commodity categories. Hence the O-D tables are simply the sum of the observed O-D flows D_{ij} by truck and rail for a particular category within the calibration communities listed in Table 1. The CFT's for each category is shown in Tables 5 to 10. All the intra-trips in the study are equal to zero, i.e., no intra-trips are considered as we are interested in modeling the external flows from each origin or source. Similarly, the zero flows in each row was replaced by the minimum value of the row, with the assumption that some flow must have taken place between the O-D pairs with respect to a specific commodity category. At the modeling(in the algorithm) stage the allotment of intra-trips by the model was avoided by giving a large distance value for the particular origin under consideration.

The distances between the O-D pairs are shown in Tables 11 and 12. A comparison of the CFT's and the distance matrices indicate that majority of the commodity flow occurs from the two major cities Edmonton and Calgary. Also, it is interesting to note that the commodity flows divide between Edmonton and Calgary depending on the nearest source. The summation of the row values in Tables 5 to 10 are considered as the productions(P_i's) and the column summation for each destination gives the consumption(C_j's); which are used as inputs to the model.

Finally, it is to be noted that commodities are produced at a city (source) to satisfy the local consumption and also to meet the demand (consumption) in other towns. Thus, the external component of production is relatively small compared to the local component. The local component is not taken into consideration in constructing the O-D table. For this reason, the tonnages shown in an O-D table may appear to be small quantities.

4.3. MODEL FORMULATION

Commodity transportation is characterized by the movement of a commodity from an origin (source) to a destination(sink). The amount of flow depends on the amounts of production and consumption and a factor that represents the spatial separation between the production zone and the consumption zone. In essence, this can be represented by a gravity model. Since commodities are of several types, the spatial separation factor is different for different commodities. The spatial separation factor that best suits the commodity category is the key element in the formulation of the model. The quantity of a commodity produced at a place is usually known. The amount of consumption of a commodity is known or can be estimated for any given place. A production constrained gravity model is appropriate under these circumstances for distributing the commodity flows.

Suppose we are given P_i, the production of commodity for each origin zone i, and C_j the consumption of commodity for each destination zone j. The spatial separation factor is defined by a function of the distance d_{ij} between origin zone i and destination zone j. In this study the highway distance between the origin and the destination is used. If T_{ij} represents the estimated interchanges between origin and destinations, it can be expressed by a singly constrained gravity model of the form (Wilson, 1970):

$$T_{ij} = A_i \; P_i \; C_j \; f(d_{ij}) \qquad \text{...............(4.1)}$$

This has to satisfy the production constraint:

$$\sum_j T_{ij} = P_i \qquad \text{.........................(4.2)}$$

Therefore, the constant term A_i and T_{ij} are given by:

$$A_i = \frac{1}{\sum_j C_j f(d_{ij})} \qquad \text{.......................(4.3)}$$

$$T_{ij} = P_i \frac{C_j \; f(d_{ij})}{\sum_j C_j \; f(d_{ij})} \qquad \text{.......................(4.4)}$$

The next step in the formulation is the definition of the spatial separation factor $f(d_{ij})$. A family of power transformations of the type(Draper and Smith, 1981), on a variable Z is introduced at this stage:

$$Z^{(\lambda)} = \begin{cases} \dfrac{Z^{\lambda} - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \ln Z & \text{for } \lambda = 0 \end{cases} \qquad \text{...............(4.5)}$$

This family of transformation is continuous on a single parameter λ if it is specified as above. Transformations can be done on Z^{λ} instead of $(Z^{\lambda} - 1)/\lambda$. In such cases, only a scale difference and an origin shift are involved and the basic nature of the subsequent analysis is unaffected. The tranformations for the case of $\lambda \neq 0$ always refer to Z^{λ}.

Let $Z = d_{ij}$ ....(4.6)

and $Z^{(\lambda)} = d_{ij}^{(\lambda)} = f(d_{ij})$ ....(4.7)

Substituting Eq. 4.7 in Eq.4. 4;

$$T_{ij} = P_i \frac{C_j \; d_{ij}^{(\lambda)}}{\Sigma_j \; C_j \; d_{ij}^{(\lambda)}} \qquad \text{...................(4.8)}$$

By performing transformation on $d_{ij}^{(\lambda)}$, specific models can be derived.

1. Transformation: $\lambda \neq 0$

Substitute for $d_{ij}^{(\lambda)}$ in Eq. 4.8, using tranformation for the case $\lambda \neq 0$:

$$T_{ij} = P_i \frac{C_j \; d_{ij}^{\lambda}}{\Sigma_j \; C_j \; d_{ij}^{\lambda}} \qquad \text{...................(4.9)}$$

Different models can be obtained by giving specific values for λ.

2. Transformation: $\lambda = 0$.

Substitute for $d_{ij}^{(\lambda)}$ in Eq. 4.8 using tranformation for the case $\lambda = 0$:

$$T_{ij} = P_i \frac{C_j \ln d_{ij}}{\Sigma_j C_j \ln d_{ij}} \qquad \text{.........................(4.10)}$$

This is a special case of transformation which gives a logarithmic form for the model.

4.4 OBJECTIVE FUCTION

The objective function chosen to calibrate the formulated distribution model for the purpose of this thesis is to minimize the sum of the squared deviations as given by;

$$\text{Min } \Sigma_i \Sigma_j (D_{ij} - T_{ij})^2 \qquad \text{.............................. (4.11)}$$

Where, D_{ij} = the observed interchanges(tons) between origin i and destination j of a particular commodity;

T_{ij} = the estimated interchanges(tons) between origin i and destination j of a particular commodity.

By minimizing Eq.4.11, the algorithm will effectively minimize the percentage errors of the largest flows and will keep the deviations of the lowest flows within limits. The large flows are of importance to the transportation planner and are expected to be simulated most accurately and the zero or negligible flows will not be substantially inflated by the model.

The objective function is achieved by comparing the observed interchanges D_{ij} and the estimated interchanges T_{ij} using Least Square Method(LSM). The value of the coefficent of determination R^2 for the various models for the different exponent values of the spatial seperation factor in Eq.4.9 are obtained. The objective function is satisfied when R^2 reaches a maximum value for a specific exponent value. This procedure of finding the maximum R^2 is explained under the section calibration of the model.

4.5. MODEL CALIBRATION

The values of λ can be varied in a specified range λ_k (k=1,2,......,n), of negtive values, including zero. A specific spatial seperation factor $d_{ij}^{\lambda_k}$ and the corresponding gravity model can be obtained for each value of λ_k. (λ) is replaced by λ_k in Eq.4.8 to indicate the range:

$$(T_{ij})_k = P_i \frac{C_j d_{ij}^{\lambda_k}}{\Sigma_j C_j d_{ij}^{\lambda_k}} \qquad \text{.............................(4.12)}$$

$$k = 1,2,3,\ldots\ldots\ldots\ldots n$$

In this range of gravity models, only one will give interchanges (T_{ij}) closest to the observed interchanges D_{ij}. The gravity model to suit this criterion can be written as follows:

$$(T_{ij})_\ell = P_i \frac{C_j d_{ij}^{\lambda_\ell}}{\Sigma_j C_j d_{ij}^{\lambda_\ell}} \qquad \text{..........................(4.13)} \qquad 1 < \ell < n$$

Where ' ℓ ' represents the optimum condition. The above model is termed as optimized gravity model in order to avoid any confusion with the standard gravity model.

The spatial separation factor $d_{ij}^{\lambda_\ell}$ in Eq.4.13 is specific to the given commodity category and therefore the optimized model. The optimum value λ is to be determined for each commodity type within the range λ_k. In order to emphasize this requirement, a notation 'm' is to be included in Eq. 4.12 to indicate the commodity category:

$$(T_{ij})_k^m = P_i^m \frac{C_j^m d_{ij}^{\lambda_k^m}}{\Sigma_j C_j^m d_{ij}^{\lambda_k^m}} \qquad \text{....................(4.14)} \qquad k = 1,2,\ldots\ldots n \quad m = 1,2,\ldots\ldots q$$

The optimized gravity model is therefore given by:

$$(T_{ij})_\ell^m = P_i^m \frac{C_j^m d_{ij}^{\lambda_\ell^m}}{\Sigma_j C_j^m d_{ij}^{\lambda_\ell^m}} \qquad \text{....................(4.15)} \qquad 1 < \ell < n$$

LSM is used to determine the optimum value λ_ℓ and the corresponding optimized gravity model for each commodity type.

Let one commodity category be considered and the parameter λ_k be given a specific value, say λ_1. The gravity model for this value λ_1 is as follows:

$$(T_{ij})_1 = P_i \frac{C_j \; d_{ij}^{\lambda_1}}{\Sigma_j \; C_j \; d_{ij}^{\lambda_1}} \quad \text{.....................(4.16)}$$

The notation ' m ' for commodity type is not included in Eq. 4.16 since the equation represents only one commodity category. The $(T_{ij})_1$ gives the matrix that shows the commodity flows between zone i and zone j when the spatial seperation factor assumes a value of $d_{ij}^{\lambda_1}$. The interchanges in the $(T_{ij})_1$ matrix can be compared with the observed interchanges D_{ij} and the differences can be obtained.

LSM is used to compare the interchanges $(T_{ij})_1$ matrix and the observed interchanges D_{ij} and to obtain the value of the coefficient of determination R^2. By varying the values of λ_k in a specified range of zero and negative values, a series of gravity models can be obtained. For each gravity model the value of R^2 is determined. The values of R^2 are plotted against the corresponding λ_k. The values of λ_k at which the R^2 reaches a maximum value is the one that gives the optimum gravity model distribution and also satistfies the objective function in Eq.4.11. The optimum gravity model is as defined earlier is:

$$(T_{ij})_\ell = P_i \frac{C_j \; d_{ij}^{\lambda_\ell}}{\Sigma_j \, C_j \; d_{ij}^{\lambda_\ell}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(4.17)$$

Where 'ℓ' represents the optimum condition. Thus, the gravity model is calibrated by optimizing the power parameter λ. The spatial separation factor $d_{ij}^{\lambda\,\ell}$ in the optimized gravity model is different for the given commodity type. The above procedure for optimization is useful because $(T_{ij})_k$ and D_{ij} values are related in a direct manner(3). The higher the value of R^2 for the optimized gravity model, the better will be the calibration.

The singly constrained gravity model formulated in Eq.4.1 for estimating the interchanges between origin and destinations is a non-linear function of the normalizing constant A_i, and the exponent λ. In order to achieve the objective function, that is, to minimize the Eq. 4.11 it would be necessary to prove that the objective function is convex in λ. Difficulties deriving the explicit relationship of the estimated flows T_{ij} as a function of the spatial seperation factor exponent have prevented such a proof. However, attempts have been made to prove that the objective function, sum of squared deviations is convex for possible matrices.

Leech has written the objective function as a explicit function of the exponent λ for a doubly constrained model with both column and row normalizing factors(25). He has proved that the second derivative of this explicit function is non-negative, proving that the objective function is convex in λ. This means that there exists a value of λ for which the objecive function in Eq.4.11 has a minimum value, meaning that it is possible to estimate the interchanges as close to the observed interchanges for a particular value of λ. The current calibration procedure attempts to the same by finding a λ value which satisfies the objective function by maximizing R^2 value and thus obtain estimated interchanges as close to the observed as possible using the developed distribution model.

As explained in the previous section, O-D tables are developed for the six categories of commodities shown in Table 3. The O-D table for each commodity category is used for the development of the gravity model. The sum of the quantities in a row of the table is P_i. The sum of the quantities in a column of the table is C_j. The distance d_{ij} between an origin and destination is known. The interchanges $(T_{ij})_k$ can be determined by the model in Eq.4.12, using the P_i, C_j and d_{ij} values. The spatial separation factor d_{ij}^{λ} is considered as a decreasing function for gravity model applications(21). A typical curve for the spatial seperation factor $d_{ij}^{-0.5}$ is shown in Fig. 6.

4.6 EXAMPLE

From Table 3 Food stuffs is considered as an example to illustrate the procedure for the development of the model. A range of negative values from -2.0 to 0.0 is assumed for λ_k since the spatial seperation factor ($d_{ij}^{\lambda_k}$) in gravity model is a decreasing function. In this range, a specific value, for example, the first value λ_1 = -2.0 is assumed and substituted into gravity model given by Eq.4. 12.

$$(T_{ij})_1 = P_i \frac{C_j \; d_{ij}^{-2.0}}{\Sigma_j \; C_j \; d_{ij}^{-2.0}} \qquad \text{.....................(4.18)}$$

The $(T_{ij})_1$ matrix is obtained using the above formula with known values for P_i, C_j, and $d_{ij}^{-2.0}$. Each cell in the matrix gives the estimated interchanges between each origin and destination pair. These interchanges in $(T_{ij})_1$ matrix is compared with the observed interchanges D_{ij} by regression analysis and the values of R^2 is obtained by LSM. The above procedure is repeated for each value of λ_k in the range of -2.0 to 0.0; however, the values of P_i, C_j and d_{ij} remains the same. The value of R^2 corresponding to the value of λ is tabulated as shown in

Table 13 for food stuffs. The variation of R^2 against λ is shown in Fig. 7. The maximum value of R^2 gives the optimum value for λ which is -0.5 for food stuffs and this value of the exponent satisfies the objective function. Thus the optimized gravity model for food stuffs can be written as:

$$(T_{ij}) = P_i \frac{C_j \; d_{ij}^{-0.5}}{\Sigma_j \; C_j \; d_{ij}^{-0.5}} \qquad \text{..................(4.19)}$$

Since this is a production constrained gravity model, the productions (P_i) of T_{ij} matrix are identical to the observed productions of D_{ij} matrix. It is found that the estimated consumption values C_j^* at the end of the calibration procedure for the T_{ij} matrix of the optimized gravity model in Eq.4.19 are in general agreement with observed consumptions C_j.

The optimized gravity model for each of the commodity categories in Table 3 is determined by the procedure described above. The variations of R^2 and r against λ for each commodity category is shown in Table 14. A high correlation coefficient is a most desirable as it indicates that the estimated flows are high when the observed flows are high and that the estimated flows are low when the observed flows are low. The correlation coefficient compares only the relative changes in two sets of figures. A high correlation value in a distribution model is of primary interest, because it accurately differentiates between the large and small flows. Also, it determines the degree to which variability in the observed flows is explained by the changes in the estimated flows by the model. The maximum value of R^2 (optimum λ) for the six commodity categories are in the range of 0.71 to 0.88. These results indicate that the estimated O-D interchanges of the optimized models and the observed O-D matrix are highly correlated. Therefore, the optimized gravity models are acceptable.

The calibration of the gravity model is further substantiated by plotting 'Commodity Haul Frequency Diagram' (CHFD) for the model and the observed data as shown in Fig. 8 and making a visual comparison. For this diagram, the distance is divided into ranges, 0 - 100 kms, 100 - 200 km. etc. The O-D interchanges in T_{ij} matrix (optimized gravity model) which have O-D distances in the given range are added and expressed as a percentage of total interchanges (total tonnage in the matrix). This percentage is shown as 'percentage of commodity haul' in Fig. 8. The percentage of commodity haul for each range is plotted against the midpoint of distance in that range. The result is the CHFD as shown in Fig. 8. A similar diagram is plotted for the observed interchanges D_{ij} . The two diagrams show general agreement by visual comparision. This indicates that the O-D interchanges given by the optimized gravity model and the observed O-D interchanges are in general agreement. Thus, the calibration of the optimized gravity model is found to be satisfactory. However, it is to be noted that in developing the observed O-D Tables 5 to 10 the zero interchanges between the origin and destinations were replaced by the minimum interchange in that particular row. This was done with the assumption that some quantity of interchange must have taken place between the origin and the destination for that particular commodity category under consideration. The gravity model allots some tons to all these cells and they alter the percentage of commodity haul in the various ranges. This could be one of the reason for the discrepancy between the two diagrams in some ranges in Fig. 8. The CHFD is similar to the so called 'Trip Length Frequency Diagram' used in the calibration of gravity models for urban transportation planning (21). Therefore, CHFD's can be used to substantiate the calibration of gravity model for commodity transportation.

The complete set of optimized gravity models for the different commodity categories are given in Table 15. The R^2 vs λ and the commodity haul frequency diagrams for the remaining five commodity categories are shown in Figs.9 to 18. There is a general aggrement between the estimated and the observed flow tons. The λ values for the optimized gravity models vary in the range of -0.25 to -1.0 for thesix commodity categories. This is significant because it shows that the spatial seperation factor d_{ij}^{λ} is different for each commodity category.

At this stage it is important to mention that, once the optimum λ value is found, any further attempts to improve the estimated flows by calculating new consumption values and conducting iterations (balancing attractions in gravity model) as suggested by Hutchinson, 1974, did not improve the results (estimated flows) obtained from the optimized gravity model. Hence the process of iteration to improve the estimated flows was not pursued. The new consumption values for the next iteration is given by the following algorithm(21);

$$\text{New } C_j^k = \frac{C_j}{S_j^{k-1}} C_j^{k-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(4.20)}$$

Where,

C_j^k = Consumption value used for iteration k,

S_j^k = actual consumption total(column summation of T_{ij}, i.e, C_j^*)

C_j = desired consumption .

4.7. EXPONENT VARIATION AND MODEL RESULTS

The fundamental question in the calibration process of the developed model is the value of the power parameter λ in the spatial seperation factor d_{ij}^{λ}. Numerous studies using the gravity model concept in urban transportation planning(21) have found that the exponent varied depending on the different type of trips under investigation. These findings clearly suggest that similar variation may occur with regard to commodity flows, as there are different types of commodities. This could also be expected by inspecting the difference in volume transported in the O-D tables in Tables 5 to 10 for each category of commodity.

The major objective of a commodity distribution model is to move commodities between points of origin(source/production) and destination (sink/consumption) such that they approximate the observed interchanges(30). If a model to determine the value of the spatial seperation exponent for given commodity category would greatly enhance the efficency of the developed distribution models(25). If it is possible to identify that there exists a relation between the exponent and certain commodity characteristics, then it will be possible to update the trip distribution model over time without repeating the complete calibration.

In the formulation of the model the exponent is the only parameter to be determined and others may be introduced if necessary. Under some situtations the value of the exponent can be inferred on a priori grounds(25,30,41). The exponent will have an effect as the number of production(source/supply) and consumption (sink/demand) points increase. Leech(25) took into consideration the following hypothetical situations to explain the variation in the exponent value. The most extreme case is the case where the commodity is produced in a very localized area due to its utilization of a local resource or raw material. Assuming that all regions demand the commodity, then it would be expected that the distance exponent will be very close to zero, since there is no other alternative supplier or nearer souce. It is conceivable that the flows of a highly specialized product could behave in much the same way if the production of the product is concentrated in one or two regions.

Now, consider a dispersed type of production and consumption, that is, the case of many soruces and many sinks. In this case a large value of the exponent forces demands to be satisfied as close to the demand points as possible. It is evident from the above reasoning that the fluctuations in the gravity models exponent can be viewed as a function of: 1) the geographical concentration of production ; 2) the geographical concentration of consumption; 3) the local demand for the commodity; and 4) characteristics of the commodity such as the value per unit of weight, commodity etc(25,30).

The variation in the value of the exponent in the study ranges from -0.25 to -1.0 for the different category of commodities. This difference in the spatial exponent can largely be due to the concept of potential accessability due to the difference in commodities and the availibility of many production and consumption points. Furthermore, the nature of distribution across the selected communities is such that the place close to the sources will try to avoid the penalty of long haul by a greater degree of local self-sufficency for certain commodities (11). The second reason is that the difference in the exponent could be expected, since the use of a single exponent value for the whole province for all commodities taken together may result in serious over-or under estimation of interchanges.

A study of commodity flows and spatial seperation factors in Britain are reported in Chishlom and O'Sullivan, 1973. In the British study, R^2 values for individual commodity flow models vary from 0.24 to 0.62, whereas in this study vary from 0.71 to 0.88 as shown in Table 15. The spatial seperation factors range from $d_{ij}^{-0.8}$ to $d_{ij}^{-1.6}$ for individual commodity flow models in the British study, whereas they vary from $d_{ij}^{-0.25}$ to $d_{ij}^{-1.0}$ as shown in Table15. In the British study the distance exponent for building materials is -1.6(high) and it is -0.8(low) for steel,scrap, and chemicals. The study states that there appears to be a slight tendency for more highly valued and processed goods, such as steel, transport vehicles and equipment and chemicals, to have low exponent value as opposed to higher exponent values for bulk goods such as coal and coke, and building materials, but it is not very marked(11). Comparing these results to the results of this study, it is found that the distance exponent for heavy machinary and bulk liquids and chemicals categories have -1.0(high) and it is -0.25(low) for highly processed product category petroleum. The values obtained seem to be in general aggrement with those of the British study.

Black(9) used Eq.2.3 which is similar to Eq.4.4, he tried to estimate the flow of 24 commodity classes across nine census regions in United States of America. The level of aggregation he used were better than those used in the British study and he had more encouraging results(11). His data was more aggregate spatially but much more finely differentiated by commodity categories. He ignored intra-regional movements similar to this study. The variance explained by his models ranged from 73% for petroleum and coal products to 95% for textile and leather products. The exponent value varied from -0.25(low) for non-electrical industrial machinery to -5.3(high) for stone, clay, and glass products. The inference was made that the value of the exponent ranged widely, the general tendency for more highly processed goods to have lower exponent values and bulkier and low value goods to have higher exponent values. Comparing Black's result with our exponet values, the following is observed. In our study the highly processed commodity category petroleum products has a low exponent value of -0.25 and bulkier goods such as constructio materials and bagged products, and heavy machinary commodity categories have high exponent value of -1.0. The exponent value for other commodity categories such as general freight, food stuff, forest proucts have exponent values between -0.25 and -1.0. The variance explained by the models range from 71% for food stuffs and 88% for forest products. Thus the results presented in the study are consistent with the results expected for the exponent values taking into consdieration the commodity category. However it is not possible to closely compare the results of studies with so much difference in data and commodity classifications used.

The exponent value variation is expected due to the differences in the commodity category and the type of data utilized to construct the O-D matrix. The difference in exponent value may also be due to sampling variance. Thus the true value of the exponent may be difficult to find taking into consideration the differences in constructing the O-D tables and the commodity category.

The significant feature of the methodology used in this study is the development of a spatial seperation factor (d_{ij}^{λ}) that is diferent for each commodity category. The term ($C_j \; d_{ij}^{\lambda}$) represents the accessibility of the commodity at the destination. The term $\Sigma_j C_j \; d_{ij}^{\lambda}$ represents the total accessability(5) . Thus the commodity produced at the given origin is distributed according to the proportion of accessibility at the given destination in the total accessibility of the entire area. Since characteristics of commodities transported are different, it is logical to expect different spatial seperation factors and this gives validity to the modelling procedure.

In this study, distance d_{ij} between origin and destination is taken as a measure of spatial seperation. However, other measures, such as transportation cost, travel time etc. can be used individually or in some combination. Fortran computer programs (Computer Program I, Appendix II) have been written for data processing and for the computation of the optimized gravity model . Standard statistical package such as Mini-Tab have been used to do the least square analysis and generate R^2 values for the different iterations during the calibration process.

4.8. SUMMARY AND CONCLUSIONS

4.8.1. SUMMARY

This study deals with the development of distribution models for commodity flows between communities in the province of Alberta, Canada. The data base for this study is obtained from the commodity flow survey done in 1977 for Alberta Transportation Department. The characteristics of production and distribution of these commodities to satisfy the demand(consumption) at various population centres, differ from one commodity to another. This fact has been recognized in this study and therefore an individual gravity model is developed for each commodity category. The gravity model is calibrated by an optimization technique which invloves the use of a power function for the spatial separation factor and regression analysis. Regression analysis is used to test the agreement between estimated O-D interchanges by the model and the observed O-D commodity flows. The calibration of the gravity models are substantiated by plotting 'commodity haul frequency diagram' for the estimated and observed O-D interchanges and comparing them by visual inspection for agreement.

A set of optimized gravity models for the various commodity categories represents the commodity flows between production centres(sources) and consumption centres(sinks) in the province of Alberta. Results show that the optimized models are acceptable.

4.8.2. CONCLUSIONS

This study is based on a comprehensive commodity flow survey done in Alberta which covers the full-range of commodities and the modes of transportation.

The survey data is grouped to specific commodity categories and a set of O-D tables have been developed. These O-D tables are the basis for the development of gravity models. The following are the conclusions of this study:

1. A set of optimized gravity models, one for each commodity category is developed to represent the commodity flows on a province wide basis.

2. The optimized gravity model for each category has a different spatial seperation factor. This factor can be determined by an optimization technique, which uses power function and regression analysis.

3. Similar to the 'Trip Length Frequency Diagram', the 'Commodity Haul Frequency Diagram' can be used to check the agreement between the estimated and observed flows.

4. The results show that the above models are acceptable and that they are consistent with those expected normally for commodity transportation models.

5. The calibration of gravity model using the optimization technique has potential for further applications.

CHAPTER 5: MODAL SPLIT ANALYSIS AND MODELS

CHAPTER 5

MODAL SPLIT ANALYSIS AND MODELS

5.1. INTRODUCTION

Modal split is the process of dividing commodity movements among competing modes. Modal split models are constructed to aid in forecasting the demand for a particular mode. There are two very important reasons why modal split analysis and models are necessary. Firstly, it could be a market research, where it is intended to establish the importance of various factors in choice of mode. Secondly, the analysis can be for forecasting, where the objective is to forecast future demand for different modes for planning purposes. The market research type analysis could establish the existing behavioural factors lying behind the observed distribution of traffic. Under forecasting it may be sufficient to establish that certain commodities or certain consignment weights are sent by a specific mode. From the point of the planner and operator, it is necessary to know why certain commodities are sent by competing transport modes(33,34,35).

Modal split analysis and modeling are basically carried out as only part of a global forecasting exercises and cannot be seperated from traffic generation, attraction, and distribution, because all three react upon each other. For example if large flows develop upon a particular route, this may encourage the use of a particular type of transport, whilst the development of a particular type of transport might affect the traffic flows. Similarly, the introduction of a new and efficient mode can affect plant location,in turn this affects the transport demand. Therefore it is evident that there is inter-dependence of the entire system of traffic generation, attraction, distribution and modal split.

The mode choice phenomenon has a very complex character because of the nature of the demand for commodity transport services. Therefore it is necessary to understand and identify the fundamental relations in the mode choice process together with other decision areas in the organization of production inputs and the organization of distribution of finished goods. If it is possible to determine these fundamental relations between the mode choice process and other relevant decision areas an explanation may be given for the actual observed mode split behaviour and the changes that may be anticipated on the basis of certain exogeneous variables in this process. Understanding mode choice is also in the interest of modern management where decisions in the logistical area are becoming more and more integrated with decisions in the production and distribution area.

The main objective of this study is to recognize the pattern of commodity movements across the province and to build suitable modal split model(s). The first phase of the study deals with the identification of the major carriers in the transportation of commodities in the province of Alberta, Canada. The factors taken in to consideration are the average shipment size, control(yes or no), loads(full load, less than full load), hire(private or for-hire), and type of commodity. For modal split analysis Cross Tabs package was used(Computer Program II, Appendix II).

In the second phase of the modal split study the data is collapsed to form multidimensional contingency tables to develop log-linear models which assist in uncovering association between the variables. After a suitable log-linear model is selected, secondary analysis is performed using the logit model to obtain a table of log-odd effects from which it is possible to determine the influence of the explanatory variables on a designated response variable. The designated response variable is the 'mode' of transport and all other variables are treated as explanatory variables. From the models developed it is possible to determine the percentage share of each mode and also identify the commodity being transported across the province.

5.2. COMMODITY MOVEMENTS ACROSS ALBERTA

The 7,493 interview records of both shippers and consignees were used.The detailed description of the data is given in chapter 3. There were 1,318 shippers and 6,175 consignees data, there were very few missing data as compared to the total length of the record. This analysis deals with the modal choice of large communities(> 16,000 population) to small communities(< 300 population).

5.2.1. Analysis of Commodity Flow

In previous studies of modal-split analysis, weight and distance have been used largely to identify general trends in modal share (6,34). In this study, average shipment size has been used along with other important factors. The type of commodity transported by each mode has not been identified under previous studies, and this analysis attempts to identify the major commodities carried by various modes under different average shipment size categories and the limitation of each mode to carry commodities. Any peculiarities in commodity flow are also recognized.

The survey data essentially gives the regular flow of commodity between origins and destinations. For the purpose of modal share analysis, the data of shippers and consignees were combined, and it is not expected to have any impact, as we are interested only in determining the mode choice (31). The obvious limitations of this analysis are confined to the fact that several other factors which may influence mode selection are omitted due to non-availability of such data and some of the factors not considered may not have much impact in conducting this particular type of analysis.

5.2.2. Average shipment Size and Commodity Type

The first two factors considered in recognizing the modal split across the province are average shipment size and commodity type, as classified under major commodity code. The survey data contained the annual tonnage and average shipment size in tons for each commodity flow. Table 4 shows the distribution of cargo movements by average shipment size and the eight modes used for transportating commodities across the province. It is evident from the table that commodities are carried exclusively by truck and rail. In view of this, only truck and rail modes are discussed in detail with occasional reference to other modes. There were also some exceptions to the above observation. A close look at the commodities shipped under various categories helped to identify the commodities and any peculiarities in shipment.

Table 16 shows the various major commodity groups, average shipment size and the modes selected. Trucks carry exclusively all goods under average shipment size below 30 tons. Under this category, rail moves a very small percentage of specific commodities, such as general freight, heavy machinery, metal products and construction materials. Trucks are exclusively used for moving all food commodities under all average shipment size categories. The use of trucks to carry all commodities under most average shipment size category indicates the perception of shippers and their choice over rail, which is the only closest competitor among all other modes. As the average shipment size increases above 30 tons, the rail mode is used more frequently to transport specific commodities as shown in Table 16. Few of the commodities which are transported by both truck and rail are in the middle range of 20 to 40 tons of average shipment size.

Petroleum products are transported by truck, rail, pipe and marine. Pipe mode is used exclusively to transport petroleum products. Bus mode is used for average shipment less than one ton for commodities such as general freight, heavy machinery, and metal products. Air mode is used for shipping general freight, food non-perishables and food perishables under two tons of average shipment size. Other modes, as mentioned in Table 4, are used exclusively to move general freight under average shipment size of 1 ton.

The average shipment size clearly indicates the preference for a particular mode and the identification of the commodities transported substantiates this observation. In concluding the preference for a particular mode, one should consider the number of flows under each mode and the average shipment size category. As the average shipment size increases above 30 tons, the number of flows reduce considerably.

5.2.3. Full Load and Less Than Full Load

Full load (FL) and less than full load (LTFL) plays an important role in the rates charged, investment in terminals, specialized pickup service, and delivery equipment. Significantly, the FL transportation costs are always much lower than LTFL, an important economic factor for the shipper. The profit margins on FL are much narrower than those for LTFL. Also LTFL requires specialization in the regional market delivery. It is to be noted that FL and LTFL markets are different in several critical aspects. Thus many of the carriers identify themselves as either FL or LTFL carriers. From previous studies, it has been seen that truck carries commodities for less than 500 miles, and rail carries commodities for much larger distances(12).

