



ADDIS ABABA UNIVERSITY
ADDIS ABABA INSTITUTE OF TECHNOLOGY
SCHOOL OF MECHANICAL AND INDUSTRIAL ENGINEERING
GRADUATE PROGRAM IN RAILWAY ENGINEERING

ADOPTING FREIGHT DEMAND FORECASTING MODEL FOR ADDIS ABABA-DJIBOUTI RAILWAY

By: AREFEAYNE AMARE

Advisor: Dr. Birhanu Beshah (Ass. Prof.)

And

Co. Advisor: Mr. Biniyam Ayalew

A Thesis submitted to the School of Graduate Studies of Addis Ababa University
in partial fulfillment of the requirements for the Degree of Masters of Science in
Mechanical Engineering **(Railway stream)**

October 2016

ADDIS ABABA UNIVERSITY
ADDIS ABABA INSTITUTE OF TECHNOLOGY
SCHOOL OF MECHANICAL AND INDUSTRIAL ENGINEERING
RAILWAY ENGINEERING
ADOPTING FREIGHT DEMAND FORECASTING MODEL FOR
ADDIS ABABA-DJIBOUTI RAILWAY

By: AREFEAYNE AMARE

APPROVED BY BOARD OF EXAMINERS

Approved by:

<u>Dr. Birhanu Beshah (Ass. prof)</u>	_____	_____
Advisor	Signature	Date
<u>Mr. Biniyam Ayalew</u>	_____	_____
Co advisor	Signature	Date
<u>Dr. Ir. Eshetie Birhan</u>	_____	_____
Internal examiner	Signature	Date
<u>Dr. Ir. Kassahun Yimmer</u>	_____	_____
External examiner	Signature	Date
<u>Dr. Birhanu Beshah (Ass. prof.)</u>	_____	_____
Railway Engineering Center Head	Signature	Date
_____	_____	_____
Director for Postgraduate Program	Signature	Date

Declaration

I hereby declare that the work which is being presented in this thesis entitled “Adopting Freight Demand Forecasting Model for Addis Ababa-Djibouti Railway” is original work of my own, has not been presented for a degree in any other university. Many assumptions and considerations has been taken by myself so that the work of the analysis truly correlate the data that has been gathered.

Arefeayne Amare

(Student)

Signature

Date

Dr. Birhanu Beshah (Ass. Prof.)

(Advisor)

Signature

Date

Mr. Biniyam Ayalew

(Co. Advisor)

Signature

Date

ACKNOWLEDGEMENT

First of all I would like to thank Merciful God for all things he makes an offer along my path, and also for my family, without all their great support I would have not been complete this thesis after my first degree.

My deepest gratitude goes to my Advisor Dr. Birhanu Beshah (Associate Professor) for his support. His office was always been open when I am in need of his help; not only during my thesis work but also on course class which led me a great interest to have this title. In addition to Dr. Birhanu Beshah I also need to thank my Co-Advisor Mr Binyam Ayalew for his valuable idea when I am in need.

The last but not the least, I also love to give special thanks to staffs from different organizations. That is Niyat from Ethiopian Commodity exchange, Emawayish from Ministry of Finance and Economic Development, Girma from Ethiopian shipping and logistic enterprise and staffs of Ethiopian transport Authority and Ethiopian customs and revenue authority. Thank you all

ABSTRACT

Freight transportation is an essential component of any economic activity which makes a continuous change as a result of growth but many economic activities depend on traffic congestion and truck travel time along the origins and destinations which makes series problem mainly for land locked countries like Ethiopia. This is because government and organization pay undesired tax for the stay of goods at the port. In order to solve this problem different countries use different forecasting models which are designed to predict the future need of the country.

This thesis work identifies major freight forecasting projection models, discusses those models, chooses the better model based on the merits and demerits of each model and forecasts with one of the classic freight forecasting model which is called four-stage aggregate model (hybrid with its own parameters) to show how it works in the context of Ethiopia by taking traffic congestion as a main problem.

As part of this work trip generated using growth factor and linear regression and mode choice along the path has been done by incorporating the truck traffic count, commodity tonnage share and an expense given to each item. This hybrid four-stage aggregate model which begins with trip generation step and gives a result of 3,351,406 trips when working with Annual Growth Factor (AGF) and 57,974 trips with linear regression method. And a comparison between this a model and a model forecasted by ERC has been compared and found ERC model which generates 119,419 trips nearly comparable with linear regression method even though the parameters used here are different. From the total number of trips generated trains took 79 percent mode share and truck took remaining modal share. The distribution and an assignment step is not done because, the demand investigating here focuses only origin and destination.