Table 17 shows the distribution of commodities by load. In case of truck mode, there are 55.3% LTFL and 44.7% FL, indicating that trucks are utilized more for LTFL. The large LTFL movements by trucks indicate that commodities are moved over shorter distances, which is the characteristic of LTFL business. The opposite is true in case of rail mode, 90.5% of all shipments are FL and only 9.5% is LTFL, indicating the use of the rail mode for FL and that the commodities are carried over larger distances. Also the use of trucks for LTFL and rail for FL can be explained from the dominance of each mode over different average shipment size. Trucks are exclusively used to transport lower average shipment size, and most of these shipments (2,330) are less than 1 ton and are LTFL as shown in Table 18. As the average shipment size increases, trucks and rail are used for full load shipments.

Table 18 clearly indicates the preference of shippers mode of choice with respect to average shipment size and loads. Table 19 illustrates the distribution of commodities by commodity type, average shipment size and loads. A few of the commodities that are predominantly sent under LTFL are general freight, food commodities, heavy machinery, metal products, petro products, forest products and construction materials.

The classification of commodities distribution according to FL and LTFL indicates clearly that shippers consider economy in terms of transportation cost, average shipment size and distance travelled in making their modal choice. Few of the commodities, such as general freight, food perishables, food non-perishables, petroleum products, construction material, etc., which are supplied regularly and those which come under lower average shipment size, are sent as LTFL indicating the intra-provincial movement of essential commodities.

5.2.4. Private and For-Hire

Commodities are carried by either owned (private) or of other (for-hire) carriers. The size and structure of the Canadian trucking industry is very large. In private trucking as the name implies, the transportation of commodities is done privately or in owned vehicles. For-hire trucking is defined as carriers who transport commodities for compensation (payment), either for general public or for shippers (12). The above broad definition of private and for-hire gives essentially the difference between the two types of carriers. The vast majority of private carriers are extensively regional.For-hire trucking carry commodities for larger distances and carry much greater capacity . This could be possibly due to the type of vehicle used, usually the private carriers use straight trucks and for-hire use tractor-trailer which have much larger capacity than straight trucks and are more suited for larger distances.

Table 20 illustrates the distribution of commodities by classifications, mode and hire. As expected, there were more commodity movement by private carriers (79.62%) than for-hire (20.38%) . The greater use of private carriers indicates that there are more local or regional flows over shorter distances. In general, private carriers are involved in short or medium distance and for-hire in long distance travel. There is no proof that one group carrier is specialized in the transportation of a particular commodity(12). The commodities that seem to be predominantly transported by for-hire includes general freight, food commodities, and petroleum products.

Table 21 shows the distribution of commodities by mode, load and hire. It is interesting to note that there are more LTFL (55.8%) by both private and for-hire carriers, indicating the use of trucks exclusively for LTFL. There is need for more study to substantiate many of the observations under type of loading (FL or LTFL) and the role of private and for-hire carriers.

5.2.5. Selection of Mode by Control

The survey contained a question asking the shipper or consignee if they decided how the commodity is to be shipped. That is by which mode the commodity has to be transported. They indicated the type of control they had by marking 'yes' or 'no' on the survey forms. Table 22 shows the percentage of mode selection by control. Truck mode is used for 2742(41.5%) shipments with control, 3867(58.5%) shipments without control. Rail is used for 146(30.9%) shipments with control and 327(69.1%) shipments without control. Trucks are used largely for shipments of lower average shipment size which are LTFL and are primarily transported by using private carriers. Whereas rail is used for shipments of higher average shipment size and are generally full loads.

This substantiates observations in previous sections as to the selection of a mode and that shippers consider economy in transporting goods. Bus and other modes are used to transport some portion of their total shipments whose average shipment size are less than 1 or 2 tons. The decision as to the selection of a mode are based on various factors which are perceived as important and advantageous in transporting goods. The control factor gives an insight into the decision making process of shipper or consignee as to the selection of a mode.

5.3. LOG-LINEAR AND LOGIT MODELS

5.3.1. Contingency Tables

In transportation planning most of the data is collected through survey or questionnaires . This data can be categorized, depending on the objective of the study, to obtain frequency counts. When this data is cross-classified according to their constituent frequency counts, they form a contingency table. When three or more variables are cross-classified, the induced table is called a multidimensional contingency table(7,13).

Let A and B be two variables with I and J levels (different categories) respectively. They can be cross-classified to form an I X J table . The data in each cell are represented by the frequency counts x_{ij}, where i=1,2....I and j=1,2....J. Table 23 is called a 2 X 2 contingency table, for variables A with A_1 and A_2 levels and B with B_1 and B_2 levels (i=j=1,2).

Measure of association among variables assist the researchers in understanding the nature and the extent of the relationships. The most commonly used measure of association is called cross product ratio or odds ratio. P_{ij} represents the true probability of an observation falling in cell(i,j) in Table 23. Let

P_{11},P_{12}, P_{21}, and P_{22} represent the true probabilities of an observation falling in cell(i,j). The ratio P_{11} divided by P_{12} is the odds(probability) of falling in the first column conditional to being in the first row. The odds ratio is represented by Θ, which is the relative odds for the two rows(7,13,16,17).

The cross product ratio or odds ratio is defined by Eq. 5.1. The range of values for Θ lies between 0 and ∞. The logarithm of the odds ratio is called logodds and it is used because of its mathematical convenience. The logarithm of odds ratio ranges between $+\infty$ and $-\infty$. Hence Θ becomes symmetric in the sense that two values of Θ, say Θ_1 and Θ_2 such that $\ln \Theta_1 = -\ln \Theta_2$ represents the same degree of association but in opposite directions. This property is used in developing the log-linear and logit models.

$$\Theta = \frac{P_{11} P_{22}}{P_{12} P_{21}} \qquad \text{...(5.1)}$$

The test of any model is its ability to fit the data and assist in interpreting the relationships between the variables. In selecting the best model for a given set of data using log-linear model, parsimony, goodness of fit, and interpretability have to be checked. Parsimony refers to having as few variables as possible in the models without losing the goodness of fit. Goodness of fit represents the ability of the model to provide estimated cell counts that are close to observed data. Finally the model should be able to explain the physical process observed, or an established theory(5).

In order to understand the association between the two variables, a log-linear model can be used. The log-linear models are a class of mathematical models designed to uncover associations that exist in contingency tables. If m_{ij} denotes the expected frequency associated with cell (i,j) for a sample of n observations, the saturated log-linear model for the two variables with I and J levels is written as:

$$\ln m_{ij} = \mu + \mu_{A(i)} + \mu_{B(j)} + \mu_{AB(ij)} \quad \text{.....(5.2)}$$

with constraints

$$\Sigma_i \mu_{A(i)} = \Sigma_j \mu_{B(j)} = \Sigma_i \mu_{AB(ij)} = \Sigma_j \mu_{AB(ij)} = 0 \quad \text{.....(5.3)}$$

The parameters on the right hand side are called μ terms. They can be interpreted by using an analogy to the analysis of variance:

$$\mu = \Sigma_{ij} \frac{\ln m_{ij}}{IJ} \quad \text{.....(5.4)}$$

$$\mu_{A(i)} = \Sigma_j\left(\frac{\ln m_{ij}}{J}\right) - \mu \qquad (5.5)$$

$$\mu_{B(j)} = \Sigma_i\left(\frac{\ln m_{ij}}{I}\right) - \mu \qquad (5.6)$$

$$\mu_{AB(ij)} = \ln m_{ij} - \mu_{A(i)} - \mu_{B(j)} + \mu \qquad (5.7)$$

The specific constraint imposed on the interaction terms $\mu_{AB(ij)}$, i=1,2....I and j=1,2,....J shows that (I-1) X (J-1) independent parameters are required to fully account for association in an I X J table. Since I=J=2 in the example, it requires only one parameter.

This is seen immediately from the relations:

$$\mu_{AB(11)} = -\mu_{AB(12)} = -\mu_{AB(21)} = \mu_{AB(22)} \qquad (5.8)$$

based on the constraints imposed on the model given by Eq. 5.1. Similarly the same could be extended for three variables A,B, and C with I, J, and K levels respectively.

The saturated model is expressed by:

$$\ln m_{ijk} = \mu + \mu_{A(i)} + \mu_{B(j)} + \mu_{C(k)} + \mu_{AB(ij)} + \mu_{AC(ik)} + \mu_{BC(jk)} + \mu_{ABC(ijk)} \qquad (5.9)$$

The higher order μ terms can be interpreted as measuring deviations from lower order μ terms. For example, to compute $\mu_{AB(ij)}$, it necessitates first to compute μ, $\mu_{A(i)}$, and $\mu_{B(j)}$. This study considers log-linear models that satisfy the hierarchy principal. The hierarchy principal stipulates that the inclusion of $\mu_{AB(ij)}$ in a model means that $\mu_{A(i)}$ and $\mu_{B(j)}$ would also appear since they are lower order μ terms of $\mu_{AB(ij)}$. On the other hand if $\mu_{A(i)} = 0$ and $\mu_{B(j)} = 0$, then $\mu_{AB(ij)}$ must be zero(7,13).

A model that contains all the μ terms is called a saturated, full or complete model as in Eqs.5.2 and 5.9. From the criteria of parsimony the μ terms which do not alter the goodness of fit can be eliminated and this helps in simplifying the model structure. Then the attention focuses on selecting the best unsaturated model in which certain μ terms are set equal to zero. For example in the unsaturated model [AB][BC] , $\mu_{AC(ik)} = 0$. The [AB][BC] model for three variables A,B, and C means that conditional on B, A and C are independent, and this also implies that the associations of A with B and B with C are included in the model.

5.3.2. Log-Likelihood Ratio

The best unsaturated model is selected on the basis of likelihood ratio statistic denoted by G^2. This is given by:

$$G^2 = 2\Sigma_i x_i \ln\left(\frac{x_i}{\hat{m}_i}\right) \qquad \text{...(5.10)}$$

Where x_i is the cell count and $\hat{m}_i$ is the expected value. G^2 is asymptotically chi-square with degrees of freedom dependent on which unsaturated hierarchical model is selected. The preference for using a G^2 statistic is based primarily on certain partitioning properties that are useful in comparing two unsaturated models. Usually the most complicated unsaturated model is chosen and the unwanted μ terms are removed by backward elimination, that is by moving towards selecting lower order models(7,13,17).

5.3.3. Logit Model

The log-linear model is useful only in determining the association among the variables and the model does not make any distinction between the variables. In order to investigate the nature of the effects the explanatory variables have on a designated response (dependent) variable it is necessary to construct a logit model. The procedure is to first select a suitable unsaturated log-linear model, then the secondary analysis is conducted using a logit model.

The logit model is constructed to obtain log odds which aids in understanding the nature and the changes in the combined levels of the explanatory variables. The log odds can be better understood by an example. Consider the three variables A,B, and C, in which variable C is designated as the response variable having two levels (k=1,2). It is required for the model to give the log odds of one level when k=1 with respect to the other when k=2 of the response variable, on the basis of the explanatory variables A and B which also can have different levels (7,13,17).

Thus log odds can be interpreted as an indicator to select a specific category of explanatory variables when the level of the response variable is given.

The logit model is deduced from the selected unsaturated log-linear model for the above three variables example described earlier as follows. Let the best parsimonious log-linear model selected be [AB][AC] indicating $\mu_{BC(jk)} = 0$. The log-linear model for this selected unsaturated model based on Eq. 5.9 is:

$$\ln m_{ijk} = \mu + \mu_{A(i)} + \mu_{B(j)} + \mu_{C(k)} + \mu_{AB(ij)} + \mu_{AC(ik)} \quad \text{.............................(5.11)}$$

The corrosponding logit model is defined by:

$$\text{logit}_{ijk} = \ln\left(\frac{m_{ij1}}{m_{ij2}}\right) = [\mu - \mu] + [\mu_{A(i)} - \mu_{A(i)}] + [\mu_{B(j)} - \mu_{B(j)}] + [\mu_{C(1)} - \mu_{C(2)}]$$
$$+ [\mu_{AB(ij)} - \mu_{AB(ij)}] + [\mu_{AC(i1)} - \mu_{AC(i2)}] \quad \text{.........................(5.12)}$$

By eliminating terms,

$$\text{logit}_{ij} = [\mu_{C(1)} - \mu_{C(2)}] + [\mu_{AC(i1)} - \mu_{AC(i2)}] \quad \text{..................................(5.13)}$$

from Eq. 5.8; the following are obtained:

$$\mu_{C(1)} = -\mu_{C(2)} \;,\; \mu_{AC(i1)} = -\mu_{AC(i2)} \;,$$

Substituting

$$\text{logit}_{ij} = 2[\mu_{C(1)} + \mu_{AC(i1)}] \qquad \text{.....(5.14)}$$

$$\text{logit}_{ij} = W + W^{A\overline{C}}_{(i)} \qquad \text{.....(5.15)}$$

Where,

$$W = 2\mu_{C(1)},\ W^{A\overline{C}}_{(i)} = 2\mu_{AC(i1)}$$

The bar over the superscripted variable identifies that variable for which the odds pertain to. In this example the odds pertain to the variable C with two levels k=1 and k=2. The μ terms from which the W terms are deduced are available as part of the output. The W terms can be converted into an odds ratio, which in turn can be converted into proportions using the method suggested by Berkson(8).

5.3.4. Example

From a preliminary observation of the data, it has been found that the selection of the mode is related to the type of load and hire. The three variables considered are (L) loads -full load or less than full load(i=1,2), (H) Hire -private or for-hire (j=1,2), and (M) Modes -truck and rail (k=1,2). All the three variables have two levels. Table 24 shows the data on 7039 freight flow records cross-classified on the basis of these variables with their constituent frequency counts. Load and hire are considered as explanatory variables with mode of transport designated as the response variable for the logit model.

The saturated log-linear model is given by:

$$\ln m_{ijk} = \mu + \mu_{L(i)} + \mu_{H(j)} + \mu_{M(k)} + \mu_{LM(ij)} + \mu_{HM(jk)} + \mu_{LM(ik)} + \mu_{LHM(ijk)} \quad \text{....(5.16)}$$

Table 25 shows the possible combination of unsaturated log-linear models M_1 to M_{17} as obtained from the computer output and two good fitting models indicated by asterisks(*) are selected for a level of significance $\alpha = 0.05$. It is necessary to select the 'best-fitting' model out of the two selected models. For this purpose, the likelihood ratio statistic G^2 in Eq. 5.10 is used.

The [LH][LM][HM], model M_{17} (where $\mu_{LHM(ijk)} = 0$) is compared with [LM][HM], model M_{15} (where $\mu_{LH(ij)} = 0$) by making the following computation.

$$G^2\left(\frac{M_{15}}{M_{17}}\right) = G^2(M_{15}) - G^2(M_{17}) = 0.77 - 0.62 = 0.15 \quad \text{........(5.17)}$$

Since $G^2\ (M_{15} / M_{17}) = 0.15 < \chi^2_{0.95;\,1} = 3.841$, this indicates that there is no substantial difference in fit between the two models. However from the point of parsimony, the model with fewer variables is selected. Therefore, the [LM][HM] model M_{15} is chosen and the extra [LH] term is deleted.

This model includes the association of 'loads' and 'hire' individually with 'modes', and this association indicates that selection of mode is related to both loads and hire. The log-linear form for the selected unsaturated model [LM][HM] is given by:

$$\ln m_{ijk} = \mu + \mu_{L(i)} + \mu_{H(j)} + \mu_{M(k)} + \mu_{LM(ik)} + \mu_{HM(jk)} \quad \text{.....................(5.18)}$$

The second part of the modeling is to construct a suitable logit model and to develop a table of log odds to understand how changes in the combined levels of explanatory variables affect the response variable. The log-linear form for the selected model is as represented in Eq. 5.17.

The logit model is defined by:

$$\text{logit}_{ij} = \ln\left(\frac{m_{ij1}}{m_{ij2}}\right) = [\mu_{M(1)} - \mu_{M(2)}] + [\mu_{LM(i1)} - \mu_{LM(i2)}] + [\mu_{HM(j1)} - \mu_{HM(j2)}] \quad \text{..(5.19)}$$

Due to constraints imposed on interaction terms,

$$\mu_{M(1)} = -\mu_{M(2)},\ \mu_{LM(i1)} = -\mu_{LM(i2)},\ \mu_{HM(j1)} = -\mu_{HM(j2)}$$

Therefore,

$$\text{logit}_{ij} = 2[\mu_{M(1)} + \mu_{LM(i1)} + \mu_{HM(j1)}] \quad \text{.....(5.20)}$$

and thus,

$$\text{logit}_{ij} = W + W^{L\bar{M}}_{(i)} + W^{H\bar{M}}_{(j)} \quad \text{.....(5.21)}$$

where, $W = 2\mu_{M(1)}$, $W^{L\bar{M}}_{(i)} = 2\mu_{LM(i1)}$, and $W^{H\bar{M}}_{(j)} = 2\mu_{HM(j1)}$

and the bar over the superscripted variable identifies the variable to which the odds refer. In this example the odds pertain to modes that have two levels k=1 (truck) and k=2 (rail). The μ terms from which the W terms can be computed are obtained as part of the output from many standard statistical package programs. For this analysis, the BMDP-4F computer statistical package (Computer Program III, Appendix II) is used.

Table 26 shows the estimated effects and logits obtained for the selected [LM][HM] model. The W term is positive for LTFL as compared to FL, indicating that trucks are preferred more often than rail for LTFL's. The odds ratio for either truck or rail for a particular combination of load and hire can be calculated using the W terms. Thus for FL (i=1) and LTFL (i=2) with type of hire as private (j=1) and for-hire (j=2) from Table 26 and Eq. 5,20:

$$\text{logit}_{ij} = \ln(\frac{\hat{m}_{1j1}}{\hat{m}_{1j2}}) = 4.798 - 1.24 - 1.86 = 1.698 \quad \text{.........................(5.22a)}$$

$$\text{logit}_{ij} = \ln(\frac{\hat{m}_{2j1}}{\hat{m}_{2j1}}) = 4.784 + 1.24 - 1.86 = 4.164 \quad \text{.........................(5.22b)}$$

The odds ratio is defined as

$$\text{Odds Ratio} = e^{\text{logit}_{ij}} = \frac{\hat{m}_{ij1}}{\hat{m}_{ij2}} \quad \text{.........................(5.23)}$$

$$\text{Odds Ratio (FL)} = e^{1.698} = 5.46 \quad \text{...............(5.24a)}$$

$$\text{Odds Ratio (LTFL)} = e^{4.164} = 64.32 \quad \text{...............(5.24 b)}$$

The odds ratio can be interpreted as, for FL trucks are preferred 5 times more than rail and for LTFL, trucks are prefered 64 times more than rail, neglecting the decimals.

The odds ratio is converted into a proportion using the transformation suggested by Berkson (1944) as:

$$\hat{P}_{ij} = \frac{e^{\text{logit}_{ij}}}{1 + e^{\text{logit}_{ij}}} \quad \text{.........................(5.25)}$$

thus for FL,

$$\hat{P}_{1j} = \frac{5.46}{1 + 5.46} = 0.84$$

for LTFL,

$$\hat{P}_{2j} = \frac{64.32}{1 + 64.32} = 0.98$$

Therefore, it can be estimated from the model [LM][HM] that 84% of FL and 98% of LTFL would prefer trucks over rail, with type of hire as private. Similarly, 99.5% FL and 99.96% LTFL would prefer trucks over rail with type of hire as for-hire.

The goodness of fit for the selected model can be verified from the standardized residuals for the model, defined as,

$$\text{Standardized Residuals} = \frac{\text{Observed Counts - Fitted Counts}}{\sqrt{\text{Fitted Counts}}} \quad \text{.........(5.26)}$$

Table 27 shows the standardized residuals and they are within acceptable limits. The standardized residuals are of concern only if the absolute value of the residuals are in excess of 2.0 , although one or two such cases in large tables should not be of concern (7,13,16,17). Finally, the model selected should be able to explain the physical process observed. In Table 24 it is observed that trucks dominate over rail in both classes of loads, and the model substantiates this observation.

5.4. ANALYSIS OF MODE SELECTION

Different models were developed in order to determine the association and influence of explanatory variable(s) alone or together with other variable(s) having different levels on the selection of a mode. Table 28 lists the different models selected based on the procedure explained . It has to be noted that the odds effect could vary with the number of variable(s) and levels of the variable(s) used in building these models. Model M1 in Table 28 refers to the [LM] [HM] model described under model building procedure. The standardized residuals for this model are shown inTable 27 which range from -0.3 to 0.2 and are within acceptable limits (7,13,17).

The questionnaire to the shippers asked if they decided the mode of transport for a commodity. In this analysis, this is called a control factor. The shipper would indicate his control by "yes" or "no" on the survey forms. It is of interest to determine the preference for a mode knowing the type of control the shipper has in transporting the commodities. Model M2 (Table 28) is the selected unsaturated model [CL][CM][LM], where (C) is control - yes or no (i=1,2), (L) is load - FL or LTFL (j=1,2) and (M) is modes - truck or rail (k=1,2). The data, degrees of freedom, and estimated effects for this model are shown in Tables 29,30 and 31 respectively. The model indicates the association of 'control' and 'load', 'control' and 'mode', and 'load' and 'mode', that is mode selection depends on control and load type. The W term is positive for 'no' in this model indicating that truck mode is preferred by receivers and that the shipments are generally asked to be sent by truck mode. From this model it is estimated that 91.3% of shippers who answered 'yes' (having control) and 99.26% of shippers who answered 'no' (having no control) prefer trucks for FL. Similarly 83.7% shippers with control and 98.5% shippers without control would prefer trucks for LTFL. These inferences can also be made during preliminary analysis by plotting the logits for the data used in developing the models as shown in Figure 19. The preference for truck mode seems to be higher with control (yes) for both types of loading. Also it is observed that the preference for trucks with LTFL is more dramatic when compared to FL. This model gives an insight into the decision making process of selecting a mode and indicates that economy is considered by shippers in selecting a particular mode in transporting commodities. The standardized residuals are as shown in Table 32, ranges from -0.5 to 0.2.

Model M3 is developed to study the relation and the influence of average shipment size and loads on the preference for a mode. In constructing this model, average shipment size up to 60 tons only are considered. Model M3 (Table 28) is represented by [SL][SM][LM] where (S) is shipment size - ranging from less than 1 ton to 60 tons (i=1,2....11), (L) is load - FL or LTFL(j=1,2), and (M) is modes -truck or rail (k=1,2). The data, degrees of freedom, and estimated effects for this model are shown in Tables 33,34 and 35 respectively The model includes the relation between 'size' and 'load', 'size' and 'mode',and 'load' and 'mode', indicating that average shipment size of the commodity and type of load have influence in the selection of a mode. As expected, the W terms are positive in Table 35 for average shipment size up to 30 tons indicating the preference for truck mode in this range. The W terms are negative for average shipment size above 30 tons, indicating that rail mode is preferred at higher average shipment size. Similarly the W terms for FL and LTFL indicate the preference for a particular mode given the type of load. The W term is positive in Table 35 for LTFL for trucks, indicating the preference for trucks under this category. Table 36 shows the mode preference by proportions for FL and LTFL under different average shipment sizes. Trucks totally dominate the movement of commodities under less than 30 tons of average shipment size for both FL and LTFL. They capture almost 100% of the market share in the LTFL for average shipment size less than 30 tons. The standardized residuals are shown in Table 37, they range from -0.3 to 1.9, with the exception of one value which is 2.1 for rail mode under LTFL category and is within acceptable range (7,13,17).

The identification of the different types of commodities by various modes is indeed difficult as there are several types of commodities and it is one of the major areas where urban and commodities transportation planning differ. The large number of commodities complicates the problem of identification. The major commodity classifications in Table 2 are considered for identification and they are expected to cover most of the different of commodities transported across the province. The identification of the type of commodity gives a clear indication about the preference for a mode, given the commodity characteristics and type of load.

Model M4 (Table 28) is represented by [KL][KM][LM] where (K) is commodity type - major commodity classifications considered are 15(i=1,2,.....,15), (L) is load - FL or LTFL (j=1,2), and (M) modes -truck or rail (k=1,2). The data, degrees of freedom, and estimated effects for this model are shown in Tables 38,39 and 40 respectively. The model includes the relation between type of 'goods' and 'load', 'goods' and 'mode', and 'load' and 'mode' indicating that the type of commodities and loads have influence in the selection of mode.

The sign of W terms in this model helped to identify the preference for a particular mode given the classification of the commodity. The W terms are negative in Table 40 for some of the major commodities such as general freight, heavy machinery, metal products, bulk liquids, bulk dry, forest products and feed products, indicating that both truck and rail mode are preferred in transporting them. This observation is well supported by the truck mode proportions, as shown in Table 41, obtained for the model with FL and LTFL. As expected, only a small percentage of the FL market for transporting the previously listed commodities is shared by rail mode. A similar observation is made in model M1 (Table 28) represented by [LM][HM], where one of the conclusions drawn is that under FL the rail shares a small percentage in the movement of commodities. From Tables 36 and 41 the proportions for trucks with LTFL indicate that trucks are exclusively used in transporting all types of commodities under different average shipment sizes. It is also observed that commodities such as food non- perishables, food perishables, petroleum products, and livestock, are transported by trucks under both FL and LTFL. This could be expected since the movements are all intra-provincial and the mean distance travelled is little over 600 kilometers between the O-D pairs. The standardized residuals for this model is shown in Table 42, ranges between -1.9 to 1.0.

Distance is considered as an important parameter in modal split analysis. It is used as spatial seperation factor in the development of distribution models, particularly in gravity model. The type of mode used is highly influenced by the distance the commodity is to be transported as seen from the earlier modal split analysis(4,27,29). The utilization of truck or rail taking into consideration the type of commodity to be transported and its average shipement size, distance between the origin and destination is an important factor considered by the transporter. The selection of private or for-hire is influenced by the distance. Similarly, the cost of transportation is tied with the distance the commodity is to be carried. Model M5 (Table 28) is represented by [DK]KM][DM] . Where, (D) is the distance between origin and destination in kilometers (i=1,2,3,4); (K) is the six commodity category (j = 1,2,3,4,5,6) shown in Table 3; (M) modes - truck or rail (k=1,2). For this model development 1927 O-D pairs were considered and distance is categorised into four ranges(0-200,201-400,401,600, and >600) as indicated in Table 45. The G^2, degrees of freedom and, estimated effects for this model are shown in Table 43 and 44. This model includes the relationship between 'distance' and 'commodity ', 'commodity' and 'mode', and 'distance' and 'mode' indicating that the selection of mode is dependent on the distance and commodity transported.

The sign of the W terms in this model in Table 44 are both positive and negative, indicating that the selection of mode is dependent on both distance and the commodity transported. Table 45 shows the proportion of trucks used in transporting the six categories of commodities. From the table it is evident that trucks are exclusively used to transport commodities across the province. The standadized residuals for this model are shown in Table 46, ranges between -0.7 and 1.9, with the exception of two values which are 2.3 and 2.4 (7,13).

5.5. SUMMARY AND CONCLUSIONS

5.5.1. SUMMARY

This study of modal split analysis has shown the extent to which some identified factors can be used for explaining the selection of a particular mode to transport commodities. The identification of the dominant mode(s) aids in making useful decisions about future requirements in the provincial transportation network. Some of the factors have not been used in similar previous studies. Although many other factors could be considered in doing analysis of this nature, unfortunately they are omitted due to the non-availability of such data. Also collection of commodity data has not yet been standardized and collected in the same manner as passenger data.

This modeling has presented a comprehensive and statistically credible method by which the vast data necessary for commodity transportation planning can be analyzed using multidimensional contingency tables. The log-linear model aids in understanding the associations among the variables. Whereas the influence of selected explanatory variables over a designated response variable could be studied by using a logit model. This method facilitates in selecting variables for modal split models and in studying the influence of variables alone or together with other variables.

There are two major advantages in using logit models. The first is that the estimated probability (proportion) of mode choice always lies within the range of 0 to 1. The second is that the sum of the probabilities of the choice of each of the mode alternatives is equal to one. From the analysis it is observed that the selection of a mode is a function of several factors such as the average shipment size, type of load, hire, control, and type of commodity. The signs of the W terms in logit models are very useful in uncovering the complex relationships that may exist among the explanatory variables and designated response variables having different levels, which is particularly true in the case of commodity transportation.

5.5.2. CONCLUSIONS

The following are the main conclusions of this study:

1. The identification of places into sources (points of production), sinks (points of consumption) or both facilitates the collection of the necessary commodity data by interviewing either shippers or consignees during the survey.

2. The average shipment size and commodity type give an indication as to the selection of a mode. Trucks and rail dominate the movement of all commodities. Rail shares a very small percentage in the movement of specific commodities above 30 tons.

3. The classification of commodity distribution by FL or LTFL helps to recognize that shippers consider the economy in terms of transportation cost, average shipment size, and distance in choosing a mode. Trucks carry more LTFL and rail carry more FL, and vice vesa. This indicates the different characteristics displayed by the two modes.

4. There are more commodities carried by private carriers than for-hire. The greater use of private carriers for transporting LTFL indicates the large movement of commodities within the province between sources and sinks. The significant movement of commodities by LTFL indicates the demand for this service in the regional market sector.

5. The control factor gives an insight into the preference and the decision making process in shipping different commodities, depending on the average shipment size and other factors perceived as important by the shipper or consignee.

6. There is very limited use of all other modes listed in Table 4. The Pipeline mode is used exclusively for the movement of petroleum products. Air

mode is used mostly for LTFL, with average shipment size less than 2 tons. Marine mode is used occasionally at higher the average shipmentsize to move general freight, petroleum products, and forest products.. The bus and other modes are generally used to move few specific commodities less than 2 tons of average shipment size.

9. The log-linear and logit model methodology suggested is a statistically valid tool for analyzing the vast commodity data necessary for transportation planning and in studying the association of variables and their influence on a designated response variable.

10. The models developed give an insight into the preference for a mode based on a set of explnatory variables at different levels.

11. The proportions estimated using logit models help to determine the percentage share of different modes and the model helps to identify the commodities shared by different modes.

12. The dominance of truck mode over all other modes is clearly evident from the analysis. One could expect an increase in market share for trucks in carrying specific commodities over larger average shipment size if there are any significant changes in the vehicle size and weight in the near future.

13. It is difficult to say from this analysis if the market of rail mode will increase. It will depend entirely on the perception of shippers or consignees as to the economy and reliability of this mode.

CHAPTER 6 : APPLICATION OF DEVELOPED MODELS

CHAPTER 6

APPLICATION OF DEVELOPED MODELS

6.1. INTRODUCTION

The study of intra-provincial commodity flow has emerged due to the result of increasing trade and commodity flows within the provincial boundaries. The objectives of these studies are to explain the rationale behind intra-provinical trade. The first major reason for the movements of commodities arises from the fact that there exists natural differences in the availability of resources between regions. Primarily the economic forces make the flow of commodities between two regions, this inspite of the spatial seperation between them. The geographical imbalance and the wide variation of commodity requirement necessitates planning if the commodities are to be distributed equally and efficiently to all regions across the province. Thus planning models are necessary with optimal forecasting capabilities. The model must be able to simulate the observed flows so that the existing transportation structures in the network can be improved for greater efficiency. An intra- provincial distribution model would have direct use for regional planners. It would enable them to project future commodity flows on the intra-provincial links and study the impact on regional economic growth(33,34,35).