Finally the thesis find out forecasting with four-stage aggregate model(Hybrid model) which uses traffic counts, commodity tonnage share and an expense growth rate given for each items as input parameter is better model in order to generate number of trips along the line. This is because import and export tonnage rates can easily affected by the influence of economy, industrial location patterns, globalization of business, fuel prices, environmental factors and others which this all cumulatively can affect the traffic congestion.

Table of Contents

ACKNOWLEDGEMENT	i
ABSTRACT.....	ii
LIST OF TABLES.....	vi
LIST OF ACRONYMS	viii
CHAPTER ONE	1
INTRODUCTION	1
1.1. BACK GROUND.....	1
1.2. STATEMENT OF THE PROBLEM	3
1.3. RESEARCH QUESTION	3
1.4. OBJECTIVE OF THE STUDY	4
1.5. SCOPE OF THE STUDY	4
1.6. SIGNIFICANCE OF THE STUDY	4
1.7. LIMITATION OF THE STUDY	5
CHAPTER TWO	6
LITERATURE REVIEW	6
2.1. FORECASTING	6
2.2. FREIGHT DEMAND FORECASTING MODELS	6
2.3. FACTORS AFFECTING FREIGHT DEMAND	10
2.3.1. FACTORS AFFECTING FREIGHT DEMAND DIRECTLY	10
2.3.2. FACTORS THAT AFFECT DEMAND INDIRECTLY	14
2.4. MODELLING AND DECISION MAKING	15

2.5. MODEL DISCUSSION	18
2.5.1. TIME SERIES MODEL	18
2.5.2. ACTIVITY-BASED MODEL	30
2.5.3. FOUR-STAGE AGGREGATE MODEL	34
CHAPTER THREE	43
METHODOLOGY	43
3.1. INTRODUCTION	43
3.2. METHOD DESIGN	43
3.3. DESCRIPTION ABOUT CASE STUDY AREA	46
3.4. DATA COLLECTION	47
3.5. ASSUMPTIONS AND CONSIDERATIONS OF THE DATA	52
CHAPTER FOUR	53
DATA ANALYSIS AND PRESENTATION	53
4.1. DATA ANALYSIS	53
4.2. COMPARISON OF THE GROWTH FACTOR MODEL AND REGRESSION MODEL	64
4.3. COMPARISON OF RESULT WITH THE MODEL FOECASTED BY ERC	67
CHAPTER FIVE	69
CONCLUSION AND RECOMMENDATION	69
5.1. CONCLUSION	69
5.2. RECOMMENDATION	70
5.3. FUTURE WORK	71

Bibliography	72
Appendix 1 Summary of trade volume in tons: (Ethiopia Commodity Exchange).2016	74
Appendix 2 Total cargo lifted by shipping sector by trade type and vessel type: (Ethiopian Shipping and Logistic Enterprise.2015	75
Appendix 3. Estimate of transit and transport costa of import and export percent of import and export value: the management of commercial road transport in Ethiopia (ERA). 2014	76
Appendix 4 Number of Vehicles along Addis Ababa-Djibouti path with their life period from the day in use and carrying tonnage and also level given by the authority: Ethiopian Transport Authority.2016	77
Appendix 5. GDP by Economic activity at constant prices and growth rate of GDP by Economic activity at constant prices; Ministry of Finance and Economic Development.2015	78

LIST OF TABLES

Table 1 Freight Model Classes by Component (re-created from NCHRP 606, 2008).....	9
Table 2 Kay's Idea which identifies additional contrast between the two basic approaches:	17
Table 3 Alternative forecasting approaches (Llvento, Titus Awokuse and Tom, 2004).....	19
Table 4 Recent Historical data from traffic counts report from Addis Ababa-Djibouti	48
Table 5 Value of export by major commodity groups measured in million birr	49
Table 6 Volume of major export by type of commodity in million ton (MT).....	50
Table 7 Value of import by type of commodity groups (in billion birr).....	51
Table 8 Volume of major import by commodity group in million ton (MT)	51
Table 9 Calculated Share of Exported commodity groups in 2007 E.C on the basis of 2002 E.C	54
Table 10 Calculated Annual Growth on the basis of exported economic variable given in million birr taken from table 5.....	54
Table 11 Expected or forecasted truck traffic counts after 15 years relating historical projection and Economic projection for exported goods	55
Table 12 Export forecasted for major commodity groups with truck traffic counts.	56
Table 13 Calculated share of Imported commodity groups in 2007 E.C on the basis of 2002 E.C data set	57
Table 14 calculated annual growth on the bases of imported economic variable given in billion birr taken from table 7.	58
Table 15 Expected or forecasted truck traffic counts after 15 years relating historical and economic projection for imported goods.....	59
Table 16 Import Forecasted for major commodity groups with truck traffic	60
Table 17 Summary of imported and exported commodity groups tonnage on the basis of the recent data and calculating for 2007 E.C tonnage	61
Table 18 Forecasted import and export for major commodity groups	62

Table 19 regression table showing commodity share in million ton with respect to annual truck traffic flow for imports and exports (2007 E.C)	63
Table 20 Expected annual forecasted trips by regression analysis	64
Table 21 Multinomial Trip Characteristics	67
Table 22 Total potential freight flows (expressed in ton).....	67

LIST OF ACRONYMS

AGF	Annual Growth Factor
CSA	Central Statistics Agency
ECRA	Ethiopia Customs and Revenue Authority
ERA	Ethiopian Roads Authority
ERC	Ethiopian Railway Corporation
ESLE	Ethiopian Shipping and Logistic Enterprise
ETA	Ethiopia Transport Authority
FAME	Freight Activity-Based Modeling
FHWA	Federal Highway Administration
FNEM	Freight Network Equilibrium Model
FTD	Freight Transportation Demand
GDP	Gross Domestic Product
GNP	Gross National Product
JIT	Just-in-Time Inventory system
LTL	Load-truck-Load
MAD	Mean Absolute Deviation
MAPE	Mean Absolute Percentage Error
MoFed	Ministry of Finance and Economic Development
MNL	Multinomial Logit
MSE	Mean Squared Error
NCHRP	National Cooperative Highway Research Program

NRNE	National Railway Network of Ethiopia
OD	Origin Destination
TAZ	Traffic Analysis Zone
VMT	Volume in Million Ton

CHAPTER ONE

INTRODUCTION

Freight demand modelling may play a particularly important role in developing countries where the efforts to increase exports and to gain access to underdeveloped areas are even more urgent. Facilitating the movements of goods in these cases is likely to have a major impact on economic development.

Freight demand forecasting models are models that are useful for forecasting purpose. Many countries use different forecasting techniques in order to get accurate measure of the figures and in order to make decisions on the demand and in order to allocate the better transport mode and also for an assigning purpose.

In this thesis the forecasting along the Addis Ababa- Djibouti rail line has been calculated by one of the classic four-stage aggregate model and hybrid that model taking the truck traffic count, commodity tonnage share and by considering an expense given to each good. This is done in order to eliminate the traffic congestion and truck travel time as a main problem.

The thesis work begin by giving general information about research conducted. Initially, the basis behind the research are discussed and presented then followed by reviewing different literatures, method design to forecast, data collection, data analysis and presentation finally give a conclusion and recommendation.

1.1. BACK GROUND

In a growing globalized context and consumption economic freight transport is of crucial importance. The need of economic activity, the prediction of more fluctuating economy and the season ability of different materials that can be transported makes a series problem of forecasting for freight transportation (White Legg.J., 1993).

History tells us even though Djibouti has a small trading population but the possession of plenty of good water and an excellent harbor, and had the additional advantage that French were believed not to have any ambition of expanding inland make Emperor Menelik to propose Ethio-Djibouti

rail line. Besides this fact different studies show the interest of Ethiopia was also importing and exporting of various kinds of goods for market of different countries with this line (Richard, Journal of Franco-Ethiopian Railway and Its History).



Figure 1 one of the locomotive employed on the Franco-Ethiopian Railway

(Richard, ([http://www.google.com/ethio-franco railway and its history/ image/html](http://www.google.com/ethio-franco%20railway%20and%20its%20history/image/html), 2008)

Due to the fact that the economic growth of Ethiopia is increasing from year to year the need of importing different construction machineries, machines to be erected for different production sectors, electronics, and many other commodities and also exporting coffee, chat, flower, animal products such as meat, finished and unfinished leather products and also different minerals has become great important this all international market is done via the Djibouti port. The total amount of commodity (measured in tons) carried by other modes of transport in any period depends up on the total output level of the commodity, structural changes (centralization and decentralization of production in relation to consumption) (Nigussie S. s., 2004).