The demand for transportation is a derived demand. The flow of commodities exists because production and consumption do not occur at the same place for most commodities. The character of the commodity flows of any region depends on the characteristics of the producers and consumers of the commodities and the spatial separation between them. Any change in the socio-economic characteristics of a place will have an impact on the pattern of flow. The flow of commodities is accomplished by the transportation system, which must be sufficiently variable to provide the specific requirements associated with the movement of any commodity. The extreme complexity of the commodity flow is due to the unique transportation requirements associated with each commodity type. Firms usually tend to minimize their costs for any given service. The pattern of modal share throughout a region will depend on the pattern of demand for commodities. In other words, the demand for various types of transportation services will depend on the specific mix of commodities for which the services are required. Thus, any attempt to forecast tranportation needs for a region must be based on anticipated flows of different commodities throughout the region(11).

Table 1 gives the list of selected communities with size (city, town , and village), population range, and the economic base of the communities. An examination of the pattern of production and consumption based on the above factors are good predictors of the zonal commodity requirement at the aggregate level. In Table 1 there are a few smaller cities and two major cities Edmonton and Calgary. More than 50 % of the population is situated in these two cities out of a total population of a little over 2 million. The smaller cities such as Camrose, Fort McMurray, Grand Prairie, LethBridge, Medicine Hat, Loyd Minister and RedDeer act as regional distribution centres (RDC). Now, consider the distances of the sinks (demand points) from the RDC's and the two major cities, the distribution of commodities takes place from the nearest source as expected. This is further substantiated by visual inspection of the O-D commodity flow tables for the six categories given in Tables 5 to 10. The flow of commodities from Edmonton and Calgary are very large as compared to the flow of commodities from all other origins. Also, it is interesting to note that the flows divide between Edmonton and Calgary depending on the distance from each other to the demand points. The conclusion that could be drawn from the above observations is that from studying the community characteristics and the distance form the nearest source, gives an indication as to the direction of flow of traffic in the provincial network.

The next step is to identify the type of commodity flow between the O-D pairs. This could be accomplished to a certain degree by checking the major economic base at the region. The major economic base of the selected communities are given in Table 1. This factor gives a good idea as to the type of commodity flow expected on a particular linkage. The identification of places into sources, and sinks as indicated in Fig.4 facilitates in collecting necessary data by interviewing either shippers or consignees depending on the type of community.

For a transport planner it is necessary that the calibrated models stimulate most accurately the larger flows within the system. That is the planner is interested in identifying the network where large tonnages are carried. These paths represent the areas where increased transportation efficiency will have to be implemented. A distribution model must be able to simulate these large flows so that the planner's attention will be drawn in that region. Also, it is necessary that the developed model be able to replicate smaller flows without inflating them to a very large extent, so as to avoid disutility of available resources in the transportation network.

The distribution model constructed in this study is based on observed tons of flow between O-D pairs. Before developing the distribution models , the commodity classification in Table 2 were condensed to form six commodity categories as shown in Table 3. The commodity categories were achieved by grouping commodities of similar characteristics. The developed optimized gravity model is as defined in Eq. 4.15 . A set of optimized gravity models are constructed for each commodity catgory in Table 3 and are shown in Table 15.

The results of the calibration process of the developed model has been tested by different methods such as, correlation coefficient, coefficent of determination, and commodity haul frequency diagram. All these tests have proven that the developed models are capable of duplicating the intra- provincial commodity flow given reliable data. Take for example the correlation coefficient between the estimated and the observed flows. They range from a low of 0.84 for Food Stuffs to 0.93 for Forest and Heavy Machinery.. These high values indicate that a great deal of the variation in the estimated flows is explained by variations in the actual flows. Hence, it can safely be assumed that the estimated flows are low when the actual flows are low and the estimated flows are large when the actual flows are large. The same could be said about the high R^2 values obtained for the optmized condition.This is substantiated by the commodity haul frequency diagrams for the six commodity categories shown in Figs. 10,12,14,16,and 18. Here, we observe that the frequency diagram of the estimated flows for all the commodities generally follow the frequency diagram of the observed flows.

Some of the discrepency in the commodity haul frequency diagram can be directly attributed to the data, and the cross-shipments inherently created by the use of gravity model. It is a fact that the gravity model concept creates cross-shipments between all origin and destinations. For example, if a destination has received any commodity from any origin, the model will invariably predict that it will receive commodities from all the other origins in the O-D commodity flow matrix. In real world situations, cross-shipments do occur, but not as much as in the gravity model.

From the above arguments it is reasonable to assume that the developed models are capable of simulating the intra-provincial commodity flows, provided reliable data input is used. The transportation planner could depend on the distribution pattern given by the model.

The next step in transportation planning after being able to distribute the total flows across the province using a distribution model is to understand the modal split between the O-D pairs. It is important to know the movement of commodities by different modes so that estimates may be made of the traffic demands on the transportation network. From Table 4 it is evident that the dominating modes of transportation in the province are the truck and rail mode. A closer look at the available data revealed that trucks completely dominate the movement of all types of commodities across the province under all categories of loads (FL or LTFL), hire (private or for-hire), and control (yes or no). Rail is only significant at very large average shipment size and is used to transport bulk materials over large distances. The majority of the commodities below 30 tons are exclusively carried by trucks, and these loads can be either FL or LTFL. Tables 16 to 22 show the modal split across the province with respect to various parameters.

The results of the selected log-linear and logit models in Table 28 substantiates these above observations and also aid in calculating the percentage of modal share among competing modes with different levels of the parameters. Consider Table 41; which gives the mode preference by commodity classifications and loads (FL or LTFL). It is observed that trucks dominate the movement of commodities under both types of loads. The W-terms of the logit model in Table 40 help to identify the specific commodities shared by the competing modes, here truck and rail. As expected the most commonly shared commodities by truck and rail are General freight, Heavy machinery, Metal products, Bulk Liquids, Bulk Dry, and Forest Products.

6.2 APPLICATION OF THE DEVELOPED MODELS

Application of the developed distribution and modal split models is illustrated in the following section with respect to the stages of transport model developement as shown in Fig.3a. This study is concerned with the trip distribution and modal split stages of the transport modeling. The trip generation stage was conducted by the Alberta Transportation Department. In this stage data that were required to develop the O-D matrix were collected as discussed in Chapters 3 and 4.

Considering the the distribution stage of transport modeling the following steps are necessary to use the developed optimized distribution models.

Step 1: First it is necessary to obtain interchanges(tons) between identified origin and destinations for different commodity categories across the province under the trip generation stage. Also, the modes used to transport the goods are to be recorded.

Step 2: This involves the development of the observed O-D matrix(Commodity Flow Tables) for each of the commodity category. At this stage it is necessary to decide which mode or modes combinations will be used in developing the O-D matrix has to be decided. The interchanges are entered in the commodity flow table cells as shown in Tables 5 to 10. In this study the actual annual tonnage of shipment between origin and destinations are used in the CFT's.

Step 3: It is necessary to determine the basis for estimating the spatial seperation factor. In this study the highway distance between the origin and destination is used as the spacial seperation factor. The distance matrix are developed as shown in Tables 11 and 12.

Step 4: From the developed O-D matrix, the input values such as the P_i's and C_j's are obtained. The d_{ij} values are obtained from the distance matrix. The only input necessary at this stage is the exponent value of the spatial separation factor d_{ij}.

Step 5: Selecting an appropriate exponent value λ to use in the distribution model. This could be done by trial and error procedure. The exponent values to be used in this study range from -2.0 to 0.0 with increments of 0.25 as shown in Table 14.

Step 6: For example let us consider one commodity category namely food stuffs. the purpose here is to estimate the interchanges between origin and destination s from the developed optimized model for food stuffs shown in Table1 5.

Step 7: The optimized gravity model for food stuffs is given by:

$$(T_{ij}) = P_i \frac{C_j \; d_{ij}^{-0.5}}{\Sigma_j \, C_j \; d_{ij}^{-0.5}} \qquad \text{..................(6.1)}$$

In this model the exponent value is found to be equal to -0.5 satisfying the objective function used in Eq.4.11. The procedure for determing the optimum exponent value is given in chapter 4 under calibration procedure.

Table 47 shows estimated interchanges as determined using the model in Eq.6.1 for few of the destinations with Edmonton and Calgary as origins. Thus using the optimized gravity models in Table 15 for each commodity category it is possible to develop estimated O-D matrices. From the calibration procedure and objective function adopted in Chapter 4, the estimated interchanges will be close to the observed interchanges for the optimized gravity models.

The third stage of the transport modeling is the modal split. At this stage the estimated flows from the optimized gravity models are to be distributed among competing modes. The modal split models constructed using logit models can be used to distribute the estimated interchanges.

For example, consider the modal split model [DK][KM][DM] in Table 28. This model contains parameters distance(D), commodity category(K), and modes(M). In this model modes is the dependent variable, commodity category and distance are the explanatory variables. Using the logit model as described under Chapter 5 it is possible to determine the modal share proportion between truck and rail. The modal share proportions are given in Table 45 for the various distance ranges and commodity categories. Using these proportions, the estimated interchanges can be divided among the competing modes truck and rail in the province. The estimated interchanges are divided between truck and rail according to the modal share proportions obtianed for the [DK][KM][DM] , the commodity category and the distance range are also considered in using the proportions. The estimated tirps are divided as shown in Table 47 for food stuffs.

The last stage in the devleopment of transport model is traffic assignment. This could be achieved by drawing desire line diagrams of the estimated interchanges along with the modal share on each link as shown in Fig. 20. The desire lines drawn in Fig. 20 shows the estimated interchanges and the modal share for food stuffs. Such desire line diagrams can be drawn for all the commodities in order to understand the pattern of commodity movements across the province using the develoed distribution and modal split models.

For traffic assignment, the estimated interchanges could be converted into number of trips along each link knowing the average shipment size carried by each mode. Thus the developed optimized gravity models and modal split models aid in constructing desire line diagrams for the various commodity categories across the province . The desire line diagrams thus drawn are practical tools to understand the pattern of commodity movements across the province.

The major problem in commodity transportation planning is the availibility of reliable data. Even though the provinces are responsible for highway system, the development of forecasting techniques and data bases specifically dealing with commodity movements by highway and railway system has been slow to come. In most provinces, the collection of truck traffic flow data, and the passenger vehicles data are combined, but not as seperate data set(34,35,36). The truck data problem is due to the inability to easily seperate automobilies and trucks when using automatic traffic-counting equipment(and the consequent necessary to count and classify trucks manually), has mitigated against developing seperate data bases for trucks and automobiles. The lack of commodity flow data pertaining to truck and rail movements is particularly critical to developing a commodity forecasting model(34,35,36). It becomes important to use existing data resources, since the avalibility of adequate dta from government and other resources is unlikely. In most provinces, future truck volumes are forecast as a percentage of aggregate traffic volumes for both existing and proposed facilities using historical data(36). Thus forecasts are made using trend extensions rather than by taking into consideration the observed commodity flows with present economic activities.

The distribution and modal split models developed in this study can be applied to other provinces if simialr data as collected by Alberta Transportation is available. The main inputs to the distribution model are the P_i's, C_j's and the d_{ij}'s. In this study the acutal tonnage is used to devleop the O-D matrix for the distribution model. In case of the modal split models, the methodlogy can be used with parameters perceived to be important for modal share across the province and develop suitable modal split models depending on the data and requirements. The problem that may be faced in using the distribution model is the value of the spatial separation factor exponent λ. The optimized gravity models can be directly used in another province provided that the observed O-D matrix in the other province is similar in nature with respect to commodity category and dimensions of the matrix. Thus it is possible to utilize the methodology in this study in developing distribution and modal split models for other provinces.

CHAPTER 7 : SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS FOR FUTURE RESEARCH

CHAPTER 7

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

7.1. SUMMARY AND CONCLUSIONS

In transportation planning, there are four major stages. They are trip generation, trip distribution, modal split, and trip assignment as shown in Fig.3a. This study is concerned with the trip distribution and modal split stages of commodity transportation planning. The data collection done by Alberta Transportation comes under trip generation, and hence that stage was not considered, except for checking the data and using it for further analysis. The last stage, namely trip assignment was not considered as there was not sufficent data available at the present time to conduct such a study.

The main data set available for analysis is based on the commodity flow survey conducted by Alberta Transportation in 1977. There are 18 major commodity classifications with several minor classifications. For this reason, the the data is well structured to contain almost all of the commodities being tranported at an reasonably aggregate level. The fact that 1977 data is used should not make much difference, because it is assumed that spatial distribution of population centres and investment in transportion facilities are stable over long periods(11). Also, the change in successive governments as it happens most often have difficulty in implementing major changes in all sectors of a transportation system. Another important reason why this data was used for the analysis was, that no other commodity flow data as this was not available for this type of study. Conducting commodity flow survey of this magnitude requires large man power and funds. Thus the results of this study are valid even though a decade has passed(11).

Chapter 4 deals with the commodity distribution models. A set of optimized gravity models, one for each category of commodity is proposed to represent the commodity flow within the province of Alberta. In this study only the intra-provincical commodity movements are considered. The O-D commodity flow tables are developed by categorizing the commodities into manageable size by combining commodities with similar characteristics. The O-D tables were constructed by considering the modes of transport, origin and destination pairs, commodity categories, and the observed annual tonnage. The developed O-D tables are the basis for the developement of optimized gravity models for commodity flows. A power function and LSM are used as the optimization technique.

The developed models use distance between the O-D pairs as spatial separation factor. Each of the six commodity categories has a different spatial seperation factor. A commodity haul frequency diagram is drawn to compare the estimated and observed flow tons. A visual inspection of the graphs shows that there is a general agreement with the estimated and observed flows. The results of this study show that the developed optimized gravity models are consistent with the results expected normally obtained for commodity transportation models. The calibration procedure adopted in this study has a potential for further applications in the development of commodity distribution models.

Chapter 5 deals with the next stage of transportation planning, that is the modal split analysis and models. In chapter 4 the commodities were distributed across the province using a distribution model, the next question was to determine how these commodities are transported across the province?. This study was conducted in two stages. The first stage consisted of modal split analysis of the commodity flow survey data and the second stage consists of constructing models.

The modal split analysis of commodity flow survey data was carried out with respect to major carriers. The study analyses the role of major carriers in the distribution of commodities within the province. The factors used are average shipment size, type of commodity, control (yes or no), loads (FL or LTFL), and hire (private or for-hire) and the distance between O-D pairs. Some of the factors used in this study have not been used in previous studies of this nature. This modal split analysis clearly showed that there exists a definite pattern in the movement of commodity across the province by different modes. As expected truck and rail dominated over all other modes. The role of each mode is well defined with respect to various factors. For example, trucks as private carriers were used to transport majority of LTFL whose average shipment size is less than 30 tons. Rail is used to transport FL of specific commadities whose average shipemt size are above 30 tons. The other modes carried insignificant percentage of commodities as compared to truck and rail mode.

In the second stage modal split models were constructed using log-linear and logit models. The main objective of this study is to propose a more comprehensive and statistically credible method to analyze the vast data required in transportation planning. It involves the application of standard statistical techniques such as the log-linear and logit models. The data are collapsed to form multidimensional contingency tables in order to develop these models. The log-linear models helps to understand the associations among the variables, whereas the influence of selected explantory variables over a designated response variable. From the developed models, it is observed that the selection of a mode is a function of several factors. The signs of the W-terms in the logit models are used to uncover complex relationships that may exist among the explanatory variables and the designated response variable (dependent variable) at different levels.

Chapter 6 deals with the applicability of the developed distribution and modal split models in commodity transportation planning. The distribution model has proved that it is capable of distributing the commodity interchanges (tons) between origin (source) and desintation (sink) using the optimized gravity model technique. There is a different spatial seperation factor d_{ij}^{λ} for each commodity category. The results of the distribution are checked using standard statistical tests and by drawing the commodity haul frequency diagram. Similarly the modal split analysis and modal split models are useful in assigning the commodities to different modes, here truck and rail. The proportion of haulage in each link by a particular mode can be calculated knowing various factors such as the distance between the origin and destinations, the commodity category, etc. Thus it is feasible to distribute the commodity interchanges with respect to different modes within the province using the developed models. The reliability of the data used for this analysis is described in Chapter 3.

The major conclusions of this study to the transportation planner are the following:

1. A set of optimized gravity models, one for each type of commodity category is developed to represent the commodity flows on a province wide basis. Each developed model has a unique spatial seperation factor d_{ij}^{λ}. The optimum value for λ is determined by an optimization technique, which uses power function and least squre method.

2. The value of the spatial seperation factor exponent λ varies from a low value of -0.25(low) for highly processed commodity petroleum products and -1.0(high) for bulkier commodities such as heavy machinery and bagged products. These values are consistent with the exponent values of other researchers in similar studies.

3. The distribution models have variance ranging from 71% for food stuffs and 88% for forest proudcts, indicating a high degree of reliability in the developed models.

4. From the commodity haul frequency diagrams it is evident that the developed models are capable of duplicating the observed flows. Also, between 60 and 70 % of the interchanges in the province take place between 100 and 400 kilometers.

5. The calibration of the developed models using the optimization technique has potential for further applications.

6. Trucks carry exclusively all commodities under average shipment size below 30 tons. Rail carry commodities whose average shipment size is above 30 tons.

7. Trucks carried 88.2% of all commodities, with 6.5% by rail and the rest shared by all other modes.

8. Trucks carried 55.3% of all less than full loads and 44.7% of full load across the province. Rail carried 9.5% of less than full load and 90.5% of all full loads.

9. Private vehicles were used for all regional movements of commodities with a high percentage of 79.26 . Whereas 20.74% all movements were by for-hire.

10. Truck mode is used 41.5% shipments with control, 58.5% shipments without control. Rail is used for 30.9% of all shipments with control, and 69.1% shipments without control.

11. Rail shares a very small percentage of the service market and is used to move specific commodity category of heavy loads over longer distances. There is very limited use of all other modes.

12. The signs of the W-terms in logit model are useful in uncovering complex relationships that may exist among the explanatory variables and a designated variable (dependent) having different levels.

13. The modal split proportions obtained from the models is useful to assign the percentage share of commodity transported by different modes.

14. Finally, the developed distribution and modal split models are useful to commodity transportation planners in desire line diagrams and in understanding the pattern of distriubtion of comodities across the province along with the modal share on each link for a partiuclar commoditiy category.

7.2. RECOMMENDATIONS FOR FUTURE RESEARCH

There is always a search for improving an existing system, especially it is true in transportation. Long range planning is an essential ingredient in advancing the tranportation efficiency in the transportation net work. However, such planning with respect to commodity flows have been hampered by lack of data. Commodity transportation has been growing at a rapid pace due to the migration of people to new places and also the increase in population. This continuous growth has put a new strain on the national and regional planners. Considering these facts, there is a need for further research in the following areas with respect to commodity transportation planning.

(1). It is necessary to develop a standard commodity data collection format similar to the one adopted in urban transportation planning for better understanding of commodity movements and for developing reliable forecasting models. Also, the characteristics of different commodities should be determined and categorized to reduce them into manageable size for analysis and modeling. The models should be constructed for the commodity or commodities in question.

(2). Research into statistical validation of developed O-D tables are to be conducted. Methods to compare the estimated flows and observed flows using different parameters are to be developed.

(3). There is a need to improve the basis on which the production, consumption at a place could be predicted using a comprehensive data base such as employment data, economic, demographic and other necessary data available at a place.

(4). In future greater emphasis should be placed in improving the existing highway network by considering all aspects of passenger and commodity transportation.

REFRENCES

REFERENCES

1. Amemiya, Takeshi. (1981)."Qualitative Response Models: A Survey." Journal of Economic Literature. Vol. x1x, pp.1483-1536, USA.

2. American Association of State Highway and Transportation Officials. (1977). "A Manual on User Benefit Analysis of Highway and Bus-Transit Improvements." Washington, D.C, U.S.A.

3. Ashtakala, B., and Murthy, A.S.N. (1986). "Intra-Provincial Trip Distribution Model." Canadian Journal of Civil Engineering, National Research Council of Canada, vol.13, No.4, pp. 438-444, Ottawa, Canada.

4. Ashtakala, B., and Murthy, A.S.N. (1987). "Modal Split Study of Commodity Movements in Alberta." Procedings of the C.S.C.E Centennial conference,Vol. 2, pp. 530-547, Montreal, Canada.

5. Ashtakala, B., and Murthy, A.S.N. (1987). "Optimized Gravity Models for Commodity Transportation." accepted, Journal of Transportation Engineering, American Society of Civil Engineering, New York, USA, July 1988.

6. Barker, Henry H. et.al. (1981). "From Freight Flow and Cost Patterns to Greater Profitability and Better Service for a Motor Carrier." Interfaces, International Journal of the Institute of Management Sciences,Vol.II, No.6, U.S.A.

7. Berenson, M.L.et.al. (1983). "Intermediate Statistical Methods and Applications-A Computer Package approach."Prentice-Hall Inc., Engelwood Cliffs, New Jersey, U.S.A.

8. Berkson, J. (1944). "Applications of Logistic Function to Bio-Assy." Journal of the American Statistical Association, Vol.39, pp.357-365, U.S.A.

9. Black, W.R. (1971). "The Utility of the Gravity Model and Estimates of its Parameters in Commodity Flow Studies." Procedings of the Association of American Geographers, Vol. 3, pp. 28-32, U.S.A.

10. Canadian Transport Commission. (1975). "Report on Transport Demand, Forecasting Program: Phase One". Working Paper No.6, Ottawa, Canada.

11. Chisholm, M., and O' Sullivan, P. (1973) . "Freight Flows and Spatial Aspects of the British Economy." Cambridge University Press, London, U.K.

12. Desauliners, Y. and Mozes, S.L. (1985). "The Canadian Trucking Industry." Canadian Transportation Research Forum, Proc.20, Toronto, Canada.

13. Dillon, William R. and Goldstein, M. (1984). "Multi-Variate Analysis; Methods and Applications." John Wiley and Sons, New York, U.S.A.

14. Draper, N.R., and Smith, H. (1981). "Applied Regression Analsysis." John Wiley & Sons, Inc. 2nd Ed. New York, U.S.A.

15. ECMT Economic Rsearch Centre. (1985). "Changes in Transport User's Motivations for Modal Choice." Report of the Sixty-Ninth Round Table on Transport Economics, OECD Publications Office, Paris, France.

16. Fox, John. (1984). "Linear Statistical Models and Related Methods." John Wiley and Sons, New York, U.S.A.

17. Fienberg, S. (1980). "The Analysis of Cross Classified Categorical Data." 2nd Ed., M.I.T Press Cambridge, Mass., MIT Press, U.S.A.

18. Hartwing, J.C., and Linton, W.E. (1974). "Disaggregate Mode Choice Models of Intercity Freight Movements". Masters Thesis, Northwestern University, U.S.A.

19. Herendeen, J.H., Jr.(1969)." Theoretical Development and Preliminary Testing of a Mathematical Model for Freight Modal Split." Pennsylvania State University and Pennsylvania Transportation and Traffic Safety Center, USA.

20. Hutchinson, B.G. (1972). "Forecasting Inter-Regional Commodity Flows". A paper written for submission to NATO Conference: The Application of Operational Research to Transport Problems, Sandefjord, Norway.

21. Hutchinson, B.G. (1974). "Principles of Urban Transport System Planning." Scripta Book Company, Washington D.C, U.S.A.

22. Isard, Walter. (1954). " Location Theory and Trade Theory: Short Run Analysis." The Quarterly Journal of Economics, Vol. 68, pp. 305-320,U.S.A.

23. Kennedy, Norman., et al.(1969)."Fundamentals of Traffic Engineering." Institute of Transportation and Traffic Engineering, University of California, Berkeley, USA.

24. Kullman, B.C. (1973). "A Model of Rail Truck Competition in the Intercity Freight Market", Vol. 15, prepared for Federal Rail Road Administration, U.S. Dot, U.S.A.

25 Leech. W. James.(1973). " A Gravity Model Approach to Canadian Freight Flows." A thesis submitted to the Graduate School of Business, Queen's University, Kingston, Canada.

26. Leontief.W., and A. Strout. (1963). "Multi-Regional Input-Output Analysis." T. Barn ed., Sturctural Interdependence and Economic Development, St. Martins Press, pp. 119-150, New York, U.S.A.

27. Meyer, J.R., et al. (1964)."The Economics of Competition in the Transportation Industries". Harvard University Press, U.S.A.

28. Milne, A.M., and Laight, J.C. (1963). "The Economics of Inland Transport." Pitman and Sons, 2nd Ed. London, U.K.

29. Milklius, W. (1969). "Estimating Freight Traffic of Competing Transportation Modes: An application of the Linear iscriminant Function." Journal of Land Economics, Vol. 45, pp. 267-273, U.S.A.

30. Miller, Leon H.(1970)."Transport Network Model." working paper number 61, Industrial Develoment association, International Bank for reconstruction and development, USA.

31. Morlok, K. Edward., and John, A. Warner. (1977). "Approximation Equations for Costs of Rail, Trailer-on-Flatcar, and Truck Intercity Freight Systems." Transportation Research Record 637, Washington, D.C, U.S.A.

32. Murthy, A.S.N., and Ashtakala, B. (1987). " Modal Split analysis using Logit Models." Journal of Transportation Engineering, ASCE, Vol. 113, No. 5, pp. 502-519, New York,U.S.A.

33. National Research Council. (1977a). "Freight Data Requirements for Statewide Transportation System Planning." Report No. 177, NCHRP, Washington D.C, U.S.A.

34. National Research Council. (1977b). "Freight Data Requirements for Statewide Transportation System Planning." Report No. 178, NCHRP, Washington D.C, U.S.A.

35. National Research Council. (1983). "Application of Statewide Freight Demand Forecasting Techniques." Report No. 260 NCHRP, TRB, Washington D.C, U.S.A.

36. Niedercorn, J.H. and Bechdolt, B.V.(1969) ."An Economic Derivation of the 'Gravity Model" of Spatial Interaction." Journal of regional Science, Vol. 9 (1969)

37. Randolph, W. Hall. (1985). "Dependence Between Shipment Size and Mode in Freight Transportation." Transportation Science, Vol. 19, No. 4, U.S.A.

38. Roth, Ronald D. (1977). "Approach to Measurement of Model Advantage." Transportation Research Record 637, Washington D.C, U.S.A.

39. Surthi, V.H., and Ebrahim, A.(1972). "Modal Split of Freight Traffic." Traffic Quarterly, Vol. 26, No. 4, pp. 575-587, U.S.A.

40. Sasaki, R. Brion.(1974)."Predicting Transporter's Choice Mode." Transportation Research Record 577, Washington D.C, U.S.A.

41. Stevens, B.H.(1961)."Linear Programming and Location Rent." Journal of Regional Science, Vol. 3, pp. 15-26.

42. Turner, R.E., et al. (1973). "Modal Choice for Canadian Freight: A CFTM Program Report." Presented at Annual Meeting of Canadian Association of Administrative Sciences, Queen's University, Kingston, Canada.

43. Trimac Consulting Services Ltd. (1977). "Pilot Study of Alberta Freight Transportation, for Alberta Transportation." Alberta Transportation, Edmonton, Alberta, Canada.

44. Trimac Consulting Services Ltd. (1978). "Final Study, Alberta Commodity Transport Study for Alberta Transportation." Alberta Transportation, Edmonton, Alberta, Canada.

45. Ullman, E.L. (1957). "American Commodity Flow." Seattle: University of Washington Press, U.S.A.

46. U.S. Department of Transportation. (1981). "An Investigation of Truck Size and Weight Limits." Final Report, Washington, D.C, U.S.A.

47. Wilson, A.G. (1970). "Advances and Problems in Distribution Modelling." Transportation Research, Vol. 4, pp. 1-18, Pergamon Press, U.K.

48. Wilson, A.G.(1969)."The use of Enthropy Maximizing Models in the Theory of Trip Distribution, Mode Split and Route Split." Journal of Transport Ecnonomics and Policy, pp. 108-126.

FIGURES

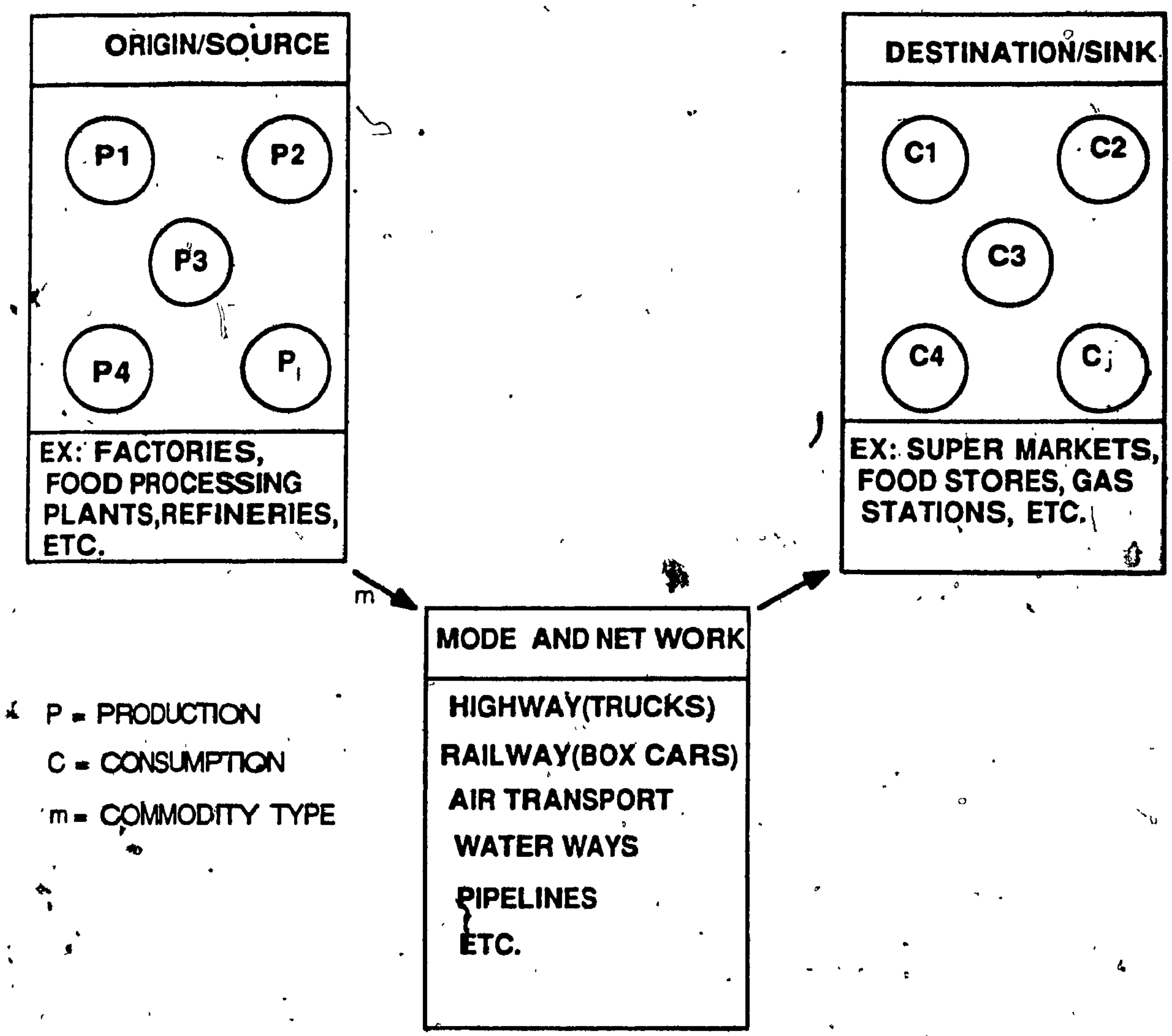


FIG. 1a COMMODITY TRANSPORTATION SYSTEM

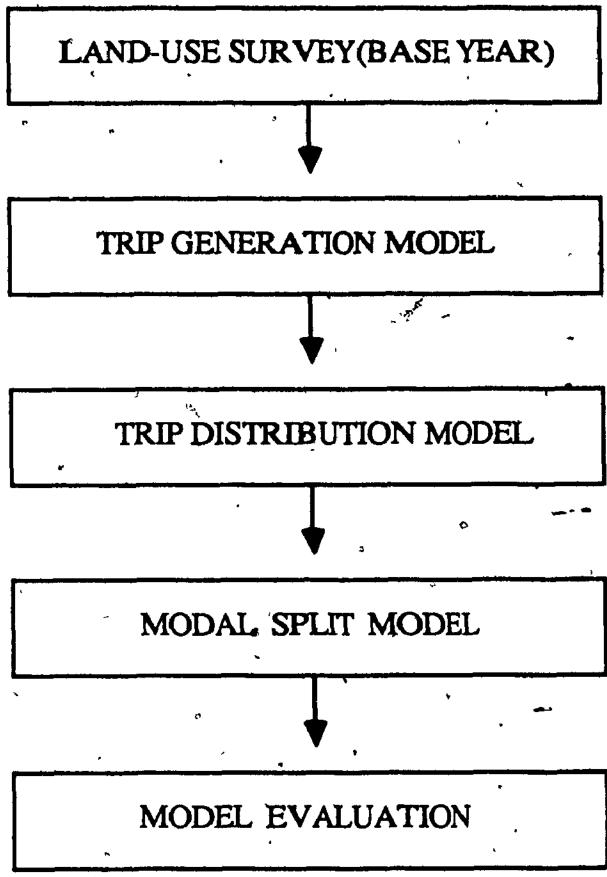
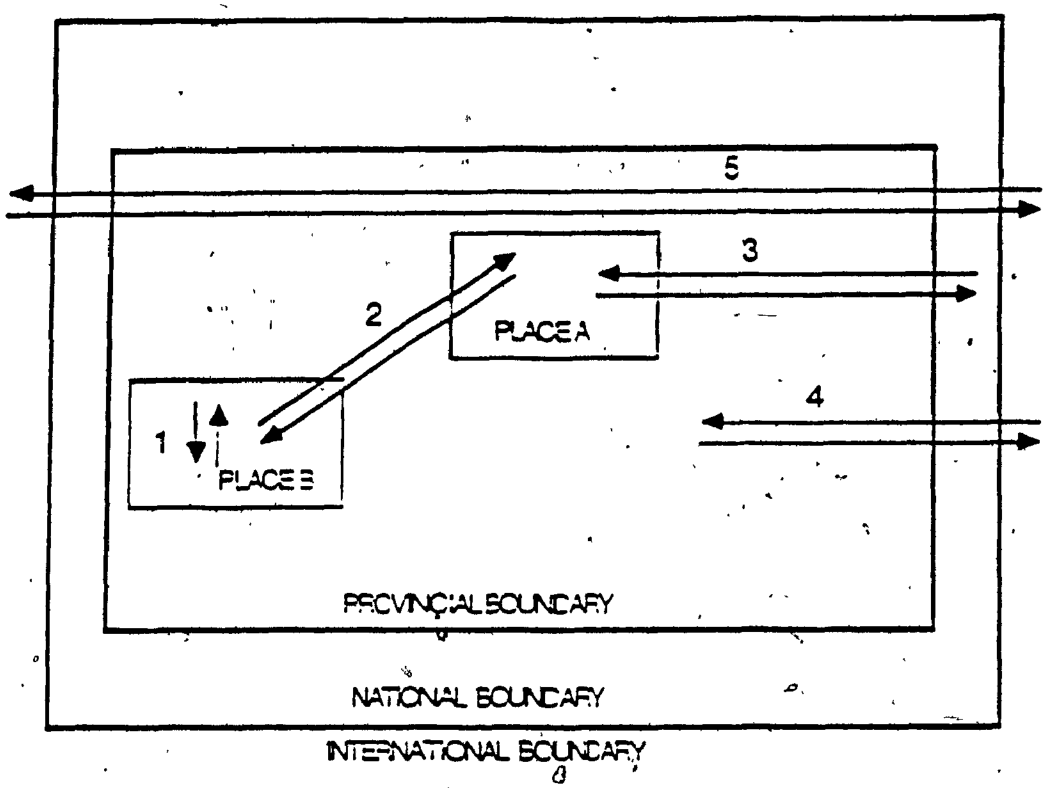


FIG. 1b STAGES IN DEVELPMENT OF TRANSPORT MODELS



CODE	PARTICULARS	CLASSIFICATION
1	Within City	Local
2	Between Cities	Intra-Provincial
3	External to Province	National
4	External to Province	International
5	Through Freight	National/International

FIG. 2 COMMODITY MOVEMENTS AND CLASSIFICATIONS

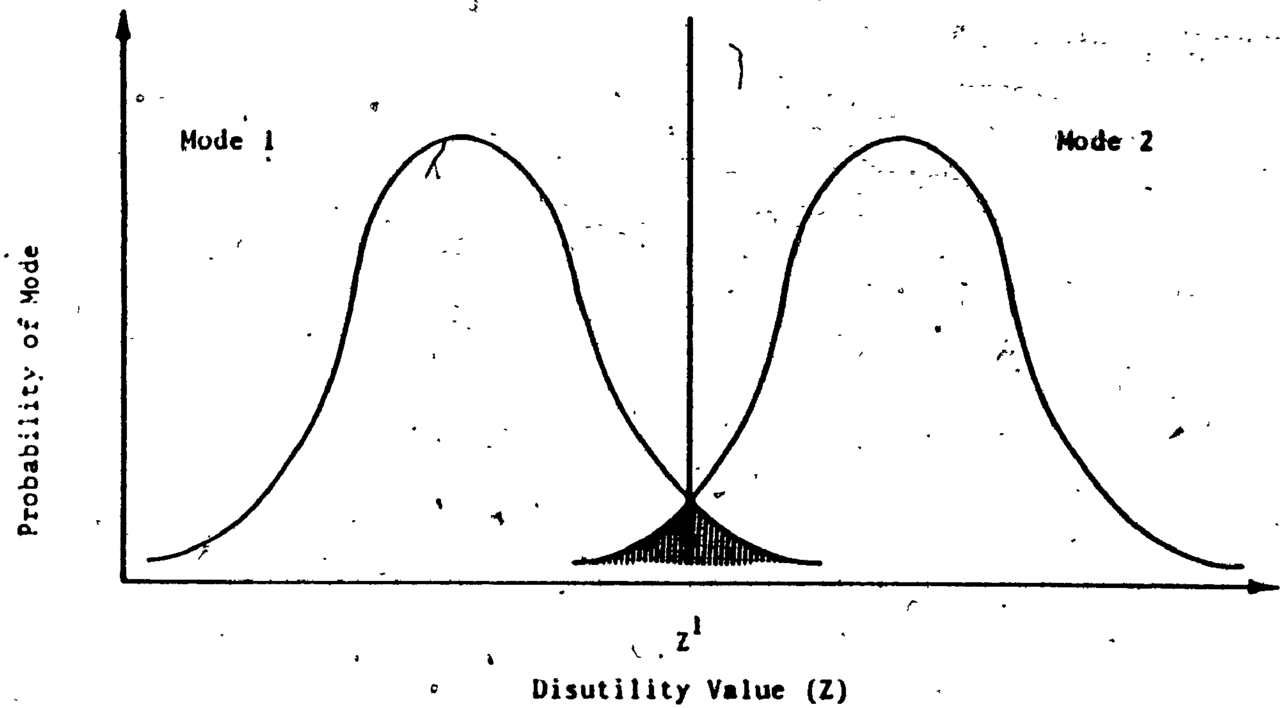


FIG. 3a DISCRIMNINANT FUNCITON METHOD

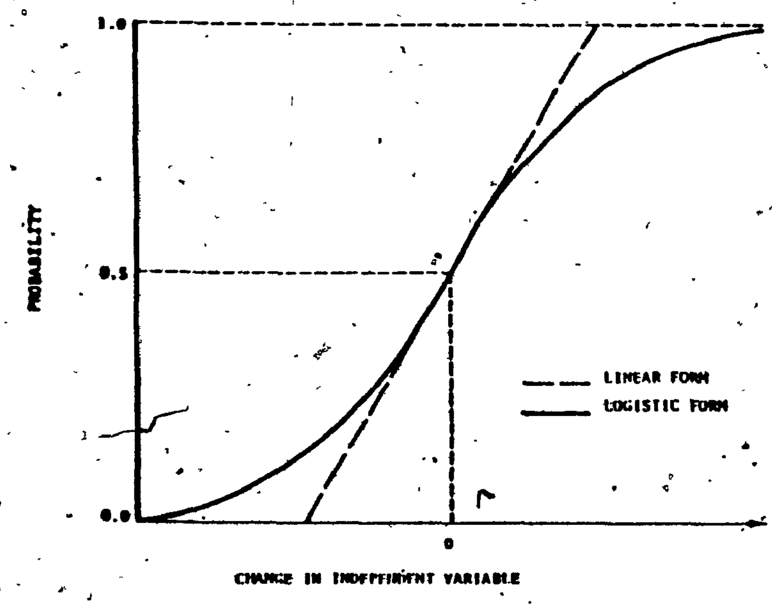


FIG. 3b LOGIT MODEL

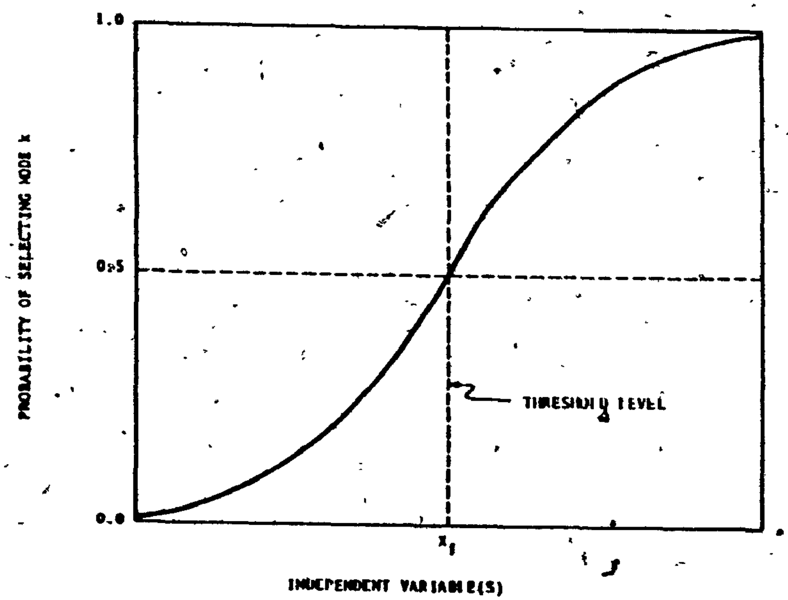


FIG. 3c PROBIT MODEL

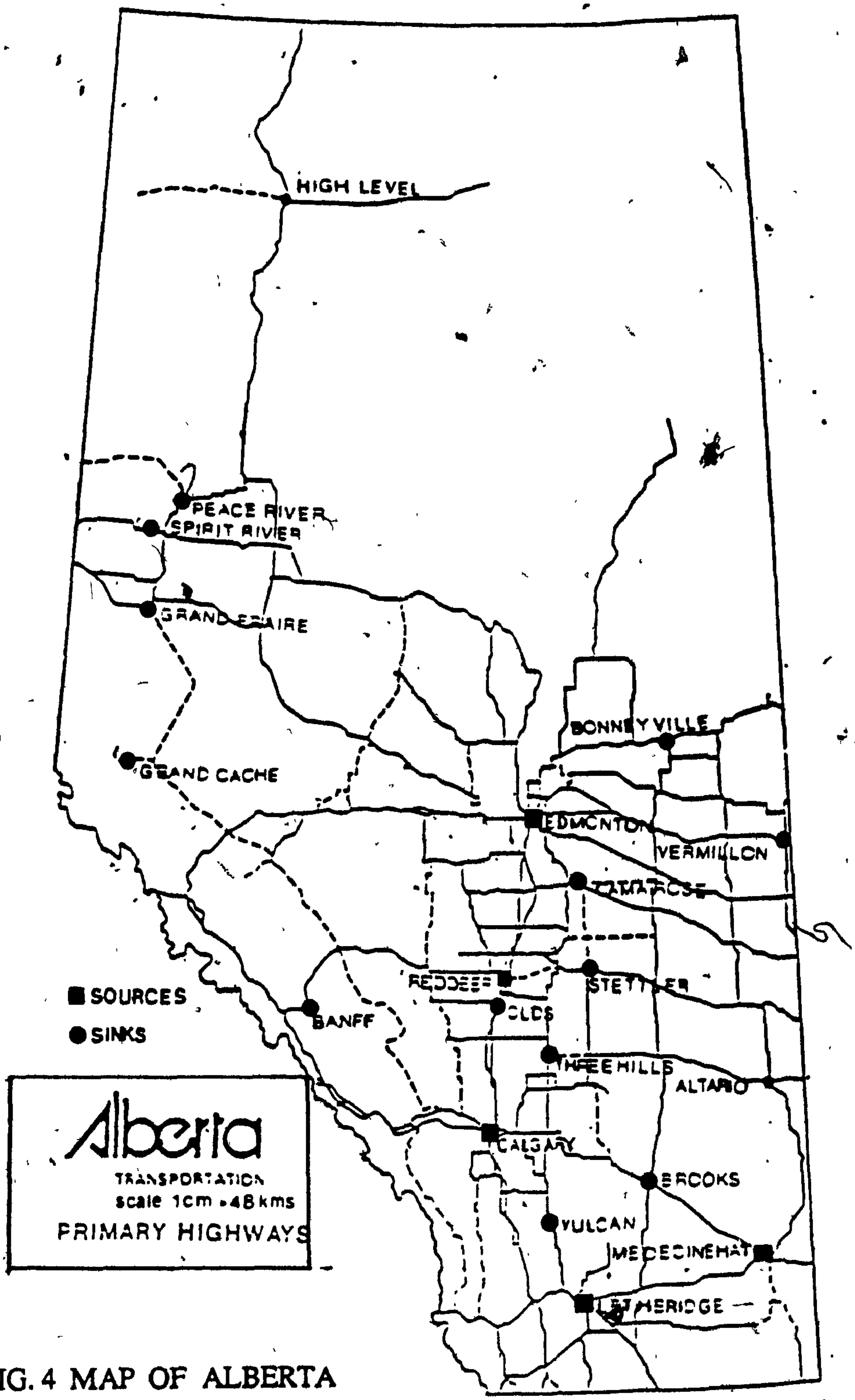


FIG. 4 MAP OF ALBERTA

ALBERTA TRANSPORTATION COMMODITY FLOW SURVEY — NAME: WAYNE BELWAY DATE:OCTOBER 31/77

	PLACE	PLACE CODE	FIRM NAME	PHONE
1	FORT CHIPWYAN	5427	BELWAY'S GENERAL STORE	TEL: 499-3003
1	2–21	22–25	26–45	46–52

	ADDRESS OF FIRM	POSTAL CODE		
21	12 MAIN STREET	K1A 0S4	05	

For Continuation Code 'C' and —
Telephone Numver —

FIRM TYPE (BY CODE)

01 FOOD STORE, 02 HARD WARE, 03 CLOTHING, 04 DEPARTMENT STORE, 05 GENERAL STORE,
06 ELEVATOR(NON-GRAIN BUSINESS), 07 BULK FUEL/ CHEMICAL SUPPLIER, 08 FARM IMPL. DEALER,
09 AUTOMATIVE/TRUCK DEALER, 10 CONSTRUCTION MACHINERY DEALER, 11 LUMBER YARD/ BUILDING SUPPLIES,
12 DAIRY, 13 FEED MILL, 14 AUCTION MART, 15 PACKING PLANT, 16 FEED LOT, 22 GAS PLANT, 23 OIL SANDS,
24 COAL, 25 IN-SITU, 26 PETROCHEMICAL/FERTILIZER, 27 PULP MILL, 99 OTHER.

	GOODS TYPE	RECEIVED(1), SHIPPED(2).	FROM/TO	COMMODITY	ANNUAL TONNAGE	(TONS) AV. SHIP. SIZE	MODE	'F' OR 'L' FL/LTFL	'Y' OR 'N' CONTROL	PRIVATE OR FOR HIRE 'P' OR 'H'	CALCULATIONS / REMARKS
			EDMONTON	CANNED GOODS	24						
3	10	1	4 2 1 3	02 01	24	2	4	L	N	H	CANNED FOOD, 2 TONS/MONTH
											BY AIR FREIGHT, PWA.
3	10	1									
3	10	1									
3	10	1									
3	10	1									

GOODS TYPE(BY CODE)
10 CONSUMER LOCAL, 11 CONSUMER RESUPPLY, 21 FARM, 22 GAS PLANT, 23 OIL SANDS,
24 COAL, 25 IN- SITU, 26 PETROCHEMICAL/FETILIZER, 27 PULP MILL, 28 FOOD PROCESSING,
29 OTHER INDUSTRY, 99 OVERLAP BROKEN CONSUMER/INDUSTRY.
MODES: 1 TRUCK, 2 RAIL, 3 PIPELINE, 4 AIR, 5 MARINE, 6 BUS, 7 WINTER ROAD, 8 OTHER.

PRINCIPAL COMPETITORS
MCBRIK'S GROCERY, ASH'S VEGETABLE STAND
FINN'S FISH MARKET, FREDMEN'S TEXACO SERVICE.

FIG. 5 SAMPLE SURVEY FORM

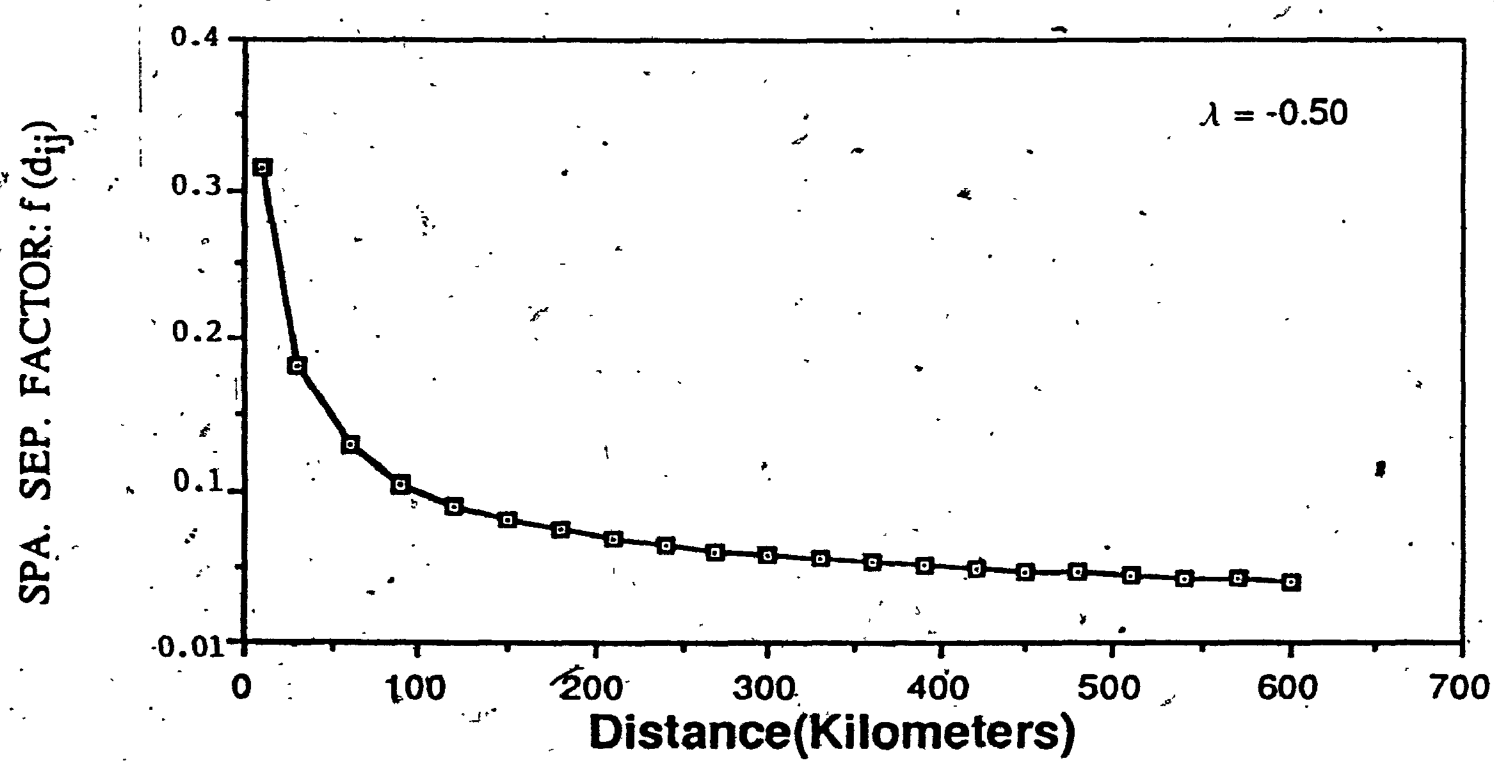


FIG. 6 SPATIAL SEPERATION FACTOR

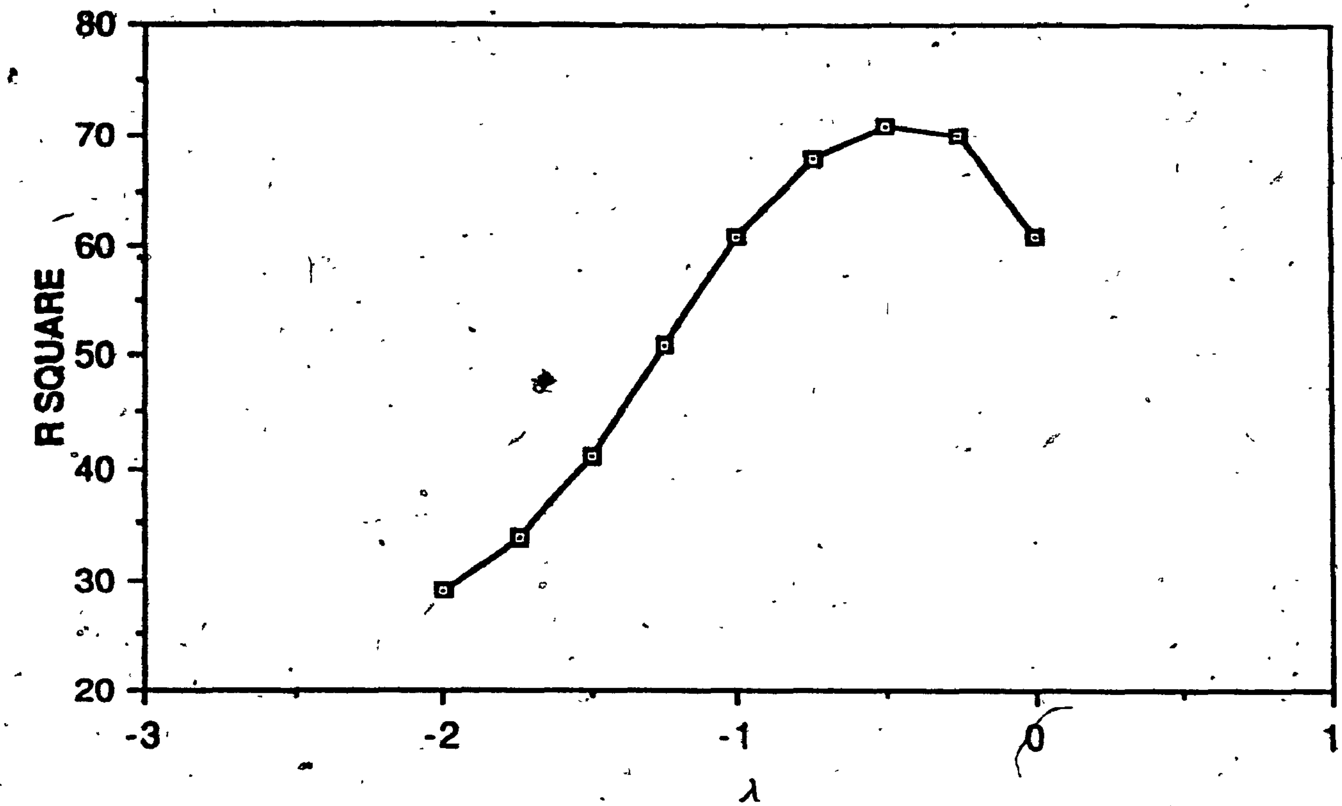


FIG. 7 R^2 Vs λ(FOOD STUFFS)

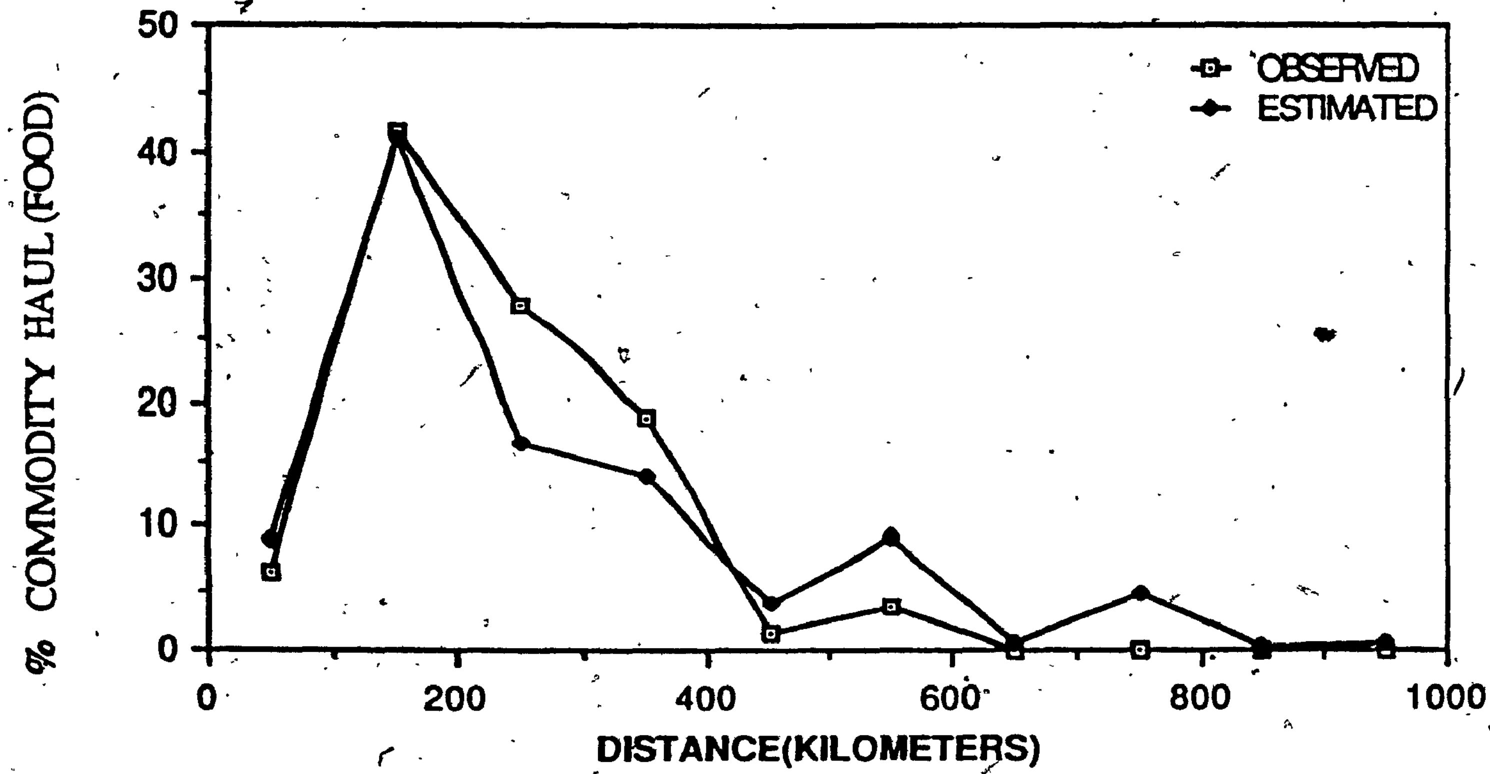


FIG. 8 COMMODITY HAUL FREQUENCY DIAGRAM(FOOD STUFFS)