All modes of domestic freight transportation are expected to experience significant growth in the coming decades; truck's share-when measured in tons and ton-miles- is expected to grow at the expense of rail freight coming the port. This reflects changes in economy that is anticipated to favor the production and shipment of higher-value-added and time sensitive goods (Seada, 2016).

Due to change of economy the freight demand will then influence the use of an appropriate planning of the models that should be used to see and come up with the best forecasting approach; so studying about different forecasting models in case of Addis Ababa-Djibouti rail line and adopting it will benefit different organizations, the society hence the country but the major freight forecasting models used and studied by different countries depend on their economy and other freight demand factors therefore seeing and assessing different models and getting countries experience about using models to Ethiopia will be helpful for developing and finding one new model for the future.

1.2. STATEMENT OF THE PROBLEM

Even though freight transportation plays a vital role in an economic growth of a country, the great challenge of not only Ethiopia but also many other countries is that commodity exchange with sea port and in land depends on the goods travel time between origin and destination of delivery and also shipment size of goods, due to this problem companies paid undesired tax for the stay of goods in the ports which may affect customer satisfaction (Nigussie S. S., November 10, 2004)

Here in Ethiopia many commodities like General cargo (including machineries, vehicles, electronics and others), petroleum, fertilizer and food aid are imported and commodities like coffee, pulses, oilseeds, meat or hides, fruits and vegetables and other goods are exported through Djibouti port and may stay for long about months due to not only /buyer/shipper/seller but mainly of traffic congestion along the line.

In order to illuminate this problem a study on freight demand forecasting model has been made by different countries in order to forecast and plan for the future so this paper is initially addressed to see different models and able to identify which will be helpful forecasting approach.

1.3. RESEARCH QUESTION

- Which freight demand forecasting model is appropriate for estimating the freight demand along Addis Ababa- Djibouti rail line?
- How the selected model works in the context of Ethiopia by taking traffic congestion as a main problem?

Basically an assumption has been made on the basis of that all the parameters in the selected model is easily influenced by many factors like that of an economic, an activity and other environmental factor which all can determine the number of traffic counts along the line.

1.4. OBJECTIVE OF THE STUDY

General Objective

The main objective of this paper is to investigate the different freight forecasting models and come up with the better model by telling their strengths and weakness and able to forecast using one of the model for Addis Ababa – Djibouti rail line.

Specific Objective

- To Estimate freight demand along the Addis Ababa-Djibouti corridor.
- To adopt the four-stage aggregate model working with traffic count together with economic variable given for commodity groups of imported and exported items.

1.5. SCOPE OF THE STUDY

There are many modelling approaches that different countries use accordingly, this paper tries to tell there are different kinds of freight forecasting model and show the models use has their own merits and demerits besides this the approaches used uses different variables in order to forecast freight demand.

The scope of this paper is to discuss those models, select the best model by comparing with models applicability, forecast with the selected model and finally give direction for those who are interested in studying for freight model.

1.6. SIGNIFICANCE OF THE STUDY

The applicability of the study depends of the nature of the problem in many cases. Freight transport is a series problem in many countries, here in Ethiopia that port in a series problem since the country is land locked studying the demand will be good. Studying this problem will then answer to have best mode choice if the data is expressive enough it can also able to allocate rail network among regions of the country which finally able to solve the traffic congestion problem along the path.

1.7. LIMITATION OF THE STUDY

- ✓ The availability of data, according to Ethiopian shipping and Logistic Enterprise, the data considered here by the organization only considers importers and exporters which uses Ethiopian ships but easily told by the organization there are some importers and exporters that use different countries 'ship.
- ✓ distribution has not be made because an assumption taken into account here only considers origins and destinations which is to mean that it doesn't consider journeys in between Addis Ababa and Djibouti.

CHAPTER TWO

LITERATURE REVIEW

2.1. FORECASTING

Forecasting is a process of estimating a future event by casting forward past data. The past data are systematically combined in a predetermined way to obtain the estimate of the future. Prediction is a process of estimating a future event based on subjective considerations other than just past data; these subjective considerations need not be combined in a predetermined way. Thus forecast is an estimate of future values of certain specified indicators relating to a decisional/planning situation, in some situations forecast regarding single indicator is sufficient, whereas, in some other situations forecast regarding several indicators is necessary. The number of indicators and the degree of detail required in the forecast depends on the intended use of the forecast (Gor.Ravi, 2009).