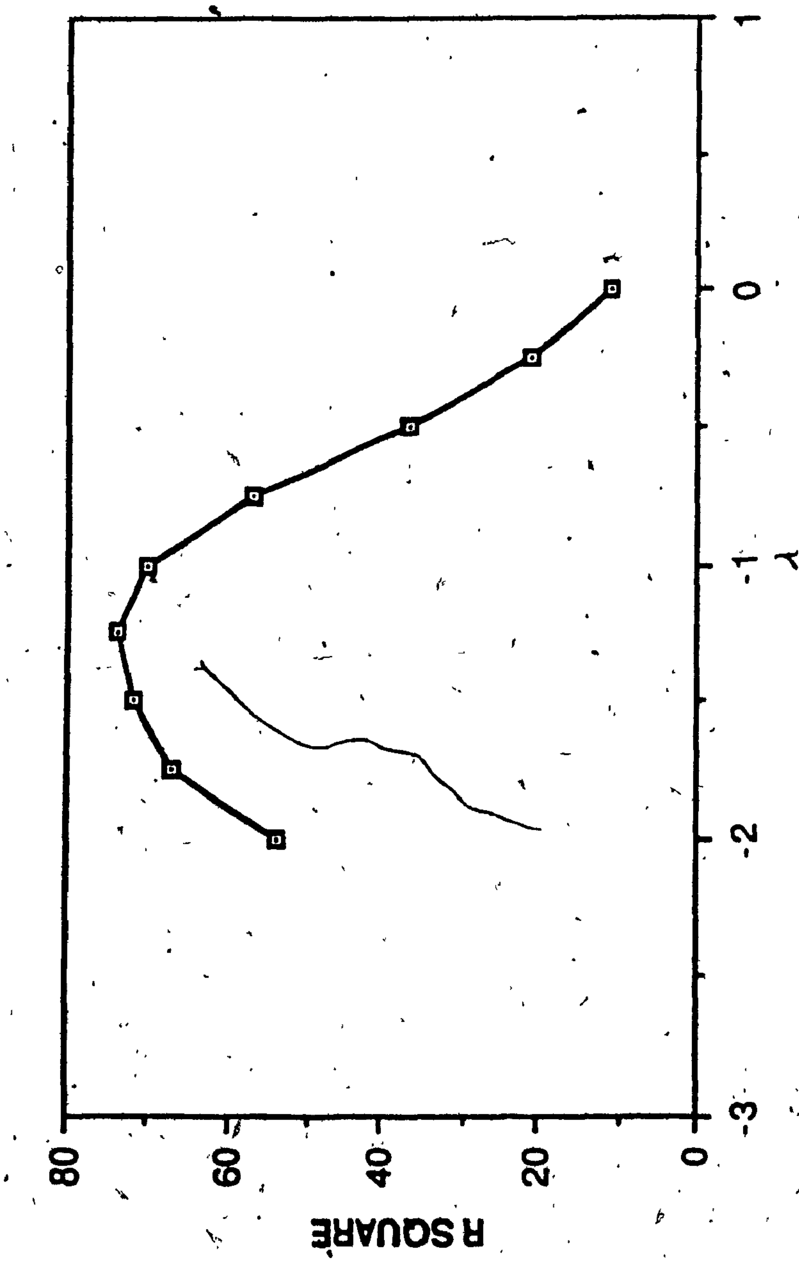


FIG. 9 R^2 Vs λ(FREIGHT)

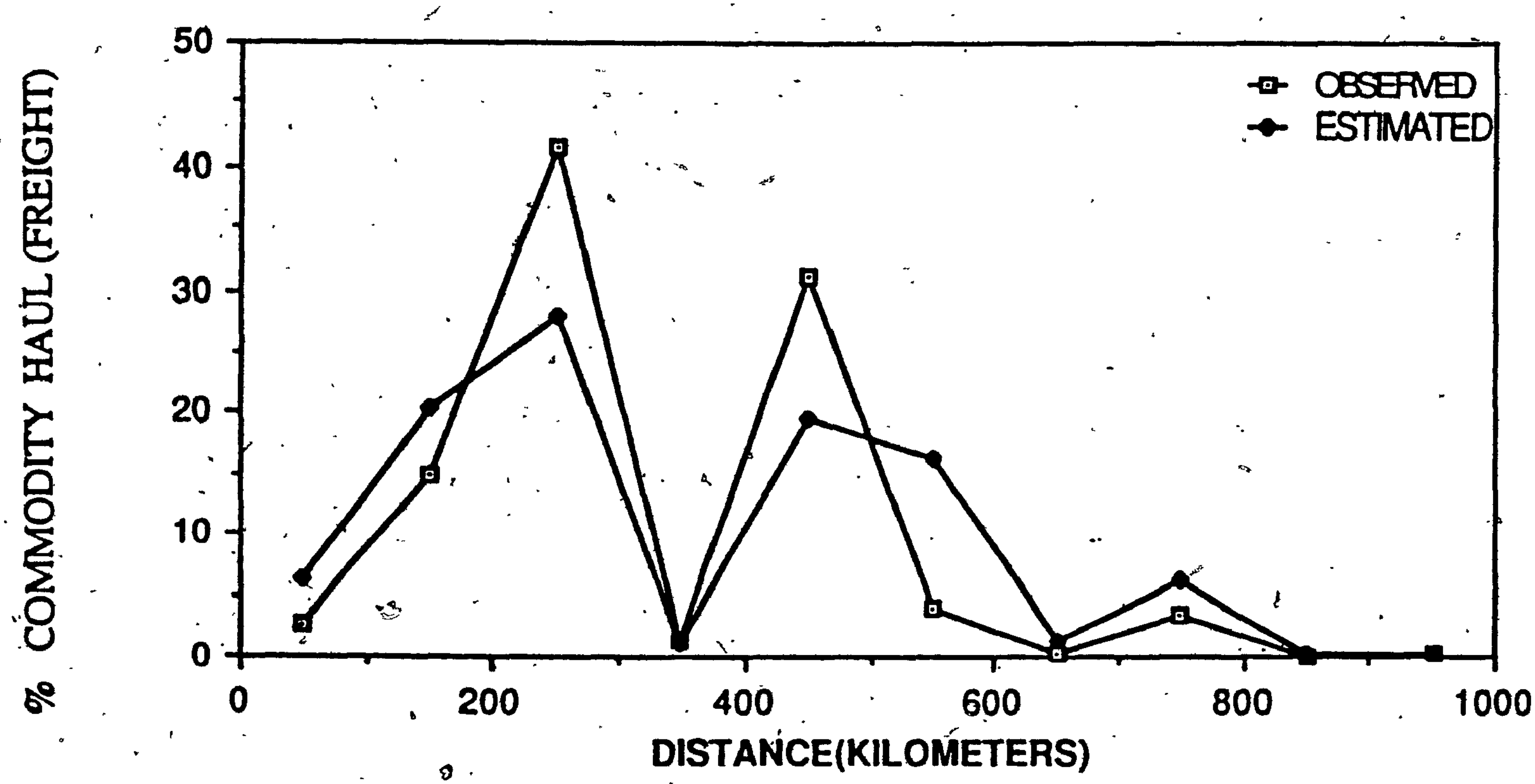


FIG. 10 COMMODITY HAUL FREQUENCY DIAGRAM(FREIGHT)

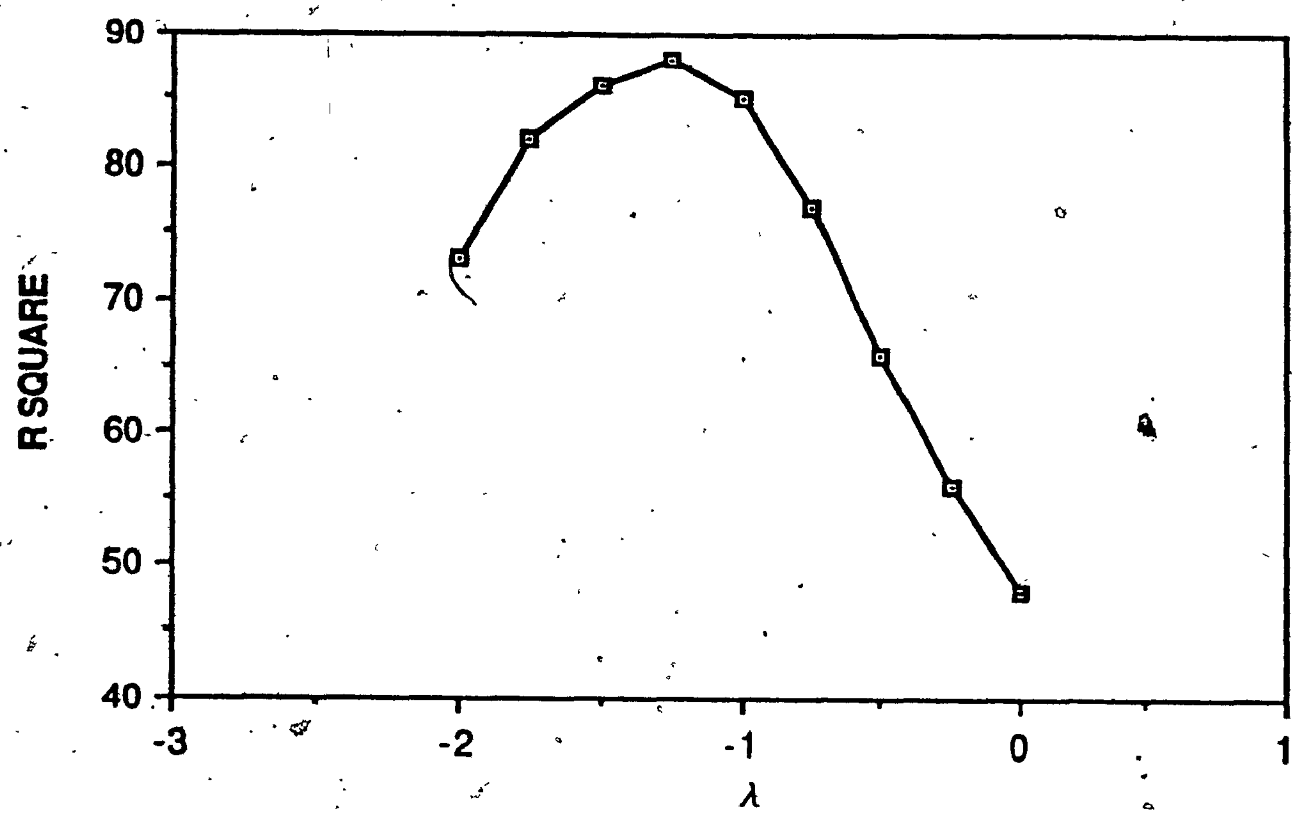


FIG. 11 R^2 Vs $\hat{\lambda}$(FOREST)

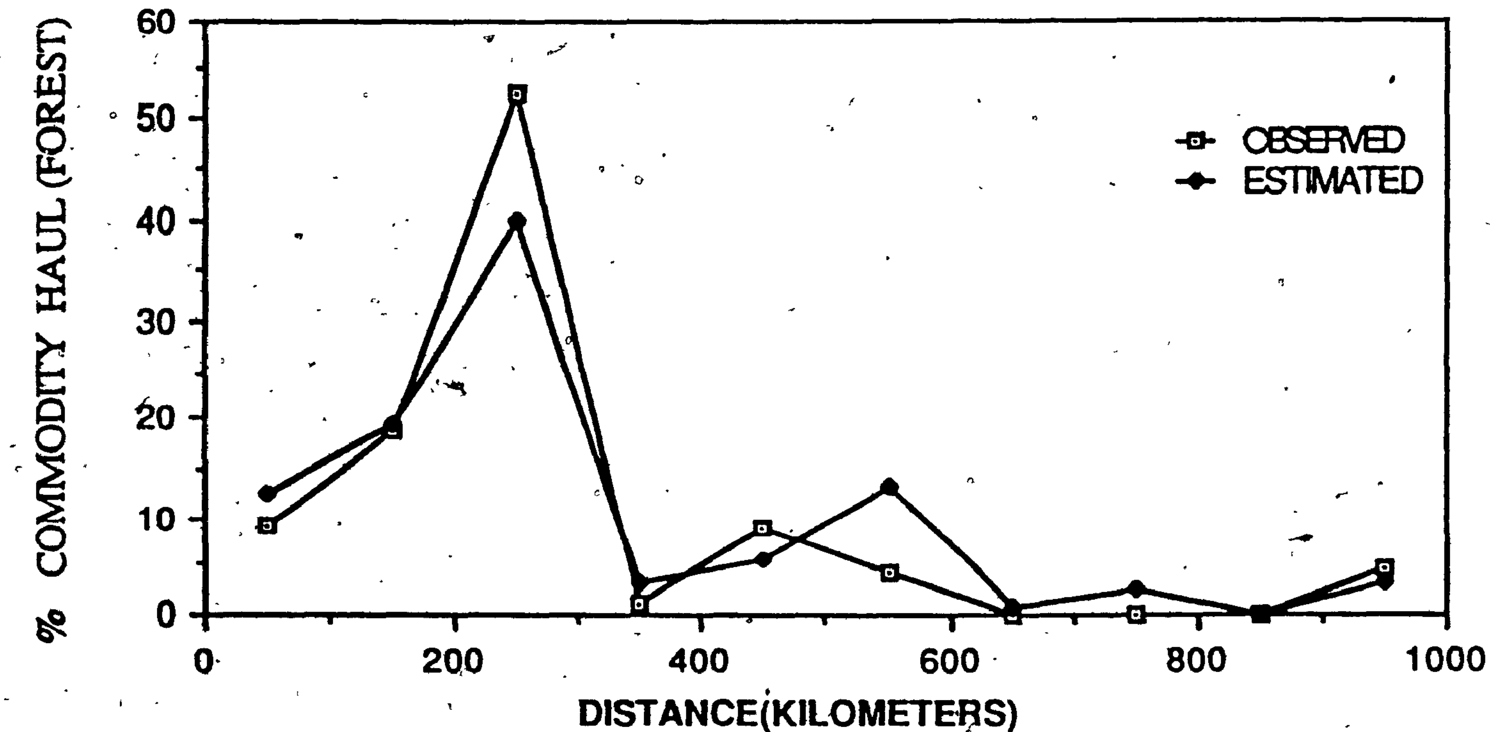


FIG. 12 COMMODITY HAUL FREQUENCY DIAGRAM(FOREST)

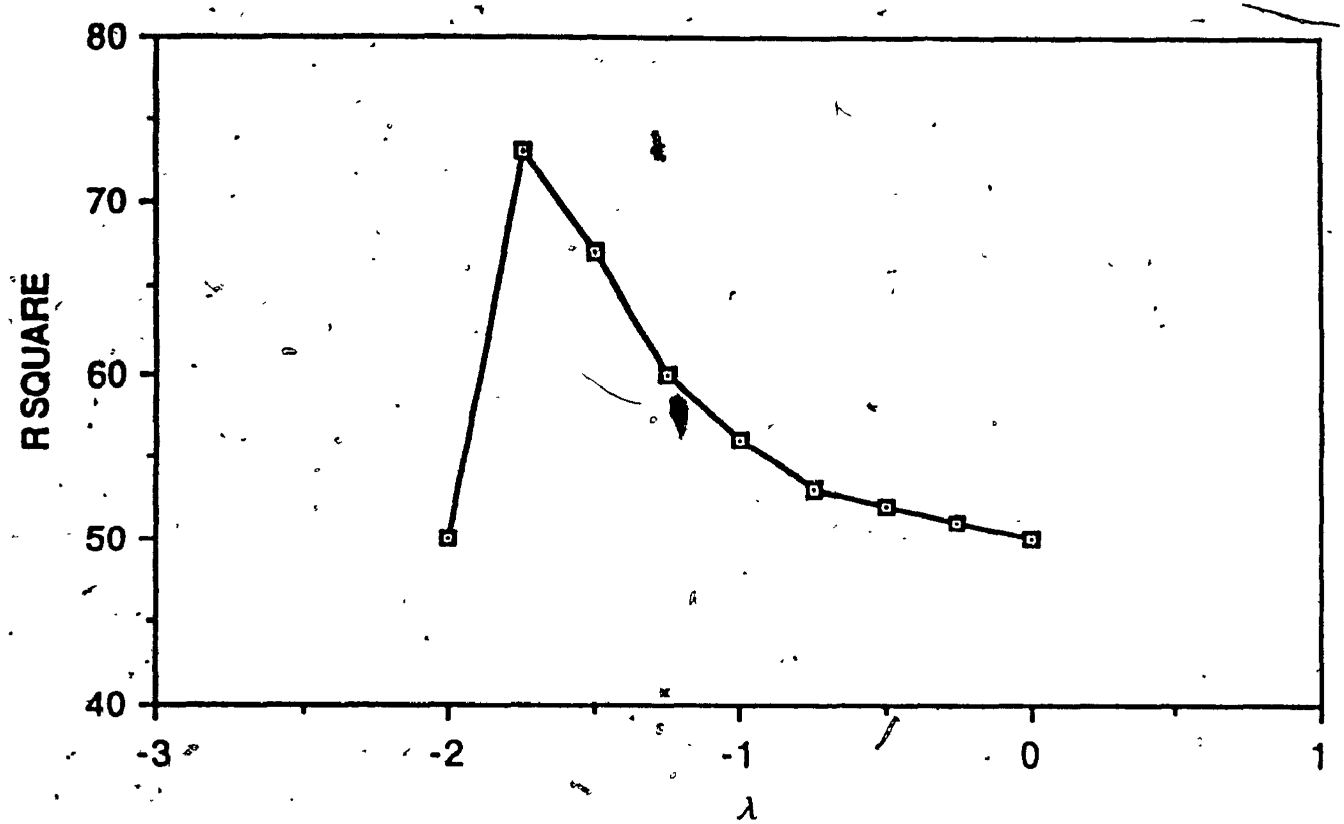


FIG. 13 R^2 Vs λ(PETROLEUM)

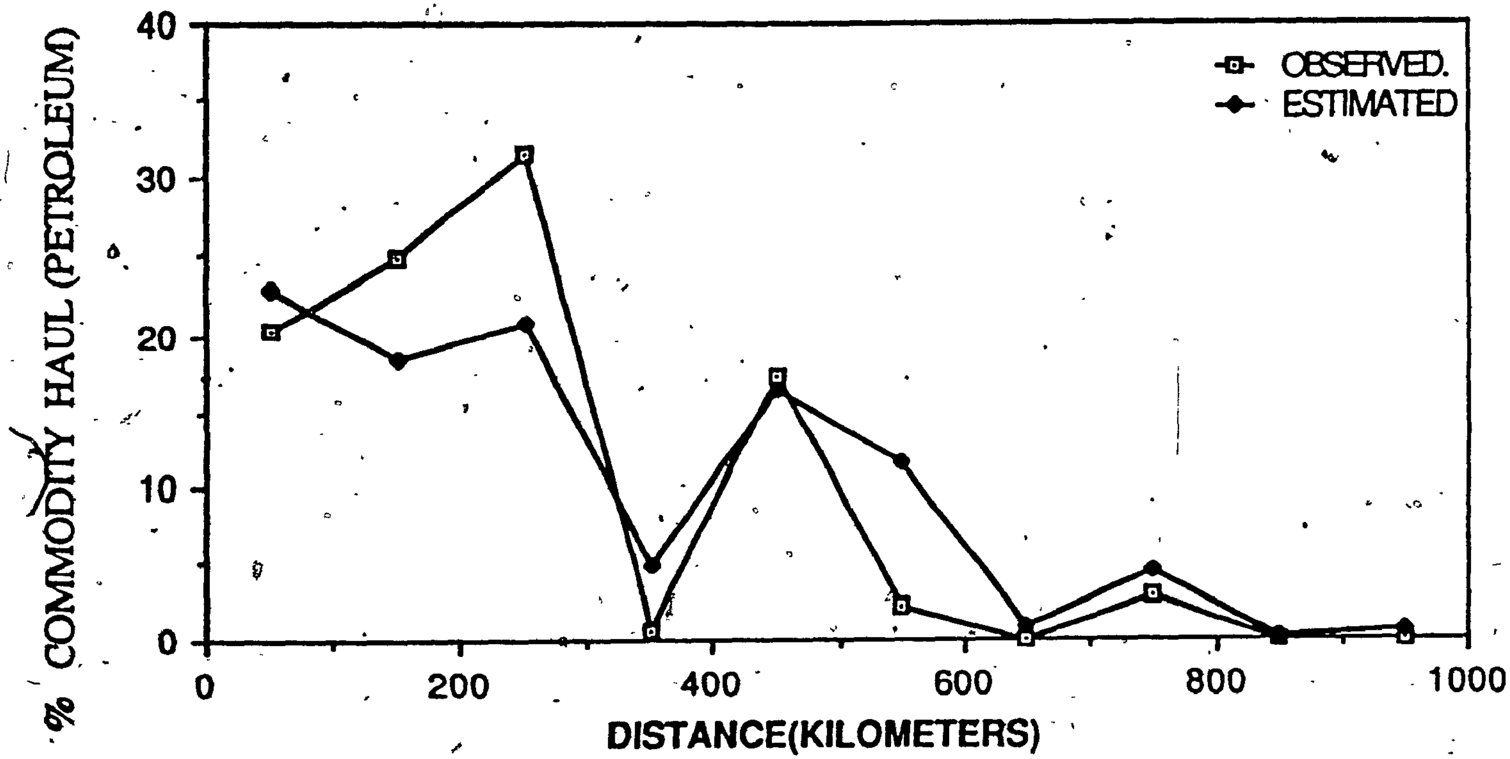


FIG. 14 COMMODITY HAUL FREQUENCY DIAGRAM(PETROLEUM)

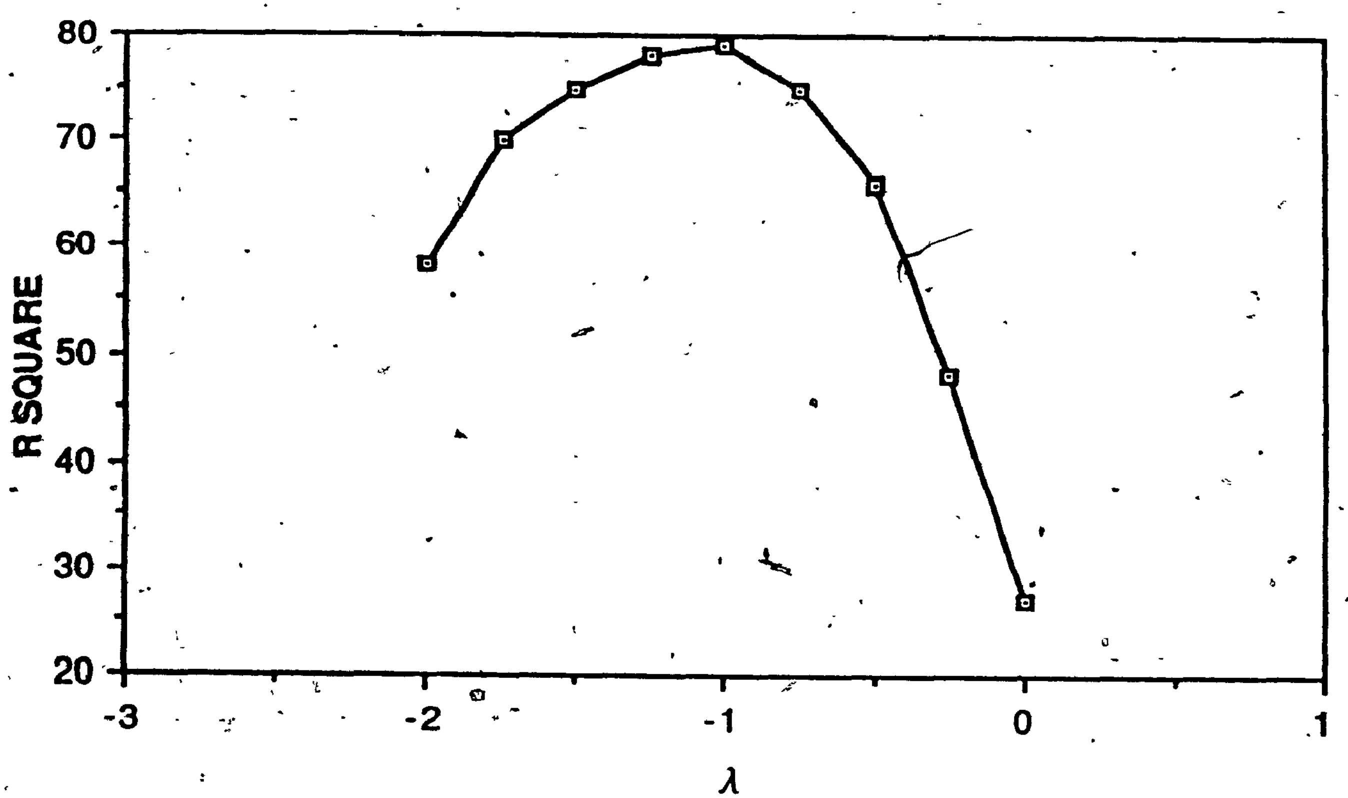


FIG. 15 R^2 Vs λ(BULK LIQ. AND DRY)

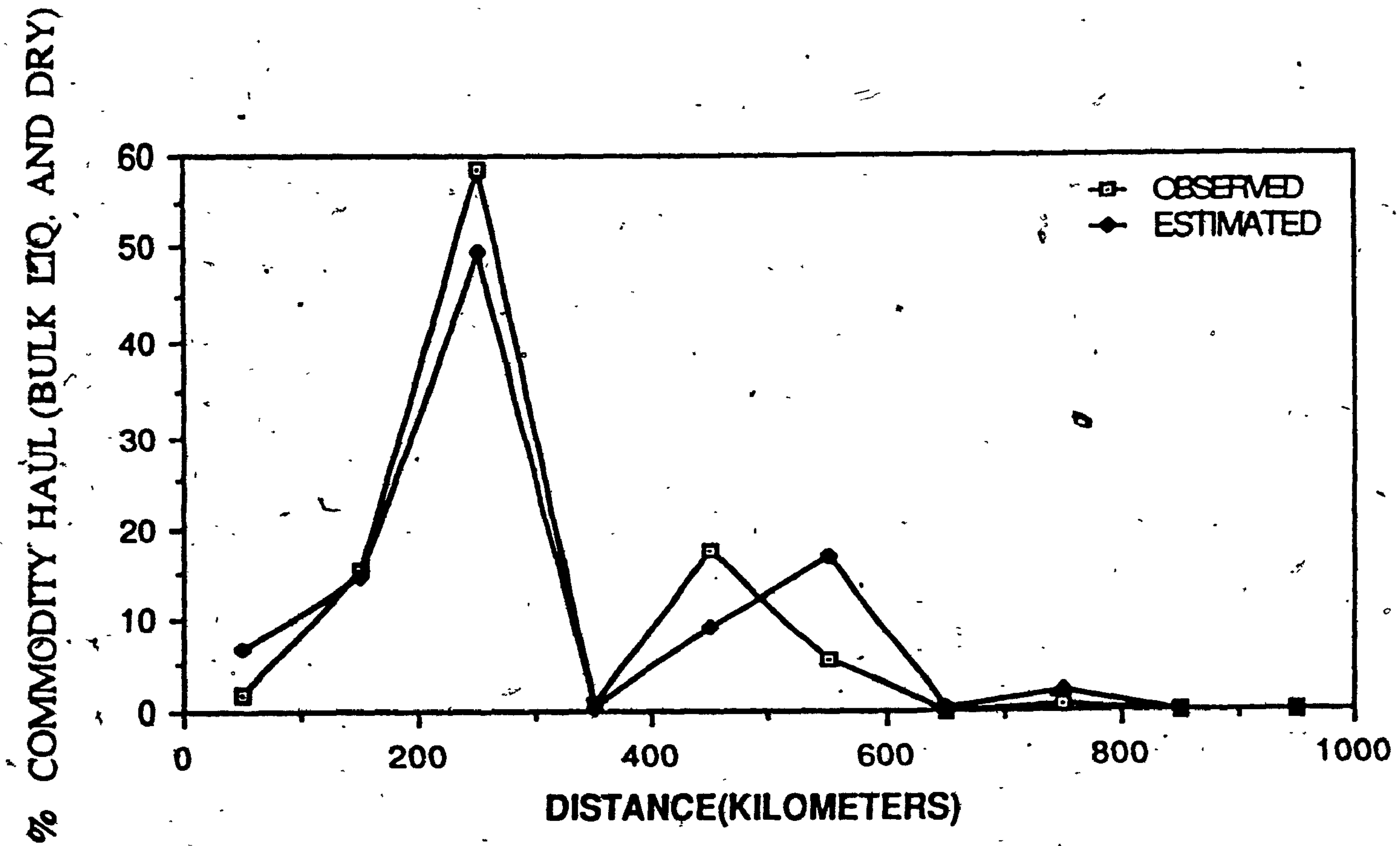


FIG. 16 COMMODITY HAUL FREQUENCY DIAGRAM(BULK LIQ. AND DRY)

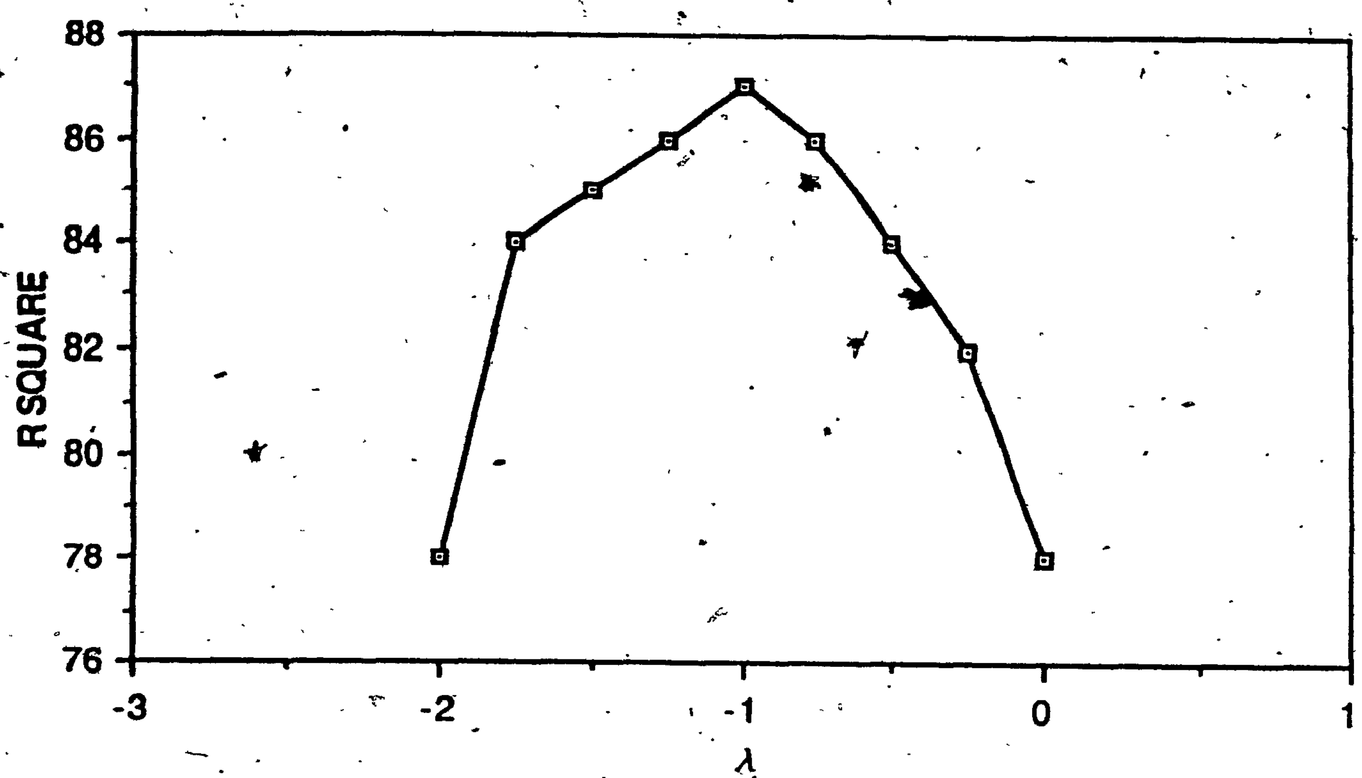


FIG. 17 R^2 Vs λ(HEAVY MACH.)

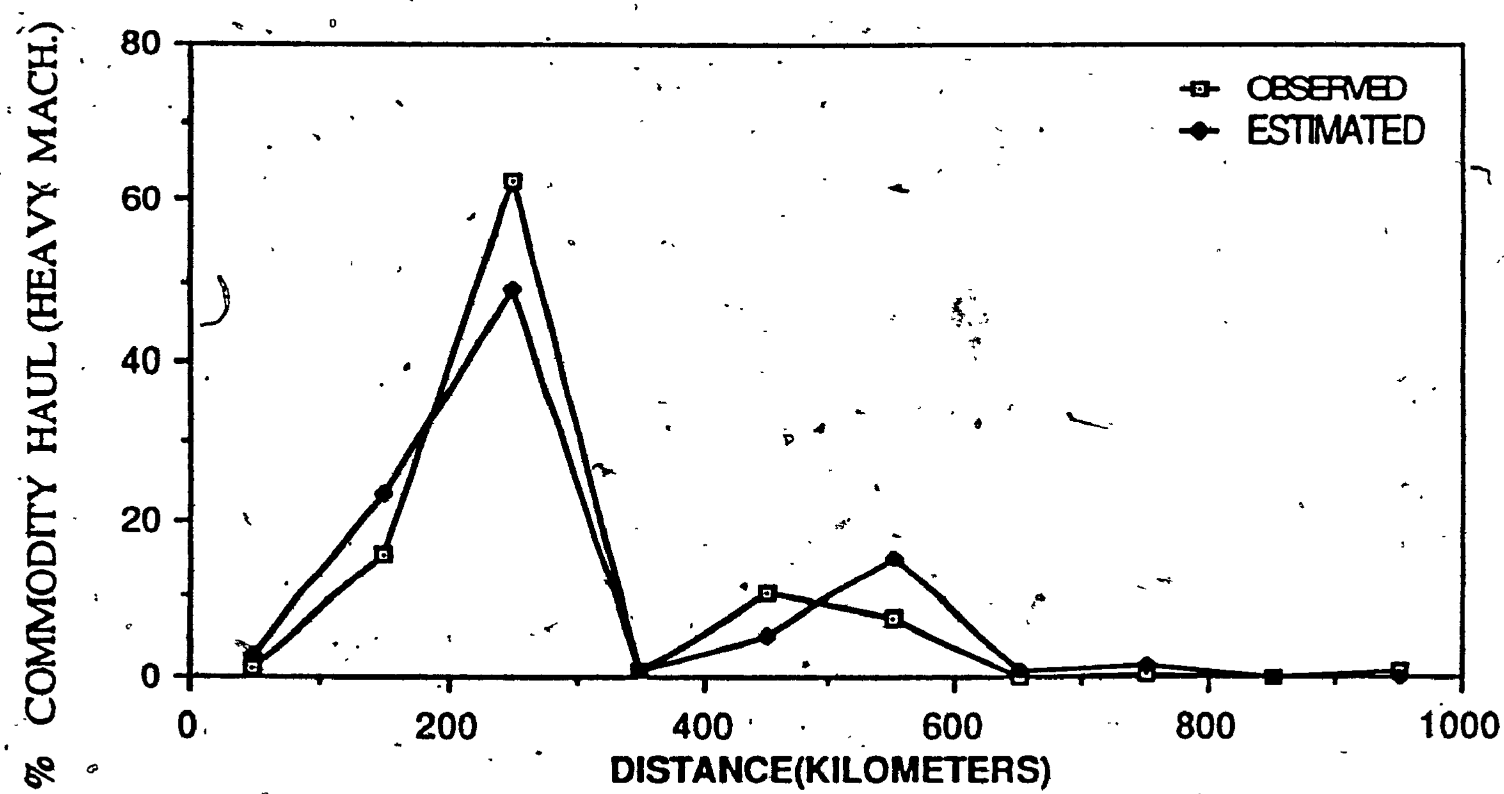


FIG. 18 COMMODITY HAUL FREQUENCY DIAGRAM(HEAVY MACH.)

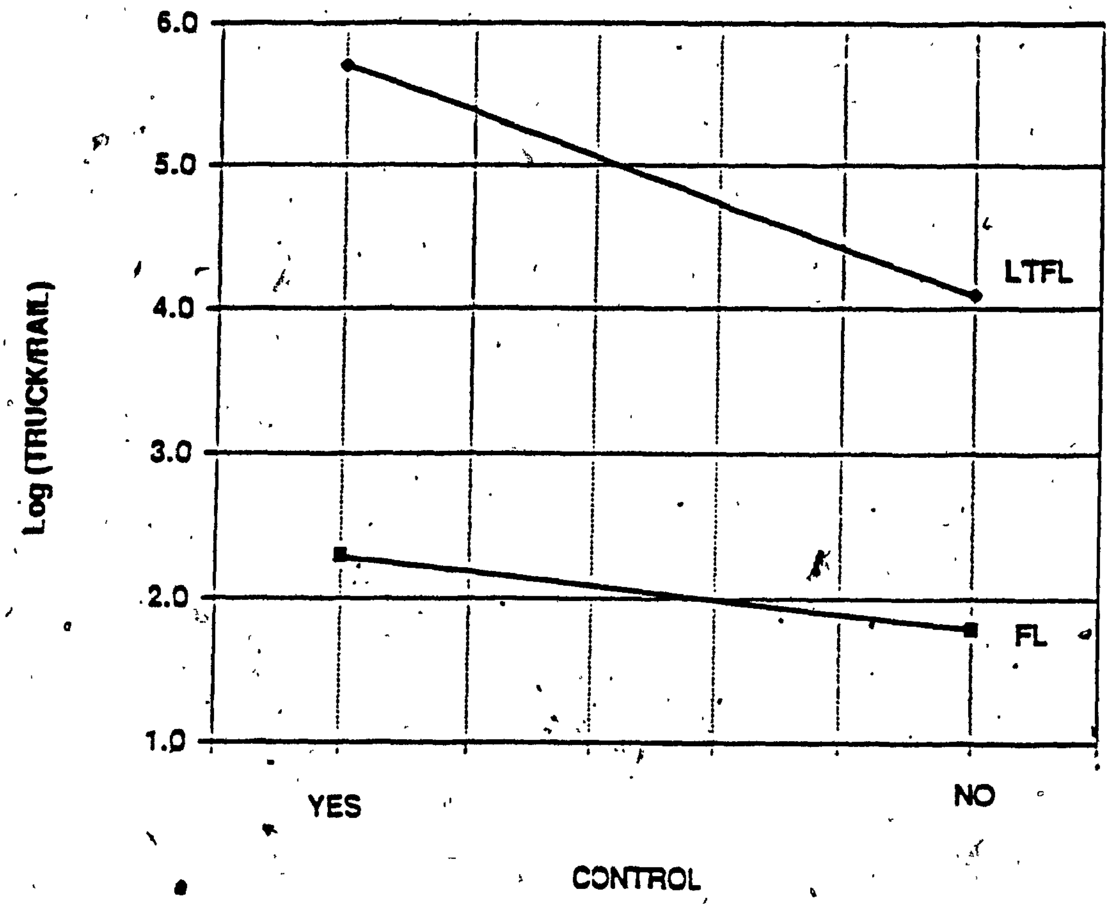


FIG. 19 PLOTTED LOGITS FOR [$\overline{CL}$][CM][HM] MODEL

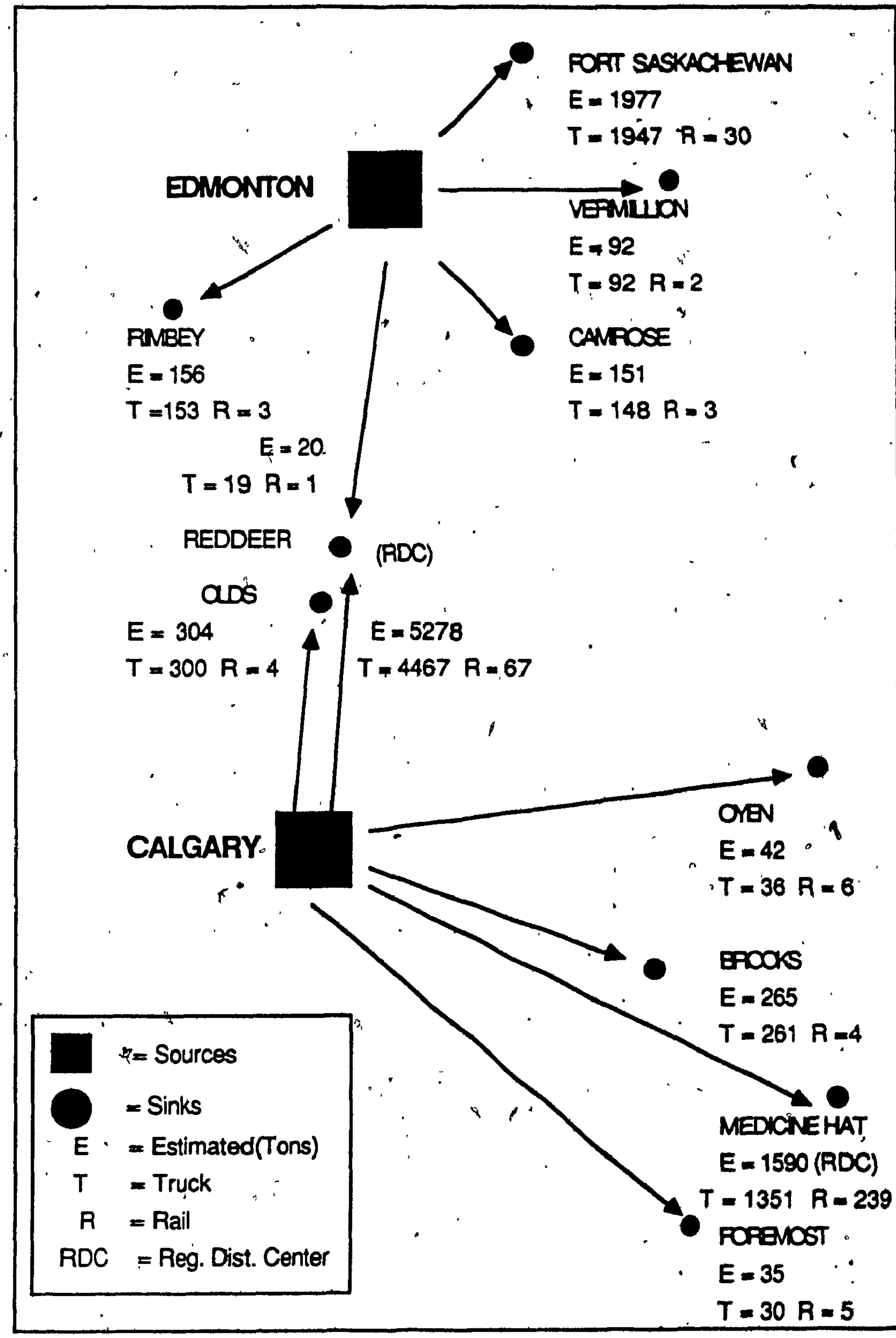


FIG. 20. EXAMPLE OF DISTRIBUTION AND MODAL SPLIT MODELS(FOOD STUFFS)

TABLES

TABLE 1. - STUDY COMMUNITY CHARACTERISTICS

Community and Community code	Community Type	Population Range	Alternate Shopping Centre	Principle Economic Base
1.Altario	Village	under 250	Redeer	Agriculture
2.Athabasca	Town(S)	1000-2500	Edmonton	Agriculture
3.Banff *	Village	under 250	Calgary	Tourist
4.Bonneyville	Town(S)	2500 t0 5000	Edomonton	Agriculture
5.Brooks	Town(M)	5000-10000	Calgary	Oil/Agriculture
6.Brudett	Village	250 to 1000	M. Hat/Calgary	Agriculture
7.Camrose	City	10000 to 25000	Edmonton	Mixed(RDC)
8.Caroline *	Town(L)	10000 to 25000	REC	Mixed(Forestry)
9.Coleman	Town(S)	1000 to 2500	LCE	Coal Mines/Mills
10.Coronation	Town(S)	1000 to 2500	REC	Agriculture
11.Coutts	Town(L)	10000 to 25000	LethBridge/Calgary	Boarder Crossing
12.Edson	Town(M)	5000 to 10000	Edmonton	Forestry/Coal Field
13.Empress	Town(S)	1000 to 2500	M.Hat/Calgary	Agr./Gas Plants
14.Foremost	Town(L)	10000 to 25000	LethBridge/Calgary	Agr.Service
15.Fort Chipway	Town(L)	10000 to 25000	Edmonton	Isolated Community
16.F. McMurray*	City	25000 to 50000	Edmonton	Oil Sands
17.Fort Sask.	Town(L)	10000 to 25000	Edmonton	Chemical/Ind. Town
18.Fox Creek *	Town(S)	1000 to 2500	Edmonton	Resource/Gas Plants
19.Grand Cache	Town(S)	2500 t0 5000	Edmonton	Coal Mines/Lumber
20.Grand Prairie	City	10000 to 25000	Edmonton/Calgary	Mixed(RDC)
21.Hardisty	Town(L)	10000 to 25000	Edmonton/Camrose	Agr./Oil Fields
22.High Level *	Town(S)	2500 to 5000	Edmonton	Resources
23.Hinton *	Town(M)	5000 to 10000	Edmonton	Forestry
24.Lacombe	Town(M)	5000 to 10000	REC	Agriculture
25.Lethbridge	City	50000 to 100000	Calgary	Mixed(RDC)
26.Medicine Hat	City	25000 to 50000	Calgary/Lethbridge	Mixed(RDC)
27.LoydMinister*	City	10000 to 25000	Edmonton	Mixed(RDC)
28.Olds	Town(S)	2500 to 5000	Calgary	Agriculture Service
29.Oyen	Town(S)	1000 to 2500	Calg./Drumheller	Agriculture Service
30.Peace River	Town(M)	5000 to 10000	G.Prairie/Edmonton	Mixed(RDC)
31.Raymond	Town(L)	10000 to 25000	Lehtbridge/Calgary	Agriculture Service
32.Redcliff	Town(S)	2500 to 5000	M.Hat/Calgary	Industrial Suburb
33.Reddeer	City	50000 to 100000	Edmonton/Calgary	Mixed(RDC)
34.Rimbey *	Town(S)	1000 to 2500	Redeer/Edmonton	Agr./Gas Plants
35.Spirit River	Town(S)	1000 to 2500	G.Prairie/Edmonton	Agr.Service
36.Swan Hills *	Town(S)	1000 to 2500	Edmonton	Agr.Service
37.Three Hills	Town(s)	1000 to 2500	Calgary	Agr.Service
38.Vermilion	Town(L)	10000 to 25000	Edmonton	Agr.Service
39.Vulcan	Town(S)	1000 to 2500	Calgary	Agr.Service
40.White Court *	Town(M)	5000 to 10000	Edmonton	Resource
41.Youngs Town	Town(M)	5000 to 10000	Calg./DrumHeller	Agr./Oil Fields

V - Village, S - Small, M - Medium, L - Large. RDC - Regional Distribution Centre.
REC - Reddeer/Edmonton/Calgary. LEC - Lethbridge/Edmonton//Calgary.
* Indicates transient population at these places.
Agr. - Agriculture, Calg. - Calgary, G.Prairie - Grande Prairie, M.Hat - Medicine Hat, F - Fort

TABLE 2. - MAJOR AND MINOR COMMODITY CLASSIFICATIONS

Major Commodity Classifications	Minor Commodity Classifications
1.GENERAL FREIGHT	- mixed freight (all categories) - waste materials (eg. returned bottles, drums etc.) - clothing - hardware (including canned paint etc.) - appliances/furnitures - sporting goods and hobby supplies - automotive parts and accessories - other industrial parts and accessories - hides - tires etc.
2.FOOD STUFFS (NON-PERISHABLES)	- food preparations (can,box,bag,bottles etc.) - non-alcoholic beverages (bottles or cans) - alcoholic beverages - tobacoo -other -- specify.
3 FOODSTUFFS (PERSISHABLES)	- meat and meat preparations - fish and seafood- poultry - dairy products (not bulk)- bakery products - fresh fruits and vegetables - frozen fruits and vegetables- other -- specify.
4.HEAVY MACHINERY	- construction and maintenance (equipment and machinery) - oil industry equipment and machinery - sutomobiles and trucks(not parts) -other industrial machinery (not parts) - agricultural machinery and equipment - other -- specify - airplanes - boats.
5.METAL PRODUCTS	- scrap metal- slag, drosses, etc. - iron and steel products - precious metals (gold,mercury etc.) - other -- specify(eg. aluminium sheets etc.).
6.PETROLEUM PRODUCTS	- petroleum (crude)- natural gas and propane - gasoline - disel and fuel oil - liquid gasses (incl. oxyacetylene) - asphalt and road oils - plastics and rubber products - lubricants - other -- specify.
7.BULK LIQUIDS AND CHEMICALS (Non-Petroleum)	- acids - alcohol (bulk industrial)- hydrated limes - anhydrous ammoni - chlorine - paint - solvents - water. - milk - liquified gases- oils and fats (non-petroleum) - other -- specify
8.BULK DRY CHEMICALS AND MINERALS	- fertilizer - sulphur - lime - potash- cement (dry) - gypsm - ores and ore concentrates - coal - sand and gravel - alum and salt - other -- specify.
9.FOREST PRODUCTS	- logs - woodchips or sawdust - firewood - planks and boards- posts and poles - plywood products - cardboard - newsprint rolls - paper products - waste paper - other -- specify.
10.LIVE STOCK	- animals (cattle,swine,horses, etc.) - poultry (chickens,turkeys, etc.) - other -- specify (e.g -reptiles) - show animals(race horses,exhibition animals, etc.
11.CONSTRUCTION MATERIALS	- bricks or blocks - landscaping materials (sod,trees,etc.) - tar products (roofing) - flooring - gypsum products(drywall) - insulation - precast concrete - prefab housing components - other --- specify.
12.BULK SEED, FEED, AND FEED PRODUCTS	- feed grains - seed grains - milling grains- industrial grains - prepared feed - meal - other -- specify - sugar beets.
13.TRAILER-MOBILE HOMES	- recreational trailer house trailer (over 20 feet) - construction buildings - prefab modular homes - other.
14.HOUSEHOLD GOODS	- example: moving vans.
15.MAIL	- mail.
16.OTHER	- empty truck - concrete mixer truck (redimix truck) - other -- specify
17.BAGGED PRODUCTS	- fertilizer - lime - cement - sand - salt - alums- seed or feed products - other -- specify.
18.UNKNOWN	- unknown.

TABLE 3. - COMMODITY CATEGORIES

Commodity Category	Number (m)
General Freight	1
Food Non-Perishables + Food Perishables	2
Forest Proudcts	3
Petroleum Products	4
Bulk Liquids and Chemicals + Bulk Dry Chemicals and Minerals + Bagged Products	5
Heavy Machineray + Metal Products + Construction Materials	6

TABLE 4.: DISTRIBUTION OF COMMODITIES BY AVERAGE SHIPMENT SIZE AND MODES

Average Shipment Size (Tons)	PERCENTAGE OF CARGO MOVEMENTS BY MODES								
	TRUCK	RAIL	PIPE	AIR	MARINE	BUS	WINTER ROAD	OTHER	Total Flows
<1	86.0	*	1.6	1.1	*	6.7	*	3.4	2756
2-3	97.8	1.1	-	*	*	-	-	*	461
4-5	98.2	1.3	-	-	*	-	-	-	447
6-7	92.2	7.2	-	-	-	-	-	*	153
8-10	96.8	2.2	*	-	*	-	*	-	593
11-20	94.3	5.4	-	-	*	-	*	*	1319
21-30	93.8	6.1	1.2	-	1.2	-	-	-	1116
31-40	74.7	22.7	1.2	-	1.2	-	-	-	166
41-50	3.7	94.4	-	-	1.9	-	-	-	54
51-60	4.4	95.6	-	-	-	-	-	-	91
61-70	-	100.0	-	-	-	-	-	-	14
71-80	-	100.0	-	-	-	-	-	-	31
81-90	-	100.0	-	-	-	-	-	-	9
91-100	10.4[1]	89.6	-	-	-	-	-	-	48

- No flows, * Less than 1%, 1.Forest Products.
Total Number of Flows Recorded 7302, Number of missing data 191.
Only average shipment size upto 100 tons is shown.

TABLE 5. - OBSERVED O-D MATRIX FOR FOOD STUFFS(TONS)[1]

I=↓ J=→	1	2	3	4	5	6	7	8	9	10	11	12
C.Code [2]	9	14	31	13	26	32	25	3	5	39	28	37
M.Hat	1	11	1	49	0[3]	1	1	1	250	1	1	1
Leth.Bri.	29	44	23	1	545	1	0[3]	1	5	2	1	1
Calgary	20	32	79	1	2505	54	1817	2029	250	83	258	68
Edmonton	8	2	3	2	750	2	5	5	4	2	116	2
Reddeer	1	1	1	1	40	1	52	1	1	1	42	1
I =↓ J=→	13	14	15	16	17	18	19	20	21	22	23	24
C.Code	29	41	10	21	24	33	34	12	20	23	7	17
M.Hat	1	1	1	1	74	1	1	1	1	1	1	1
Leth.Bri.	1	1	1	1	1	180	1	1	1	1	1	1
Calgary	67	16	18	1	1	5278	1	1	1	1	1	1
Edmonton	26	2	22	62	47	2285	170	159	218	344	192	1266
Reddeer	13	1	97	1	26	0[3]	80	1	1	1	1	1
I=↓ J=→	25	26	27	28	29	30	31	32	33	34	35	36
C.Code	27	38	4	2	18	36	40	20	30	35	15	16
Med. Hat	1	1	1	1	1	1	1	1	1	1	1	1
Leth.Brid.	1	1	1	1	1	1	1	1	1	1	1	1
Calgary	1	1	1	1	1	1	1	9	16	1	1	1
Edmonton	689	161	214	221	272	208	368	629	910	240	43	3450
Reddeer	1	1	1	1	1	1	1	1	1	1	1	1

1 The above matrix is of dimension 5 x 36, where I=5(origins) and J=36(destinations).
2 C.Code indicates the destination communities as given in Table 1. Example, C.Code 9 is Coleman.
3. All intra trips are equal to zero. M. Hat - Medicine Hat, and Leth.Bri. - LethBridge

TABLE 6. - OBSERVED O-D MATRIX FOR FREIGHT(TONS)[1]

I= ↓ j=→	1	2	3	4	5	6	7	8	9	10	11	12
C.Code [2]	9	14	31	13	26	32	25	3	5	39	28	37
M.Hat	2	1	1	10	0[3]	1	1	1	12	1	1	1
Leth.Bri.	2	14	16	1	8	1	0[3]	77	7	5	1	1
Calgary	103	8	14	2	1078	386	3847	2	388	241	123	41.
Edmonton	70	4	1	1	18	1	422	1	61	38	33	64
Reddeer	1	1	1	1	5	1	16	1	6	1	[illegible]	1

I= ↓ j=→	13	14	15	16	17	18	19	20	21	22	23	24
C.Code	29	41	10	21	24	33	34	8	20	23	7	17
M.Hat	1	1	1	1	1	9	1	1	1	1	1	1
Leth.Bri.	1	1	1	1	1	3	1	1	1	1	1	1
Calgary	49	4	36	2	45	705	10	2	2	2	25	2
Edmonton	8	4	19	17	60	628	53	192	1058	307	189	188
Reddeer	1	1	3	1	7	0[3]	8	1	1	1	1	1

I= ↓ j=→	25	26	27	28	29	30	31	32	33	34	35	36
C.Code	27	38	4	2	18	36	40	20	30	35	15	16
Med. Hat	1	1	1	1	1	1	1	1	1	1	1	1
Leth.Brid.	1	1	1	1	1	1	1	60	1	1	1	1
Calgary	262	2	2	2	2	2	6	86	20	2	2	606
Edmonton	2093	584	188	146	315	95	189	1584	305	93	11	3509
Reddeer	32	1	3	1	1	1	1	9	19	2	1	5

1 The above matrix is of dimension 5 x 36, where i=5(origins) and j=36(destinations).
2 C.Code indicates the destination communities as given in Table . Example, C.Code 9 is Coleman.
3. All intra trips are equal to zero. M. Hat - Medicine Hat, and Leth.Bri. - LethBridge

TABLE 7. - OBSERVED O-D MATRIX FOR FOREST(TONS)[1]

I=↓ J=→	1	2	3	4	5	6	7	8
C.Code [2]	9	31	26	32	25	32	5	39
Leth.Bri.	1	4	120	1	0	1	1	1
Calgary	8	122	635	48	2274	6	103	457
Edmonton	2	2	2	2	260	2	27	2
G.Prairie	1	1	12	1	260	1	1	1
I =↓ J=→	9	10	11	12	13	14	15	16
C.Code	28	29	10	24	33	34	7	17
Leth.Bri.	1	1	1	1	1	1	1	1
Calgary	83	48	13	18	79	2	2	2
Edmonton	4	2	5	339	43	6	287	210
G.Prairie	1	1	1	1	1	1	1	1
I=↓ J=→	17	18	19	20	21	22	23	24
C.Code	38	4	40	35	16	12	22	15
Leth.Brid.	1	1	1	1	1	1	1	1
Calgary	2	2	2	2	2	2	2	2
Edmonton	6	225	48	2	536	85	10	4
G.Prairie	1	1	1	56	1	1	1	1

1 The above matrix is of dimension 4 x 24, where I=4(origins)and j=24(destinations).
2 C.Code indicates the destination communities as given in Table 1. Example: C.code 9 is Coleman 3. All intra trips are equal to zero.
Leth.Bri. - LethBridge, G.Prairie - Grand Prairie

TABLE 8. - OBSERVED O-D MATRIX FOR PETROLEUM(TONS)[1]

I=↓ j=→	1	2	3	4	5	6	7	8	9	10	11	12
C.Code[2]	9	14	34	26	25	3	5	39	28	37	29	41
Leth.Bri.	4	432	313	1	0[3]	1	1	1	8	1	1	1
Calgary	578	600	330	5859	11075	2136	4358	1513	2133	1011	278	184
Edomonton	25	2	12	1780	32	2	101	33	41	2	230	2262
Reddeer	1	1	1	1	40	1	52	1	1	1	1	332
I=↓ j=→	13	14	15	16	17	18	19	20	21	22	23	24
C.Code	10	21	24	33	34	8	19	23	7	17	27	38
Leth.Bri.	1	1	1	1	1	1	1	1	1	1	1	1
Calgary	2	2	25	1484	2	155	2	2	2	2	2	2
Edmonton	1257	506	2015	5424	1144	617	2137	5622	1777	23120	6556	3357
Reddeer	3	0[3]	50	1	1	1	1	1	1	1	1	1
I=↓ j=→	25	26	27	28	29	30	31	32	33	34	35	36
C.Code	4	2	18	36	40	20	30	35	15	16	22	12
Leth.Brid.	1	1	1	1	1	1	1	1	1	1	1	1
Calgary	2	2	2	2	2	2	2	2	2	2	2	2
Edmonton	1953	2687	2448	992	7515	5763	2494	940	274	12263	3652	4538
Reddeer	1	1	1	1	1	1	1	1	1	1	1	1

1 The above matrix is of dimension 4 x 36, where i=4(origins) and j=36(destinations).
2 C.Code indicates the destination communities as given in Table 1. Example, C.Code 9 is Coleman.
3. All intra trips are equal to zero.Leth.Bri. - LethBridge

TABLE 9. - OBSERVED O-D MATRIX FOR BULK LIQ. AND DRY(TONS)[1]

I=↓ J=→	1	2	3	4	5	6	7	8	9	10	11	12
C.Code[2]	9	14	31	13	26	32	6	27	5	39	28	37
M.Hat	1	1	1	15	0[3]	1	10	208	280	1	5	1
Leth.Bri.	1	11	1	9	1	1	2	0[3]	21	1	150	1
Calgary	39	20	162	1	13600	113	1	7649	217	404	767	563
Edmonton	8	2	3	2	750	2	5	5	4	2	116	2
F.Sask.	1	1	1	1	40	1	52	1	1	1	42	1

I=↓ J=→	13	14	15	16	17	18	19	20	21	22	23	24
C.Code	29	41	10	21	24	33	34	8	23	7	17	23
M.Hat	13	1	30	20	1	1	1	1	1	1	1	1
Leth.Bri.	1	1	1	1	1	110	1	1	1	1	1	1
Calgary	118	3	1	21	357	1944	42	37	1	80	1	100
Edmonton	2	1	2	4	1	2888	105	1	772	249	470	4123
F.Sask.	1	4	13	1	90	720	40	1	1	30	1	1

I=↓ J=→	25	26	27	28	29	30	31	32	33	34	35	36
C.Code	38	4	2	18	36	40	20	30	35	15	16	12
Med. Hat	1	72	1	1	1	1	1	1	1	1	1	1
Leth.Brid.	65	1	1	1	1	1	1	1	1	1	1	590
Calgary	255	35	40	1	1	1	266	1	1	1	1	500
Edmonton	1809	25329	243	216	20	3749	9140	302	13	2500	226	91
F.Sask.	1	1	1	256	1	1	392	41	1	1156	1	16

1 The above matrix is of dimension 5 x 36, where I=5(origins) and J=36(destinations).
2 C.Code indicates the destination communities as given in Table 1. Example, C.Code 9 is Coleman.
3. All intra trips are equal to zero. M. Hat - Medicine Hat, Leth.Bri. - LethBridge, and F.Sask. - Fort Saskachewan

TABLE 10. - OBSERVED O-D MATRIX FOR HEAVY MACH.(TONS)[1]

I= ↓ J= →	1	2	3	4	5	6	7	8	9	10	11
C.Code[2]	9	14	31	26	32	25	5	39	28	29	10
M.Hat	1	3	1	0[3]	1	18	16	1	1	1	1
Leth.Bri.	1	2	1	753	1	0[3]	85	1	2	1	1
Calgary	72	14	124	2751	192	4972	625	216	432	24	59
Edmonton	2	2	2	22	2	1246	44	22	27	4	2
Reddeer	1	1	1	1	250	1	1	1	7	1	11

I = ↓ J= →	12	13	14	15	16	17	18	19	20	21	22
C.Code	21	24	33	34	8	19	23	7	17	27	38
M.Hat	1	1	100	1	1	1	1	1	1	1	1
Leth.Bri.	1	1	1	1	1	1	1	1	1	1	1
Calgary	16	79	5139	2	183	2	2	230	2	3900	8
Edmonton	7	162	2705	70	2	72	354	459	43	4563	352
Reddeer	1	1	0[3]	12	1	1	1	1	1	1	1

I= ↓ J= →	23	24	25	26	27	28	29	30	31	32	33
C.Code	4	2	18	36	40	20	30	35	16	12	22
Med. Hat	1	20	1	1	1	1	1	1	1	1	1
Leth.Brid.	1	1	1	1	572	1	1	1	1	1	1
Calgary	20	43	2	2	2	183	2	2	104	13	2
Edmonton	30236	140	93	89	134	2666	1283	87	3175	291	96
Reddeer	1	1	1	1	1	1	1	1	1	1	1

1 The above matrix is of dimension 5 x 33, where i=5(origins) and j=33(destinations).
2 C.Code indicates the destination communities as given in Table 1. Example, C.Code 9 is Coleman.
3. All intra trips are equal to zero. M. Hat - Medicine Hat, and Leth.Bri. - LethBridge

TABLE 11- DISTANCE MATRIX - I (Kilometers)

I= ↓ J=→	1	2	3	4	5	6	7	8	9	10	11	12
C.Code [1]	9	14	31	13	26	32	25	3	5	39	28	37
M.Hat	315	117	183	133	*	11	167	424	108	264	378	309
Leth.Bri.	150	104	38	298	167	176	*	338	155	99	313	267
Calgary	222	325	259	332	288	279	223	138	183	126	93	127
Edmonton	525	606	560	467	539	530	524	429	434	427	223	261
Reddeer	369	471	404	382	412	403	369	273	307	272	68	105
I = ↓ J=→	13	14	15	16	17	18	19	20	21	22	23	24
C.Code	29	41	10	21	24	33	34	8	20	23	7	17
M.Hat	199	262	328	390	437	411	471	458	925	789	439	538
Leth.Bri.	364	331	365	425	393	367	428	393	839	703	440	549
Calgary	325	285	313	358	174	148	208	174	639	503	277	329
Edmonton	401	346	274	215	134	161	152	236	462	296	97	24
Reddeer	316	276	189	214	30	*	64	83	508	401	133	185
I= ↓ J=→	25	26	27	28	29	30	31	32	33	34	35	36
C.Code	27	38	4	2	18	36	40	20	30	35	15	16
Med. Hat	490	447	560	686	816	775	736	1017	1044	1083	973	912
Leth.Brid.	608	551	636	664	773	732	692	973	1001	1039	1013	952
Calgary	522	465	535	445	553	512	473	754	782	820	793	733
Edmonton	247	190	233	149	261	261	181	462	495	528	491	430
Reddeer	378	321	391	301	409	368	329	610	637	676	649	588

1. C.Code indicates the destination communities as given in Table 1. Example, C.Code 9 is Coleman.
* Internal distances. M. Hat - Medicine Hat, and Leth.Bri. - LethBridge

TABLE . 12- DISTANCE MATRIX - II (Kilometers)

I= ↓ j=→	1	2	3	4	5	6	7	8	9	10	11	12
C.Code [1]	9	14	31	13	26	32	6	25	5	39	28	37
F.Sask.	549	605	584	466	538	529	574	549	433	452	248	252
G.Prairie	974	1076	1009	933	1017	1008	564	973	912	876	672	386

I= ↓ j=→	13	14	15	16	17	18	19	20	21	22	23	24
C.Code	29	41	10	21	24	33	34	8	23	7	17	27
F.Sask.	400	345	273	214	159	185	177	260	316	96	*	244
G.Prairie	867	826	754	676	583	610	538	574	343	564	470	711

I= ↓ j=→	25	26	27	28	29	30	31	32	33	34	35	36
C.Code	38	4	2	18	36	40	20	30	35	16	22	12
F.Sask.	186	217	156	269	228	189	470	502	536	426	799	229
G.Prairie	654	676	654	203	344	283	*	1008	81	756	457	386

1. C.Code indicates the destination communities as given in Table 1. Example, C.Code 9 is Coleman. F.Sask. - Fort Saskatchewan, and G.Prairie - Grand Prairie. * Internal distance.

TABLE 13. - λ, R^2, and r VALUES FOR FOOD STUFFS

λ	R^2	r
-2.00	0.29	0.53
-1.75	0.34	0.58
-1.50	0.41	0.64
-1.25	0.51	0.71
-1.00	0.61	0.78
-0.75	0.68	0.82
-0.50	0.71	0.84
-0.25	0.70	0.83
0.00	0.61	0.78

TABLE 14.- λ, R^2, and r VALUES FOR COMMODITY CATEGORIES

Commodity Category and m	λ								
	-2.00	-1.75	-1.50	-1.25	-1.00	-0.75	-0.50	-0.25	0.0
1.General Freight									
R^2	0.11	0.21	0.37	0.57	0.70	0.74	0.72	0.67	0.54
r	0.33	0.45	0.61	0.75	0.84	0.86	0.85	0.82	0.74
2.Food Stuffs									
R^2	0.29	0.34	0.41	0.51	0.61	0.68	0.71	0.70	0.61
r	0.53	0.58	0.64	0.71	0.78	0.82	0.84	0.83	0.78
3.Forest									
R^2	0.48	0.56	0.66	0.77	0.85	0.88	0.86	0.82	0.73
r	0.70	0.75	0.81	0.88	0.92	0.93	0.93	0.91	0.85
4.Petroleum									
R^2	0.50	0.51	0.52	0.53	0.56	0.60	0.67	0.73	0.50
r	0.71	0.71	0.72	0.73	0.75	0.78	0.82	0.85	0.71
5.Bulk Dry and Liq.									
R^2	0.27	0.48	0.66	0.75	0.79	0.78	0.75	0.70	0.58
r	0.52	0.69	0.81	0.87	0.89	0.88	0.87	0.84	0.76
6.Heavy Mach.									
R^2	0.78	0.82	0.84	0.86	0.87	0.86	0.85	0.84	0.78
r	0.88	0.90	0.92	0.93	0.93	0.93	0.92	0.91	0.88

TABLE 15. - OPTIMIZED GRAVITY MODELS

Commodity Category	m	Models
General Freight	1	$T_{ij} = P_i \frac{C_j d_{ij}^{-0.75}}{\Sigma_j C_j d_{ij}^{-0.75}}$
Food Non-Perishables and Food Perishables	2	$T_{ij} = P_i \frac{C_j d_{ij}^{-0.5}}{\Sigma_j C_j d_{ij}^{-0.5}}$
Forest Proudcts	3	$T_{ij} = P_i \frac{C_j d_{ij}^{-0.75}}{\Sigma_j C_j d_{ij}^{-0.75}}$
Petroleum Products	4	$T_{ij} = P_i \frac{C_j d_{ij}^{-0.25}}{\Sigma_j C_j d_{ij}^{-0.25}}$
Bulk Liquids and Chemicals + Bulk Dry Chemicals and Minerals + Bagged Products	5	$T_{ij} = P_i \frac{C_j d_{ij}^{-1.0}}{\Sigma_j C_j d_{ij}^{-1.0}}$
Heavy Machineray + Metal Products + Construction Materials	6	$T_{ij} = P_i \frac{C_j d_{ij}^{-1.0}}{\Sigma_j C_j d_{ij}^{-1.0}}$

TABLE 16. - MODAL SHARE BY AVERAGE SHIPEMENT SIZE AND COMMODITY CLASSIFICATIONS

Commodity Classifications	Average Shipment Size(Tons) <1	2-3	4-5	6-7	8-10	11-20	21-30
General Freight	2	2	2	2	2	2	2
Food Non-Perishables	T	T	T	T	T	T	T
Food Perishables	T	T	T	T	T	T	T
Heavy Machinery	2	2	2	2	2	2	2
Metal Products	T	T	T	2	2	2	2
Petro. Products	T	T	T	T	T	T	2
Bulk Liquids	T	T	T	T	T	T	T
Bulk Dry	T	T	T	T	T	T	T
Forest Products	T	T	T	T	T	T	2
Live Stock	T	T	T	T	T	T	2
Construction Materials	T	T	T	T	T	2	2
Feed Products	T	T	T	T	T	T	T
Travel Mobile Homes	T	T	T	T	T	T	T
Bagged Products	T	T	T	T	T	T	2
Other	T	T	T	T	T	T	T
Total Flows	2756	461	447	153	593	1319	1116

Commodity Classifications	31-40	41-50	51-60	61-70	71-80	81-90	>90
General Freight	R	R	R	R	R	R	R
Food Non-Perishables	2	2	2	2	2	2	2
Food Perishables	T	T	T	T	2	2	2
Heavy Machinery	2	R	R	R	R	R	R
Metal Products	R	R	R	R	R	R	R
Petro. Products	2	R	R	R	R	R	R
Bulk Liquids	2	R	R	R	R	R	R
Bulk Dry	2	2	R	R	R	R	R
Forest Products	2	R	R	R	R	R	R
Live Stock	R	-*	-	-	-	-	-
Construction Materials	2	R	R	R	R	R	R
Feed Products	2	R	R	R	R	R	R
Travel Mobile Homes	-	-	-	-	-	-	-
Bagged Products	2	R	R	R	R	R	R
Other	-	-	-	-	-	-	-
Total Flows	166	54	91	14	31	9	92

2-represents both Truck and Rail. * indicates no flows.

TABLE 17. DISTRIBUTION OF COMMODITIES BY MODE AND LOADS

	LOADS		
MODES	FULL LOAD(%)	LESS THAN FULL LOAD(%)	TOTAL FLOWS
TRUCK	44.7	55.3	6629
RAIL	90.5	9.5	473
PIPE	88.9	11.1	9
AIR	0.0	100.0	34
MARINE	12.5	87.5	16
BUS	1.1	98.9	181
WINTER ROAD	42.9	57.1	7
OTHER	5.3	94.7	94
TOTAL FLOWS	3000	4460	7460

1. Number of Missing Data 33

TABLE 18. DISTRIBUTION OF COMMODITIES BY AVERAGE SHIPMENT SIZE, MODE AND LOAD

AVERAGE SHIPMENT SIZE (TONS)	TRUCK MODE			RAIL MODE		
	FL	LTFL	TOTAL	FL	LTFL	TOTAL
<1	39	2330	2369	3	21	23
2-3	18	326	344	0	4	4
4-5	35	160	195	1	2	3
6-7	102	333	435	2	8	10
8-10	368	286	654	15	2	17
11-20	1152	92	1244	69	2	71
21-30	1039	8	1047	65	3	68
31-40	123	1	124	38	0	38
41-50	2	0	2	49	2	51
51-60	4	0	4	87	0	87
61-70	0	0	0	14	0	14
71-80	0	0	0	31	0	31
81-90	0	0	0	9	0	9
91-100[1]	5	0	5	43	0	43
TOTAL FLOWS	2887	3536	6423	426	44	470

1. Average shipments up to 100 tons are only tabulated.

TABLE 19. - DISTRIBUTION BY AVERAGE SHIPEMENT SIZE, COMMODITY CLASSIFICATIONS, AND LOADS

Commodity Classifications	Average Shipment Size(Tons) and Full Load						
	<1	2-3	4-5	6-7	8-10	11-20	21-30
General Freight	11	7	23	3	40	125	46
Food Non-Perishables	1	1	7	0	18	81	22
Food Perishables	1	1	8	11	13	85	32
Heavy Machinery	4	3	13	13	68	261	48
Metal Products	3	4	2	1	11	69	70
Petro. Products	13	1	14	36	42	105	167
Bulk Liquids	1	2	0	1	8	46	47
Bulk Dry	2	0	0	0	12	61	143
Forest Products	4	4	3	0	8	134	149
Live Stock	0	1	2	4	19	46	55
Construction Materials	2	2	11	2	38	60	108
Feed Products	2	2	5	1	19	45	97
Travel Mobile Homes	4	4	12	6	30	36	2
Bagged Products	2	0	4	0	6	64	117
Other	0	0	0	0	0	0	0

Commodity Classifications	Average Shipment Size(Tons) and Less Than Full Load						
	<1	2-3	4-5	6-7	8-10	11-20	21-30
General Freight	1837	185	92	11	78	8	1
Food Non-Perishables	166	29	39	12	25	12	2
Food Perishables	259	49	34	6	17	10	1
Heavy Machinery	93	27	40	24	28	72	0
Metal Products	46	24	28	2	7	2	0
Petro. Products	71	24	40	15	34	29	1
Bulk Liquids	30	12	8	0	19	2	2
Bulk Dry	[illegible]	1	6	0	4	2	0
0Forest Products	54	18	10	0	13	8	1
Live Stock	3	2	2	0	0	0	0
Construction Materials	77	43	29	2	18	2	1
Feed Products	5	1	2	1	4	2	0
Travel Mobile Homes	0	5	1	2	3	2	0
Bagged Products	12	8	16	0	11	2	0
Other	3	4	0	0	0	0	0

TABLE 20. DISTRIBUTION OF COMMODITIES BY COMMODITY CLSASSIFICATION, MODE AND HIRE

		MOVEMENT BY TRUCKS		
NO.	COMMODITY	PRIVATE	FOR-HIRE	TOTAL FLOW
1	[illegible] - Freight	1944	249	2193
2	Food - Non Perishable	213	210	423
3	Food - Perishables	280	272	552
4	Heavy Machinery	499	56	555
5	Metal Products	230	31	261
6	Petro Products	409	260	669
7	Bulk Liquids	151	30	181
8	Bulk Dry	214	37	251
9	Forest Products	353	46	399
10	Live Stock	101	25	126
11	Construction Materials	356	45	401
12	Feed Products	147	42	189
13	Trav. Mobil Homes	99	11	110
14	Bagged Products	222	20	242
15	Other	3	12	15
	TOTAL	5221	1346	6557

TABLE 21.- DISTRIBUTIOIN OF COMMODITIES BY MODES, LOADS, AND HIRE

	TRUCK MODE		
LOADS	PRIVATE	FOR-HIRE	TOTAL FLOWS
FL	2313	588	2901
LTFL	2907	758	3665
TOTAL	5220	1346	6566

TABLE 22. MODE SELECTION BY CONTROL

	CONTROL		
MODES	YES (%)	NO (%)	TOTAL FLOWS
TRUCK	41.5	58.5	6609
RAIL	30.9	69.1	473
PIPE	47.7	52.3	44
AIR	67.6	32.4	34
MARINE	62.5	37.5	16
BUS	23.2	76.8	181
WINTER ROAD	42.9	57.1	7
OTHER	13.5	86.5	96
Total Flows	3000	4460	7460

1. Number of Missing Data 33.

TABLE 23. CONTINGENCY TABLE

Variable A	Variable B		
	B1	B2	Total
A1	X_{11}	X_{12}	X_{1+}
A2	X_{21}	X_{22}	X_{2+}
Total	X_{+1}	X_{+2}	X_{++}*

* Total number of observations

TABLE 24. MODE PREFERENCE FOR [LM][HM] MODEL

LOADS(L)	HIRE(H)	MODES(M)		
		TRUCK(T)	RAIL(R)	TOTAL
FL	Private	2313	425	2738
	For - Hire	588	3	591
LTFL	Private	2907	45	2952
	For - Hire	758	0	758
Total		6566	473	7039

Table 25. DEGREES OF FREEDOM AND COMPUTED G^2 STATISTICS

MODEL		D.F	G^2
M_1 :	[L]	6	9778.56
M_2 :	[H]	6	6919.60
M_3 :	[M]	6	3508.87
M_4 :	[L][H]	5	6898.97
M_5 :	[H][M]	5	629.27
M_6 :	[M][L]	5	3488.23
M_7 :	[L][H][M]	4	608.64
M_8 :	[LH]	4	6890.83
M_9 :	[LM]	4	3061.35
M_{10}:	[HM]	4	448.29
M_{11}:	[L][HM]	3	427.66
M_{12}:	[H][LM]	3	181.75
M_{13}:	[M][LH]	3	600.50
M_{14}:	[LH][LM]	2	173.61
M_{15}:	[LM][HM]	2	0.77 *
M_{16}:	[HM][LH]	2	419.51
M_{17}:	[LH][LM][HM]	1	0.62 *

* Selected Models for Comparison

TABLE 26. ESTIMATED EFFECTS AND LOGISTICS FOR [LM][HM] MODEL

VARIABLE	MODE PREFERENCE		W TERMS
	TRUCK	RAIL	
μ	+ 2.392	- 2.392	+ 4.784
$\mu_{LM(i1)}$			
Full Load (FL)	- 0.620	+ 0.620	- 1.240
Less Than Full Load (LTFL)	+ 0.620	- 0.620	+ 1.240
$\mu_{HM(j1)}$			
Private	- 0.930	+ 0.930	- 1.860
For - Hire	+ 0.930	- 0.930	+ 1.860

TABLE 27. STANDARDIZED RESIDUALS FOR [LM][HM] MODEL

MODES	HIRE	LOADS	
		FULL LOAD	LESS THAN FULL LOAD
Truck	Private	0.1	-0.1
	For-Hire.	-0.3	0.2
Rail	Private	0.0	0.0
	For-Hire	0.2	-0.5

TABLE 28. SELECTED MODELS

Model Number	List of Variables	Explanatory Variables	Response Variables	Selected Model	D.F	α
M1	L, H, M	L,H	M	[LM] [HM]	2	0.05
M2	C, L, M	C,L	M	[CL] [CM] [LM]	1	0.05
M3	S, L, M	S,L	M	[SL] [SM] [LM]	8	0.05
M4	K, L, M	K,L	M	[KL] [KM] [LM]	2	0.01
M5	D, K, M	D,K	M	[DK] [KM] [DM]	15	0.05

TABLE 29. MODE PREFERENCE [LC][CM][LM] MODEL

LOADS(L)	CONTROL(C)	MODES(M) TRUCK(T)	RAIL(R)	TOTAL
FL	YES	1424	142	1566
	NO	1519	286	1805
LTFL	YES	1318	4	1322
	NO	2347	41	2388
Total		6608	473	7081

TABLE 30. DEGREES OF FREEDOM AND COMPUTED G^2 STATISTICS [LC]CM][LM] MODEL

MODEL	D.F	G^2
M_1 : [C]	6	5697.14
M_2 : [L]	6	5997.75
M_3 : [M]	6	805.68
M_4 : [C][L]	5	5604.29
M_5 : [L][M]	5	779.95
M_6 : [M][C]	5	513.05
M_7 : [C][L][M]	4	499.81
M_8 : [CL]	4	5424.63
M_9 : [CM]	4	495.44
M_{10}: [LM]	4	370.84
M_{11}: [C][LM]	3	134.92
M_{12}: [L][CM]	3	480.31
M_{13}: [M][CL]	3	438.15
M_{14}: [CL][CM]	2	400.10
M_{15}: [CM][LM]	2	115.04
M_{16}: [LM][CL]	2	48.95 *
M_{17}: [CL][CM][LM]	1	4.74 *

* Selected Models for Comparison

TABLE 31. ESTIMATED EFFECTS AND LOGISTICS FOR [LC][CM][LM] MODEL

VARIABLE	MODE PREFERENCE: TRUCK	MODE PREFERENCE: RAIL	W TERMS
μ	+1.638	-1.638	+3.276
$\mu_{CM(i1)}$			
YES	-0.638	+0.638	-1.276
NO	+0.638	-0.638	+1.276
$\mu_{LM(j1)}$			
FULL	+0.176	-0.176	+0.352
LTFL	-0.176	+0.176	-0.352

TABLE 32. STANDARDIZED RESIDUALS FOR [LC][CM][LM] MODEL

MODES	CONTROL	LOADS: FL	LOADS: LTFL
Truck	YES	-0.2	0.2
	NO	0.1	-1.0
Rail	YES	0.5	-1.8
	NO	-0.3	1.0

TABLE 33. MODE PREFERENCE FOR [SL][SM][LM] MODEL

LOADS(L)	AV.SHIPMENT SIZE(TONS)	MODES(M) TRUCK	RAIL	TOTAL
FL	A	39	3	42
	B	18	0	18
	C	13	0	13
	D	23	2	25
	E	250	8	258
	F	366	18	384
	G	243	31	274
	H	1759	68	1827
	I	171	40	201
	J	3	25	8
	K	6	135	141
	TOTAL	2891	330	3221
LTFL	A	2330	21	2371
	B	326	4	330
	C	105	1	106
	D	61	1	62
	E	412	9	421
	F	236	2	238
	G	40	1	41
	H	32	2	34
	I	2	1	3
	J	0	0	0
	K	0	0	0
	TOTAL	3544	42	3586

A = <1 TON, B = 2 TO 3 TONS, C=3-4, D=4-5, E=5-10, F=10-15,
G=15-20,H=20-30,I=30-40,J=40-5O, K=50- 60(DENOTES AVERAGE SHIPMENT SIZE).

TABLE 34. DEGREES OF FREEDOM AND COMPUTED G^2 STATISTICS [SL][SM][LM] MODEL

MODEL	D.F	G^2
M_1: [S]	6	15289.13
M_2: [L]	6	51290.39
M_3: [M]	6	26103.15
M_4: [S][L]	5	15226.78
M_5: [L][M]	5	25839.77
M_6: [P][S]	5	10591.19
M_7: [S][L][M]	4	10880.39
M_8: [SL]	4	5989.44
M_9: [SM]	4	4952.11
M_{10}: [LM]	4	25018.18
M_{11}: [S][LM]	3	955.65
M_{12}: [L][SM]	3	4946.76
M_{13}: [M][SL]	3	2850.79
M_{14}: [SL][SM]	2	30.26 *
M_{15}: [SP][LM]	2	4854.28
M_{16}: [LM][SL]	2	1484.01
M_{17}: [SL][SM][LM]	1	15.02 *

* Selected Models for Comparison

TABLE 35. ESTIMATED EFFECTS AND LOGISTICS FOR [SL][SM][LM] MODEL

	MODE PREFERENCE		
VARIABLE	TRUCK	RAIL	W TERMS
μ	+1.229	- 1.229	+ 2.458
$\mu_{SM(i1)}$			
A	0.842	- 0.842	1.684
B	0.798	- 0.798	1.596
C	0.997	-0.997	1.994
D	0.383	-0.383	0.766
E	0.598	-0.598	1.196
F	0.575	-0.575	1.150
G	0.055	-0.055	0.110
H	0.625	-0.625	1.250
I	-0.275	0.275	-0.550
J	-2.051	2.051	-4.102
K	-2.547	+2.547	-5.094
$\mu_{LM(j1)}$			
FULL	-0.239	+0.259	-0.478
LTFL	+0.239	-0.239	+0.478

TABLE 36. MODE PREFERENCE BY AVERAGE SHIPMENT SIZE AND LOAD

No.	Average Shipment Size(Tons)	Truck Mode Proportions*	
		Full Load	Less Than Full Load
		P_{i1}	P_{i2}
1	< 1	97.48	99.20
2	2 - 3**	97.27	98.91
3	3 - 4	98.16	99.26
4	4 - 5	93.96	97.54
5	5 - 10	95.98	98.37
6	10 - 15	95.79	98.30
7	15 - 20	88.93	98.35
8	20 - 30	96.15	98.46
9	30 - 40	80.52	91.37
10	40 - 50	10.06	23.26
11	50 - 60 @	4.21	10.15

* i=1,2,3.........11, j=1,2 and k=1,2. Model Selected [SL][SM][LM]
** Denotes Average Shipment Size greater than 2 tons and less than 3 tons
@ Average Shipment Size upto 60 tons are only considered.

TABLE 37. STANDARDIZED RESIDUALS FOR [SL][SM][LM] MODEL

MODES	AV.SHIPMENT SIZE(TONS)	LOADS FL	LOADS LTFL
Truck	A	-0.3	0.0
	B	0.1	0.0
	C	0.1	0.0
	D	-0.1	0.1
	E	0.1	-0.1
	F	-0.1	0.1
	G	-0.1	0.1
	H	0.0	-0.3
	I	0.1	-0.5
	J	0.0	0.0
	K	0.0	0.0
Rail	A	0.5	-1.8
	B	-0.3	1.0
	C	-0.5	0.3
	D	0.4	-0.4
	E	-0.7	0.9
	F	0.5	-1.0
	G	0.2	-0.6
	H	-0.2	2.1
	I	-0.1	1.5
	J	0.0	0.0
	K	0.0	0.0

TABLE 38. MODE PREFERENCE FOR [KL][KM][LM] MODEL

		MODES(M)		
LOADS(L)	COMMODITY	TRUCK	RAIL	TOTAL
FL	1	242	31	273
	2	144	9	153
	3	192	23	215
	4	336	83	419
	5	147	24	171
	6	454	45	499
	7	104	29	133
	8	235	69	304
	9	300	55	355
	10	120	9	129
	11	222	12	234
	12	172	22	194
	13	96	1	97
	14	8	0	8
	15	191	16	207
TOTAL		2963	428	3381
LTFL	1	1951	31	1982
	2	299	1	300
	3	403	0	403
	4	219	9	228
	5	114	0	114
	6	215	0	215
	7	77	0	77
	8	16	0	16
	9	99	3	102
	10	6	0	6
	11	178	1	179
	12	17	0	17
	13	14	0	14
	14	7	0	7
	15	0	15	15
TOTAL		3544	42	3586

1-General Freight, 2-Food Non-perishables, 3-Food Perishables,4-Heavy Machinery,5-Metal Products,6-Petroleum Products,7-Bulk Liquids,8-Bulk Dry,9-Forest Products,10-Live Stock, 11-Construction Materials,12-Feed Products,13-Travel -Mobile Homes,14-Bagged Products, 15-Other.

TABLE 39 : DEGREES OF FREEDOM AND COMPUTED G^2 STATISTICS [KL][KM][LM] MODEL

MODEL		D.F	G^2
M_1:	[K]	6	9889.57
M_2:	[L]	6	34329.95
M_3:	[M]	6	16971.65
M_4:	[K][L]	5	9807.45
M_5:	[L][M]	5	16450.97
M_6:	[M][S]	5	3343.57
M_7:	[K][L][M]	4	3350.50
M_8:	[KL]	4	5426.22
M_9:	[KM]	4	2718.60
M_{10}:	[LM]	4	15314.03
M_{11}:	[K][LM]	3	2719.35
M_{12}:	[L][KM]	3	2712.29
M_{13}:	[M][KL]	3	560.48
M_{14}:	[KL][KM]	2	258.95
M_{15}:	[KM][LM]	2	2461.30
M_{16}:	[M][KL]	2	139.05 *
M_{17}:	[KL][KM][LM]	1	16.33 *

* Selected Models for Comparison

TABLE 40. ESTIMATED EFFECTS AND LOGISTICS FOR [KL][KM][LM] MODEL

VARIABLE	MODE PREFERENCE		W TERMS
	TRUCK	RAIL	
μ_M	+1.728	- 1.728	+ 3.456
$\mu KM(i1)$			
1	-0.178	0.178	-0.356
2	0.259	- 0.259	0.590
3	0.029	-0.029	0.058
4	-0.454	0.454	-0.908
5	-0.181	0.181	-0.362
6	0.053	-0.053	0.106
7	-0.448	0.448	-0.896
8	-0.509	0.509	-1.018
9	-0.291	0.291	-0.582
10	0.172	-0.172	0.344
11	0.328	-0.328	0.656
12	-0.092	0.092	-0.184
13	1.163	-1.163	2.326
14	0.0	0.0	0.0
15	0.144	-0.144	0.228
$\mu LM(j1)$			
FULL	-0.603	+0.603	-1.206
LTFL	+0.603	-0.603	+1.206

TABLE 41. MODE PREFERENCE BY COMMODITY CLASSIFICATION AND LOAD

No	Major Commodity	Truck Mode Proportions	
		Full Load	Less Than Full Load
		P_{i1}	P_{i2}
1	General Freight *	86.84	98.66
2	Food - Non Perishables	94.44	99.47
3	Food Perishables	90.09	99.11
4	Heavy Machinery	79.14	97.70
5	Metal Products	86.77	98.65
6	Petro Products	91.26	99.15
7	Bulk Liquids	73.30	97.74
8	Bulk Dry	77.23	97.44
9	Forest Products	84.10	98.32
10	LiveStock	93.02	99.33
11	Construction Materials	94.79	99.50
12	Feed Products	88.75	98.86
13	Travel-Mobile Homes	98.97	99.90
14	Bagged Products	**	**
15	Other	92.25	99.25

* i=1,2,3,..............15, j= 1,2, and k=1,2. Model Selected [GL][GM][LM].

** Not Calculated

TABLE 42. STANDARDIZED RESIDUALS FOR [KL][KM][LM] MODEL

MODES	AV.SHIPMENT SIZE(TONS)	LOADS PL	LTFL
Truck	1	0.3	-0.1
	2	0.0	0.0
	3	-0.3	0.2
	4	0.2	-0.3
	5	-0.1	0.1
	6	-0.1	0.1
	7	-0.2	0.2
	8	0.0	-0.1
	9	0.1	-0.1
	10	0.0	0.0
	11	0.0	0.0
	12	0.0	0.0
	13	0.0	0.0
	14	0.0	0.0
	15	0.0	0.0
Rail	1	-0.8	0.9
	2	0.2	-0.4
	3	0.8	-1.9
	4	-0.4	1.7
	5	0.3	-1.2
	6	0.3	-1.3
	7	0.3	-1.3
	8	0.0	-0.6
	9	-0.2	1.0
	10	0.0	-0.2
	11	0.0	0.1
	12	0.0	-0.4
	13	0.0	-0.1
	14	0.0	0.0
	15	0.0	0.0

TABLE 43 : DEGREES OF FREEDOM AND COMPUTED G^2 STATISTICS [DK][KM]DM] MODEL

MODEL	D.F	G^2
M_1 : [D]	44	2915.78
M_2 : [K]	42	3279.08
M_3 : [M]	46	1448.01
M_4 : [D][K]	39	2556.60
M_5 : [K][M]	41	1098.84
M_6 : [M][D]	43	735.54
M_7 : [D][K][M]	38	386.36
M_8 : [DK]	24	2229.14
M_9 : [DM]	40	724.89
M_{10}: [KM]	36	1089.41
M_{11}: [D][KM]	33	376.93
M_{12}: [K][DM]	35	375.72
M_{13}: [M][KL]	23	48.90
M_{14}: [M][DK]	20	38.25
M_{15}: [DK][DM]	30	366.29
M_{16}: [KM][DK]	18	39.47
M_{17}: [DK][DM][KM]	15	21.04*

* Selected model

TABLE 44 . ESTIMATED EFFECTS AND LOGISTICS FOR [DK][KM][DM] MODEL

VARIABLE	MODE PREFERENCE		W TERMS
	TRUCK	RAIL	
μ_M	+2.104	- 2.104	+ 4.208
$\mu_{DM(i1)}$			
A *	0.803	-0.803	+1.606
B *	-0.392	0.392	-0.784
C *	0.326	-0.326	+0.652
D *	-0.737	0.737	-1.474
$\mu_{KM(j1)}$			
GEN. FREIGHT	0.556	-0.556	+1.112
FOOD STUFFS	-0.822	+0.882	-1.644
FOREST	-0.294	+0.294	-0.568
PETROLEUM	0.030	-0.030	+0.06
BULK LIQ./DRY	0.111	+0.111	-0.222
HEAVY MACH.	0.642	-0.642	+1.284

* Are the distances ranges given in Table 45

TABLE 45. MODE PREFERENCE BY DISTANCE AND COMMODITY CATEGORY

	Truck Mode Proportions			
	Distance(Kilometers)			
Commodity Category	0 -200	201- 400	401- 600	>600
GEN. FREIGHT	99.9	98.9	99.7	97.9
FOOD STUFFS	98.5	85.0	96.0	75.0
FOREST	99.4	94.5	98.6	89.7
PETROLEUM PRODUCTS	99.7	96.9	99.3	94.2
BULK LIQ. AND DRY	99.6	96.1	99.0	93.0
HEAVY MACH.	99.9	99.1	99.7	98.2

TABLE 46. STANDARDIZED RESIDUALS FOR [DK][KM][DM] MODEL

Modes	Commodity Category	0 -200	201- 400	401- 600	>600
Truck	GEN. FREIGHT	0.0	-0.1	0.0	0.1
	FOOD STUFFS	0.1	-0.9	0.1	-0.4
	FOREST	0.0	0.0	-0.2	-0.5
	PETROLEUM	0.0	0.0	0.1	0.2
	BULK LIQ. AND DRY	-0.1	0.1	-0.1	0.2
	HEAVY MACH.	-0.1	0.0	0.0	0.1
Rail	GEN. FREIGHT	-0.2	0.5	-0.4	-0.6
	FOOD STUFFS	-0.5	2.3	-0.4	0.7
	FOREST	-0.5	0.0	1.4	-0.7
	PETROLEUM	-0.7	-0.2	-0.7	1.9
	BULK LIQ. AND DRY	1.6	-0.5	1.1	-0.7
	HEAVY MACH.	2.4	-0.1	-0.5	-0.7

TABLE 47. - EXAMPLE OF DISTRIBUTION AND MODAL SPLIT FOR FOOD STUFFS

		Sink				
Source	Com.Code	17	38	7	33	34
EDMONTON	Observed*	1266	161	192	2	170
	Estimated*	1977	92	151	20	156
	By Truck**	1947	90	148	19	153
	By Rail**	30	2	3	1	3
	Distance***	24	190	151	161	152

		Sink					
Source	Com. Code	28	33	29	5	26	14
CALGARY	Observed*	258	5278	67	250	2505	32
	Estimated*	304	4467	42	265	1590	35
	By Truck**	300	4400	36	261	1351	30
	By Rail**	4	67	6	4	239	5
	Distance***	93	148	325	183	288	325

* Interchanges in Tons. ** Modal Split Proportions for estimated flows are obtained from Table 45. *** Distance between source and sink in Kilometers. Com.Code =Community Code, from Table 1. For example, 17 represents Fort Saskachewan.

APPENDIX - I

APPENDIX - I

CODING INSTRUCTIONS AND SURVEY DATA FORMAT

The following is the coding instructions delivered by the Alberta Transportation Department to the shippers and consignees across Alberta to collect necesssary data during commodity flow survey. It consists of general information, item by item description of coding requested and methodology to fill the survey form shown in Fig.6.. A sample of the data compiled on tape from the commodity flow survey foms is illustrated.

THE ALBERTA TRANSPORTATION COMMODITY FLOW SURVEY
CODING INSTRUCTIONS

1.0 GENERAL INFORMATION

- The Survey is directed to the major tonnage shippers and consignees within each community. Smaller tonnage firms may be either: (1) Interviewed by telephone only.
  (2) Sampled and forms coded for them on the basis of extrapolated information (indicate thereon).

- Emphasis is on 'non-local trips. Where a business handles <u>exclusively</u> local freight (e.g. - a gas station supplied by the local bulk dealer), no form need be coded (to avoid double counting). However, an industry that gathers some of its inputs locally and also ships externally should be completely coded for the sake of material balance.

e.g. - a feed mill	
<u>INPUTS</u>	<u>OUTPUTS</u>
x tons grain - local sources	(x + y) tons - feed - to local feed lots and farmers.
y tons supplements - Edmonton	

- Explain to shippers/consignees that the interview will assist Alberta Transportation to plan highways for their area. That volume information requested is on the basis of approximate estimates only, is not directed at any confidential corporate information, and all responses will remain strictly confidential in nature.

- Be attentive to shipper concerns and comments and note these on the back of the form.

- Do not write in shaded areas of the form.

- Be flexible. Different shippers have different methods of estimating tonnage. (Note in Calculations/Remarks)

  e.g. - Bulk Fuel - x million gallons/year.
  Foodstores - x times per week, y lbs. average estimated shipment size.
  Others - average tariff $y/hundred. Monthly freight bill is approximately $z.

- For 'Firm Type' and 'Goods Type' fields, select that category which predominantly describes the operation or product involved. If no category fits very well, code 99 in these fields.

- In coding the interview form, print the information neatly in the spaces provided. Remember that a keypunch operator must be able to read the information coded. Some fields have been arranged for direct coding as you interview (e.g. - goods type, received or shipped, mode, LTL vs. FL, control, private/for hire, rate, market share) while others have a place to print the information longhand for later coding (after the interview) (e.g. - place, commodity, tonnage).

- Cards 1 and 2 may be pre-coded in the office prior to the field trip based on the telephone book listing and telephone interview. Provision is made to note the individual's name in order to assist you when you make your visit.

- If information is not available (e.g. - rate, market % estimate), leave these items blank. Attempt, if possible, to get the tonnage, shipment size, mode, place data, in all cases.

- Keep forms for each community in a separate manila file and note any important community features or concerns on the file.

- If it is necessary to code more than 5 '3' category cards, the interviewer must continue on a second form. To facilitate this, re-code the telephone number (columns 46 - 52 of card 1) and code 'C' in column 54 of card 2. The balance of the header information (cards 1 and 2) may be left blank. An unlimited number of such continuations is allowed. Staple the forms together sequentially.

- The question on principal competitors (lower right hand portion of the form) is to ensure that all major firms are interviewed.

2.0 ITEM BY ITEM DESCRIPTION

Card	Item	Space	Instructions
CARD 1	Place Name	Cols. 2 - 21	Alphabetical place name. Abbreviate if necessary (consistently).
	Place Code	Cols. 22 - 25	Numeric. From list[7] supplied.
	Firm Name	Cols. 26 - 45	Alphabetic firm name. Abbreviate if necessary.
	Telephone	Cols. 46 - 52	Numeric.
CARD 2	Street Address	Cols. 2 - 45	Code one character per space provided.
	Postal Code	Cols. 46 - 51	Code one character per space.
	Firm Type	Cols. 52, 53	Numeric using code on sheet.
	Continuation	Col. 54	Leave blank if first card. Code 'C' if continuation card.
CARD 3	Goods Type	Cols. 2, 3	Numeric as per code on sheet.
	Received/Shipped	Col. 4	'1' or '2' as instructed on sheet.
	From/To 'Place'	Cols. 5 - 8	Alphabetic in space provided above. Numeric from list provided.

Commodity	Cols. 9 - 12	Alphabetic in space above. Numeric from list provided.
Annual Tonnage	Cols. 13 - 20	Derived (see calculations/ remarks).
Average Shipment Size	Cols. 21 - 24	In tons. Decimal place is between column 23 and 24.
Mode	Col. 25	Numeric using code on sheet.
LTL or FL	Col. 26	L or F depending if shipment shares space on the vehicle.
Control	Col. 27	Ask the shipper if he decides how the product is shipped. Y - Yes, N - No.
Private or For Hire Carriage	Col. 28	P or H.
Rate	Cols. 29 - 33	Ask this item last. It antagonizes some shippers.
Estimated % Market Share	Cols. 34 - 36	Explain that the purpose is to provide for all traffic to/from a community and that this question assists with this estimate.

```
    166LETHBRIDGE             2627SHELL AGRI-CHEMICALS3284735
3004 1 AVE SOUTH                                      07
211300501100000002500501LNH
    167LETHBRIDGE             2627SHELL AGRI-CHEMICALS3284735
3004 1 AVE SOUTH                                      07
211300501010000005002501FNH
    168LETHBRIDGE             2627LAVERS BULK FUELS   3272762
1915 2 AVE SOUTH                                      07
111300506030000056000250 1FNH
    169LETHBRIDGE             2627LAVERS BULK FUELS   3272762
1915 2 AVE SOUTH                                      07
111300506030000520002502FNH
    170LETHBRIDGE             2627LAVERS BULK FUELS   3272762
1915 2 AVE SOUTH                                      07
111300508040000315006002FNH
    171LETHBRIDGE             2627LAVERS BULK FUELS   3272762
1915 2 AVE SOUTH                                      07
111300506080000021601801FNH
    172LETHBRIDGE             2627VANTAS MEAT MARKET  3294545
904 7 AVE SOUTH                                       01
101421303010000001600061LNH
    173LETHBRIDGE             2627RILEY + MCCORMICK   3285644
CENTRE VILLAGE MALL                                   99
101300501030000001300051LNH
    174LETHBRIDGE             2627CONLAN CONTRACTING  3286403
207 12 AVE NORTH                                      99
101421316030000006000200 1FNH
    175LETHBRIDGE             2627GENTLEMEN 3 MENSWEAR3272232
314 7 STREET SOUTH                                    03
101300501030000000800031LNH
    176LETHBRIDGE             2627MCGUIRES MENSWEAR   3273761
310 5 STREET SOUTH                                    03
101300501030000001000041LNH
    177LETHBRIDGE             2627BOND STREET MEN WEAR3293411
COLLEGE MALL                                          03
101300501030000000800031LNH
    178LETHBRIDGE             2627BLACK*S MENS SHOP   3273551
619 4 AVE SOUTH                                       03
101300501030000001300051LNH
    179LETHBRIDGE             2627BLACK SHEEP BOUTIQUE3290511
LETHBRIDGE CENTRE                                     03
101421301030000001300051LNH
    180LETHBRIDGE             2627SWEET 16            3282166
COLLEGE SHOPPING MALL + 113 LETHBRIDGE CENTR          03
101300501030000002600011LNH
    181LETHBRIDGE             2627RICKI*S LADIES WEAR 3293272
67-400 4 AVE                                          03
101300501030000005200101LNH
    182LETHBRIDGE             2627SALLY SHOP          3280770
CENTRE VILLAGE MALL                                   03
101300501030000000800031LNH
    183LETHBRIDGE             2627REITMAN*S LTD       3282653
506 4 AVE SOUTH                                       03
101300501030000000800031LNH
    184LETHBRIDGE             2627PRIMROSE SHOP LTD   3272244
613 4 AVE SOUTH                                       03
101300501030000000800028LNH
    185LETHBRIDGE             2627PENNINGTON*S        3270677
LETHBRIDGE CENTRE                                     03
101300501030000001200051LNH
```

DATA COMPILED ON TAPE FROM COMMODITY FLOW SURVEY FORM

APPENDIX - II

APPENDIX - II-

INSTRUCTIONS AND LISTING OF COMPUTER PROGRAMS

COMPUTER PROGRAM - I: The purpose of this program is to algorithmically calibrate a singly-constrained gravity model of the form;

$$(T_{ij}) = P_i \frac{C_j \, d_{ij}^{\lambda}}{\Sigma_j \, C_j \, d_{ij}^{\lambda}} \qquad 1 \leq \quad < n$$

Where,

T_{ij} - Estimated interchanges between origin and destination,tons;

P_i - Production at origin, tons;

C_j - Consumption at Destination, tons;

d_{ij} - Distance between origin and destination;

λ - Power parameter; and

- Optimum condition for λ.

DESCRIPTION:

Computer program I is written in Fortran Language. The program can be modified to handle flow matrices of dimension i X j (i = number of origins and j = number of destinations) depending on the capacity of the computer utilized.

Provided the following inputs are given. Some of the instructions are included in the program itself.

The inputs for this program consists of:

1. P_i - Productions at each origin , tons.

2. C_j - Cosumption at each destination, tons.

3. d_{ij} - Distances between origin and destinations, kilometers.

4. D_{ij} - Observed interchanges, tons. and

5. λ – power parameter value.

The outputs from this program are written on a file called Tape5 in the form of tables. The output consists mainly of the friction factors, estimated tons, difference between estimated and observed tons, new production and consumption values, and other required data. The observed and estimated tons are also written on a seperate file called Tape6 for LSM to determine the optimum coefficient of determination (R^2) and correlation coefficient r values for a specific λ value.

```
      PROGRAM NANDU(INPUT,OUTPUT,TAPE206,TAPE207,TAPE208,
     *TAPE209,TAPE5,TAPE6)
C
      PARAMETER(N=5,D=36,M=12)
C
      DIMENSION P(5),C(36),OT(5,36),D(5,36),F(5,36),SUMI1(5),SUMI2(5)
      DIMENSION CF(5,36),SI(5),SJ(36),SCF(36),R(5,36),SUMJ1(36)
      DIMENSION TT(5,36),ET(5,36),DT(5,36),SUMJ2(36)
      DIMENSION SI1(4),SI2(4)
C
C THE FOLLOWING PROGRAM IS WRITTEN TO OBTAIN ESTIMATED TONS USING
C GRAVITY MODEL OF THE TYPE DESCRIBED BELOW; ESTIMATED TRIPS ET(I,J)
C IS EQUAL TO P(I) MULTIPLED BY (C(J)*(D(I,J)** LAMBDA)) DIVIDED BY
C SIGMA OF (C(J)*(D(I,J)**LAMBDA)).
C
C THE FOLLOWING BRIEFLY DESCRIBES THE SYMBOLS USED IN THIS PROGRAM:
C
C P(I) = PRODUCTIONS AT ORIGN, C(J) = CONSUMPTION AT DESTINATION,
C OT(I,J) = OBSRVED TONS OF O-D MATRIX. ET(I,J) = ESTIMATED TONS,
C D(I,J) = DISTANCE BETWEEN ORIGIN AND DESTINATIONS, F(I,J) = FRICTION
C FACTOR GIVEN BY D(I,J) ** LAMBDA, CF(I,J) = C(J) * F(I,J) THE NUMERATOR,
C R(I,J) = CF(I,J) DIVIDED BY (SIGMA OF CF(I,J), ET(I,J) = P(I)*R(I,J),
C DT(I,J)= OBJECTIVE FUNCTION VALUE, GIVEN BY (ET(I,J) - OT(I,J) SQUARE.
C ALL THE ABOVE VALUES ARE PRINTED IN SERIES.
C
C ****************READ*STATEMENTS .... DATA INPUT ***************
C
C       ******** READ PRODUCTION P(I)'S ***********
C
      READ(206,*) (P(I),I=1,N)
C
C *********** READ CONSUMPTION C(J)'S **********
C
      READ(207,*)(C(J),J=1,D)
C
C     *********** READ OT(I,J),*ORIGINAL*TRIPS *********
C
      J1=1
      J2=M
      DO 10 K=1,3
      DO 20 I=1,N
      READ(208,*) (OT(I,J),J=J1,J2)
20    CONTINUE
      J1=J1+M
      J2=J2+M
10    CONTINUE
C
C     ****** READ D(I,J) ***********
C
      J1=1
      J2=M
      DO 50 K=1,3
      DO 60 I=1,N
      READ(209,*) (D(I,J),J=J1,J2)
60    CONTINUE
      J1=J1+M
      J2=J2+M
50    CONTINUE
C
C END OF DATA INPUT (KNOWN)AND READ STATEMENTS
```

COMPUTER PROGRAM I

CONTINUED;

```
C
C CALCULATE ***************F(I,J)'S **************
C
      DO 53 I=1,N
      DO 53 J=1,O
      F(I,J)=((D(I,J)**(-0.75)))
C
C ******* THE FOLLOWING EQ. FOR LOG D(I,J) *******
C
C      F(I,J)=((LOG(D(I,J))))
C
53    CONTINUE
C
C ********* BEGINNING*OF*PRINT*STATEMENTS -I*********
C
         WRITE(5,9990)
9990     FORMAT(2X,/////////////)
C
        WRITE(5,8798)
8798     FORMAT(30X,'LAMBDA=-0.75,CC=FREIGHT,M=1,200SERIES',)
C
         WRITE(5,9999)
9999     FORMAT(30X,'***************************************',//)
      WRITE(5,103)
103   FORMAT(30X,'*****FRICTION FACTORS,F(I,J)S*****',////)
C
        J1=1
        J2=M
        DO 400 K=1,3
      DO 22 I=1,N
       WRITE(5,14)(F(I,J),J=J1,J2)
14     FORMAT(5X,12F10.2,//)
22     CONTINUE
          J1=J1+M
          J2=J2+M
400       CONTINUE
C
C ***********END*OF*DATA SET UP FOR PROGRAM **************
C
C      ************ C(J)*F(I,J) ************
C
       DO 56 I=1,N
       DO 56 J=1,O
       CF(I,J)=(F(I,J)*(C(J)))
56     CONTINUE
C
       WRITE(5,8992)
9992   FORMAT(2X,////////////)
C
       WRITE(5,104)
 104   FORMAT(30X,'*****C(J)XF(I,J)=CF(I,J)S******',//)
C
        J1=1
        J2=M
        DO 402 K=1,3
      DO 23 I=1,N
        WRITE(5,15)(CF(I,J),J=J1,J2)
 15    FORMAT(5X,12F10.5,//)
 23    CONTINUE
          J1=J1+M
```

CONTINUED;

```
         J2=J2+M
 402     CONTINUE
C
C ******************CALCULATE SIGMA C(J)*F(I,J)=SCF(J),ROW*SUM*******
C
        SCF(I)=0
        DO 58 I=1,N
        DO 58 J=1,D
        SCF(I)=(SCF(I)+CF(I,J))
58      CONTINUE
C
        WRITE(5,9993)
9993    FORMAT(2X,////)
C
        WRITE(5,999)
999      FORMAT(30X,'*****SIGMA*OF,CF(I,J)=SCF(I,J),*****',////)
C
        DO 4441 I=1,N
        WRITE(5,77)(SCF(I))
77      FORMAT(35X,F15.4,//)
4441     CONTINUE
C
C ***************CALCULATE R(I,J)=C*F/SIGMA C*F ****************
C
        DO 666 I=1,N
        DO 666 J=1,D
        R(I,J)=(CF(I,J)/SCF(I))
666     CONTINUE
C
        WRITE(5,909)
909      FORMAT(30X,'*****CF(I,J)/SCF(I,J)=R(I,J)S,*****',////)
C
          J1=1
          J2=M
          DO 404 K=1,3
       DO 25 I=1,N
        WRITE(5,17)(R(I,J),J=J1,J2)
17      FORMAT(5X,12F10.5,//)
25         CONTINUE
           J1=J1+M
           J2=J2+M
404        CONTINUE
C
C ************CALCULATE ET(I,J), ESTIMATED TONS********************
C
        DO 144 I=1,N
        DO 144 J=1,D
        ET(I,J)=(P(I)*(R(I,J)))
144     CONTINUE
C
66     FORMAT(25X,'*****EST. MINUS ORI. TONS*****',////)
C
        DO 9878 I=1,N
        DO 9878 J=1,D
        TT(I,J)=((ET(I,J))-(OT(I,J)))
9878    CONTINUE
C
        DO 2343 I=1,N
        DO 2343 J=1,D
        WRITE(6,2344)(OT(I,J)),(ET(I,J))
```

CONTINUED;

```
40        CONTINUE
          J1=J1+M
          J2=J2+M
118        CONTINUE
C
         WRITE(5,9995)
9995     FORMAT(2X,////)
C
          WRITE(5,8464)
8464        FORMAT(25X,'***** ESTIMATED TONS,ET(I,J) ******',////)
C
          J1=1
          J2=M
         DO 1325 K=1,3
          DO 1324 I=1,N
          WRITE(5,1889)(ET(I,J),J=J1,J2)
1889        FORMAT(5X,12F9.2,//)
1324        CONTINUE
          J1=J1+M
          J2=J2+M
1325         CONTINUE
C
          WRITE(5,9996)
9996      FORMAT(2X,////)
C
          WRITE(5,7876)
 7876     FORMAT(25X,'***** [ET(I,J) MINUS OT(I,J)]*****',////)
C
            J1=1
            J2=M
          DO 9821 K=1,3
           DO 9820 I=1,N
            WRITE(5,8373)(TT(I,J),J=J1,J2)
 8373       FORMAT(5X,12F10.2,//)
 9820       CONTINUE
           J1=J1+M
           J2=J2+M
 9821      CONTINUE
C
          WRITE(5,9997)
9997      FORMAT(2X,////)
C
          WRITE(5,8393)
 8393     FORMAT(25X,'****** SQUARE OF[ET(I,J) - DT(I,J)]*****',////)
C
            J1=1
            J2=M
         DO 7838 K=1,3
          DO 7837 I=1,N
          WRITE(5,3938)(DT(I,J),J=J1,J2)
 3938     FORMAT(5X,12F10.2;//)
 7837     CONTINUE
          J1=J1+M
          J2=J2+M
 7838      CONTINUE
C
C
          WRITE(5,9998)
9998      FORMAT(2X,////)
C
```

CONTINUED;

```
2344     FORMAT(F9.2,2X,F9.2,//)
2343     CONTINUE
C
C   ********** CALCULATE DT(I,J)= DIFF SQUARE **************
C
          DO 271 I=1,N
          DO 271 J=1,O
          DT(I,J)=((TT(I,J)**(2)))
271       CONTINUE
C
C COLUMN AND ROW TOTAL OF ESTIMATED TRIP MATRIX
C
          SUMI2(I)=0
           DO 987 I=1,N
           DO 987 J=1,O
           SUMI2(I)=(SUMI2(I)+ET(I,J))
987        CONTINUE
C
C ********** CALCULATION OF COL.TOTAL =EST.C(J) ****
           DO 487 J=1,O
           SUMJ1(J)=0
           SUMJ2(J)=0
           DO 487 I=1,N
           SUMJ1(J)=(SUMJ1(J)+DT(I,J))
            SUMJ2(J)=(SUMJ2(J)+ET(I,J))
487       CONTINUE
C
C ******** TO CALCULATE SIGMA OF DT(I,J),ET(I,J) AND DT(I,J) ***
C
        SUMS=0
        SUMS1=0
        SUMS2=0
        DO 500 I=1,N
        SI(I)=0
        SI1(I)=0
        SI2(I)=0
        DO 600 J=1,O
        SI1(I)=(SI1(I)+DT(I,J))
        SI2(I)=(SI2(I)+ET(I,J))
        SI(I)=(SI(I)+DT(I,J))
 600    CONTINUE
        SUMS=SUMS+SI(I)
         SUMS1=SUMS1+SI1(I)
         SUMS2=SUMS2+SI2(I)
500     CONTINUE
C
C ******** BEGINNING OF PRINT STATEMENTS II *************
C
        WRITE(5,9994)
9994    FORMAT(2X,////)
C
         WRITE(5,101)
101      FORMAT(30X,'***** OBSERVED TONS ,DT(I,J) ******',///)
C
         J1=1
         J2=M
        DO 118 K=1,3
         DO 40 I=1,N
         WRITE(5,99)(DT(I,J),J=J1,J2)
 99      FORMAT(20X,12F5.0,//)
```

CONTINUED;

```
      WRITE(5,177)
 177  FORMAT(25X,'****** DISTANCES D(I,J) ******',////)
C
      J1=1
      J2=M
     DO 872 K=1,3
      DO 70 I=1,N
      WRITE(5,80)(D(I,J),J=J1,J2)
 80   FORMAT(20X,12F5.0,//)
 70   CONTINUE
      J1=J1+M
      J2=J2+M
 872   CONTINUE
C
       WRITE(5,9910)
9910   FORMAT(2X,////)
C
      WRITE(5,988)
988   FORMAT(20X,'***** ORIGINAL AND ESTIMATED P(I)*****',////)
C
      WRITE(5,989)
989   FORMAT(25X,'    P(I)   ',5X,'    EP(I)    ',//)
C
      DO 454 I=1,N
      WRITE(5,645)(P(I)),(SUMI2(I))
645     FORMAT(25X,F9.2,5X,F9.2,//)
454   CONTINUE
C
      WRITE(5,967)
967   FORMAT(25X,'******* ORIGINAL AND ESTIMATED C(J) ***',//)
C
      WRITE(5,968)
968   FORMAT(25X,'    C(J)    ',5X,'    EC(J)    ',//)
C
      DO 333 J=1,O
      WRITE(5,203) (SUMJ1(J)),(SUMJ2(J))
203   FORMAT(25X,F9.2,5X,F9.2,//)
333   CONTINUE
C
    WRITE(5,501)
501 FORMAT(25X,' SIG.DT(I,J),SSD VALUE',////)
C
    DO 504 I=1,1
    WRITE(5,502) SUMS
502 FORMAT(25X,F15.2,//)
504   CONTINUE
C ******** END OF RUN **********
     STOP
     END
```

SPSS

COMPUTER PROGRAM - II: The purpose of this program is to anlayse the vast commodity flow data using a sub-program called CROSS TABS from the SPSS (Statistical Package for Social Sciences). This program is used for modal split analysis of the commodity movements across Alberta.

DESCRIPTION:

The following gives general instructions of the basic inputs required for using the cross-tabs, various other options are available under cross-tabs.

RUN NAME	-	Title of the project;
VARIABLE LIST	-	List variables by name;
NUMBER OF CASES	-	Numbe of cases;
INPUT FORMAT	-	Specify format for reading the listed variables;
MISSING VALUES	-	If any; and
VALUE LABELS	-	This is list of names to the variables according to their codes in the data file.

The output for this program is obtained by specifying the type of Cross-Tabs requiired depnding on the purpose of the researcher. The output consists of tables with percentage of distribution of specified variables.

```
RUN NAME          CROSSTABS
VARIABLE LIST     PLACE1,FIRM,GOODS,PLACE2,MAJOR,MINOR,TONS,SIZE,
                  ,MODES,LOADS,CONTROL,HIRE,MARKET
N OF CASES        309
INPUT FORMAT      FIXED(F4.0,F2.0,F2.0,X,F4.0,F2.0,F2.0,F7.0,X,F3.0,X,
                  F1.0,F1.0,F1.0,F1.0,F3.0)

ACCORDING TO YOUR INPUT FORMAT, VARIABLES ARE TO BE READ AS FOLLOWS

VARIABLE  FORMAT  RECORD   COLUMNS.

PLACE1    F 4. 0     1      1-   4
FIRM      F 2. 0     1      5-   6
GOODS     F 2. 0     1      7-   8
PLACE2    F 4. 0     1     10-  13
MAJOR     F 2. 0     1     14-  15
MINOR     F 2. 0     1     16-  17
TONS      F 7. 0     1     18-  24
SIZE      F 3. 0     1     26-  28
MODES     F 1. 0     1     30-  30
LOADS     F 1. 0     1     31-  31
CONTROL   F 1. 0     1     32-  32
HIRE      F 1. 0     1     33-  33
MARKET    F 3. 0     1     34-  36

THE INPUT FORMAT PROVIDES FOR    13 VARIABLES.    13 WILL BE READ.
IT PROVIDES FOR  1 RECORDS ("CARDS") PER CASE.
A MAXIMUM OF     36 "COLUMNS" ARE USED ON A RECORD.

MISSING VALUES ALL(BLANKS)
VALUE LABELS      MAJOR(1)GEN FREIGHT(2)FOOD NONPER(3)FOOD PER
                  (4)HEAVY MAH(5)METAL PRO(6)PETRO PRO(7)BULK LIQ
                  (8)BULK DRY(9)FOREST PRO(10)LIVE STOCK(11)
                  CONSTRU MAT(12)FEED PRO(13)TRAV-MOBI HOMES(14)
                  HOUSE GOODS(15)MAIL(16)OTHER(17)BAGGED PRO
                  (18)UNKNOWN(19)COAL(20)IN-SITU(21)PETRO&FERT
                 (22)PULP(23)OTHER&GOODS/MINOR(1)ONE(2)TWO(3)THREE(4)FOUR
                  (5)FIVE(6)SIX(7)SEVEN(8)EIGHT(9)NINE(10)TEN
                  (11)ELEVEN(12)TWELVE/MODES(1)TRUCK(2)RAIL
                  (3)PIPE(4)AIR(5)MARINE(6)BUS(7)WINTER ROAD
                  (8)OTHER/LOADS(1)FULL(2)LTFL/CONTROL(1)YES
```

COMPUTER PROGRAM II

BMDP

COMPUTER PROGRAM - III: The purpose of this program is to anlayse the vast commodity flow data using a sub-program called BMDP-4F from the BMDP (Bio Medical Computer Programs). This program was used for the analysis and development of modal split models(log-linear and logit models) of the commodity movements across Alberta.

DESCRIPTION:

The following gives general instructions of the basic inputs required for using the BMDP-4F, various other options are available under this package.

PROBLEM	- Title of the project;
INPUT	- List variables by name (if cell frequencies are used it should be named as a varaible), name of data file,format for reading, and number of cases;
CATEGORY	- Provide labels for each variable to be analyzed;
TABLE	- Gives the multiway table desired;
FIT	- Used to specify the model(s) to be investigated. The word model requires the symbols for the terms to be included in the model.
PRINT	- Provides for additional out put such as expected values, standardized residuals, lambda (required for logit model), etc.

The output from this program consists of, degress of freedom, G^2 statistic (for likelihood ratio test), logits(W terms) and standardized residual values. Also, it gives some other useful statistics necessary for log-linear and logit analysis. The important part of using this program is to make sure the inputs are in the required format as specified in the program listing.

```
TWO-WAY FREQUENCY TABLES -- MEASURES OF ASSOCIATION
MULTIWAY FREQUENCY TABLES -- LOGLINEAR MODELS (INCLUDING STRUCTURAL ZEROS)

COPYRIGHT (C) REGENTS OF UNIVERSITY OF CALIFORNIA.

  BMDP STATISTICAL SOFTWARE, INC.
  1964 WESTWOOD BLVD.  SUITE 202        PHONE (213) 475-5700
  LOS ANGELES, CALIFORNIA  90025        TELEX 4992203

PROGRAM VERSION,  APRIL, 1985.

THIS VERSION OF BMDP-85 HAS BEEN CONVERTED FOR
CDC NOS AND NOS/BE OPERATING SYSTEMS
BY BMDP PROJECT, VOGELBACK COMPUTING CENTER
NORTHWESTERN UNIVERSITY
2129 SHERIDAN ROAD
EVANSTON, ILLINOIS 60201
(312) 491-2197

CDC RELEASE DATE, JUNE 1986
MANUAL EDITION, 1983, 1985 REPRINT.  STATE NEWS IN THE PRINT
  PARAGRAPH FOR A SUMMARY OF NEW FEATURES.
EXECUTED ON  87/11/07 AT 11.26.18.

PROGRAM CONTROL INFORMATION

/PROBLEM           TITLE IS 'MODE PREFERENCE BY DISTANCE'.
/INPUT             VARIABLES ARE 4.
                   FILE=DATA2.
                   FORMAT IS '(F2.0,1X,F1.0,1X,F1.0,1X,F3.0)'.
                   CASES ARE 48.
/VARIABLE          NAMES ARE DIST,COMM,PREF,FREQ.
/CATEGORY          NAMES(1) ARE A,B,C,D.
                   CODES(1) ARE 1,2,3,4.
                   NAMES(2) ARE GEN,FOOD,FOREST,PETRO,BULK,HEAV.
                   CODES(2) ARE 1,2,3,4,5,6.
                   NAMES(3) ARE TRUCKS,RAIL.
                   CODES(3) ARE 1,2.
/TABLE             INDICES=DIST,COMM,PREF.
                   SYMBOLS=D,K,M.
                   COUNT IS FREQ.
/FIT               ALL.
                      MODEL IS DK,DM.
                   ADD=SIMPLE.
                   DELETE=SIMPLE.
/PRINT             EXPECTED.
                   STANDARDIZED.
                   LAMBDA.
/                  END.
```

COMPUTER PROGRAM III