There are two basic reasons for the need for forecasting any field.

1. **Purpose** – Any action devised in the PRESENT to take care of some contingency accruing out of a situation or set of conditions set in future. These future conditions offer a purpose / target to be achieved so as to take advantage of or to minimize the impact of (if the foreseen conditions are adverse in nature) these future conditions.
2. **Time** – To prepare plan, to organize resources for its implementation, to implement; and complete the plan; all these need time as a resource. Some situations need very little time, some other situations need several years of time. Therefore, if future forecast is available in advance, appropriate actions can be planned and implemented ‘in time’ (Ilvento, Titus, & Tom, 2004).

2.2. FREIGHT DEMAND FORECASTING MODELS

Goods carried by rail, water and air are generally considered freight, while goods transported by truck may considered freight only if the truck in question carries goods that are also likely to be carried by other modes or not limited to local delivery.

In transportation planning, goods transported incidental to the primary purpose of trip, such as luggage accompanying an airline passenger on a business trip or tools accompanying a workman

on service call are generally not considered freight. It also excludes shipments of agricultural products from the farm to a processing center or terminal elevator most likely short distance or local distance.

Demand models are one of the key components of transportation planning at the strategic, tactical and operational levels. Government and private sectors needs forecasts of demand for transportation services to anticipate among others future financial requirements, equipment acquisition and labor requirements (Ortuzar & Willumsen, 2011).

Freight transportation is commonly measured and described by either commodity movements or vehicle movements. Commodity movements are typically represented by an origin-destination matrix that contains both the type and quantity of goods moved *vis-à-vis* vehicle movements, which are represented by traffic flows in different modes. Freight demand is derived from the socioeconomic system in which raw materials, intermediate inputs and finished products are needed at specific locations at specific times. Therefore the primary focus of freight transportation demand should be commodity movements because vehicle movements are triggered by the need to move commodities (Garrido, Amelia, & Regan, 2002).

A Model is a simplified representation of a part of the real world—the system of interest—which focuses on certain elements considered important from a particular point of view. Models are, therefore, problem and viewpoint specific.

Designing a model may fall into the physical model which the range spans from mental models in daily interaction of Human being with the real world also give typically analytical model which is the representation of some theory about the system of interest and how it work. These models attempt to replicate the system of interest and its behavior by means of mathematical equations based on certain theoretical statements about it. Although they are still simplified representations, these models may be very complex and often require large amounts of data to be used.

Published freight transportation studies reflect different modeling approaches. According to Harker (1985), freight models can be divided into three categories: econometric model, spatial price equilibrium model and network equilibrium model. A thorough review of econometric freight transportation demand (FTD) models can be found in Zlatoper and Austrian (1989).

A review of the spatial price equilibrium approach can be found in Friesz *et al.* (1985) and Harker and Friesz (1986a and b) while comprehensive survey of network equilibrium models for FTD can be found in Crainic (1987). The family of Operations Research models developed for strategic, tactical and operational freight transportation planning and decision making are carefully reviewed in Crainic and Laporte (1987). Strong, Harrison and Mahmassani (1996) divide freight mode split models into econometric models and network based models. Strong, et.al(1996)define econometric models as those that attempt to identify and analyze cause-and-effect and correlative relationships between freight demand and various factors and network models as those that apply and optimization rule to an objective function, governed by a system of equations with an appropriate set of data, in order to predict the distribution of freight traffic at some point in the future (Garrido, Amelia, & Regan, 2002).

Mazzarino (1997), in recent survey of literature, identifies two main groups; the macroeconomic models and the micro economic models. Similarly Winston (1983) classifies the studies as either aggregate or disaggregate based on the nature of the data used for estimation purposes. The aggregate models' basic unit of observation is an aggregate share of a freight mode at a certain geographical level. Disaggregate models consider the individual decision maker's choice of a freight mode for a given shipment as a basic unit of observation. Aggregate models tend to be either adhoc (empirical) or the solution to firm-based cost minimization problem, whereas disaggregate models focus more on behavioral aspects of the decision maker. Even though the disaggregate approach may seem theoretically more appealing than the aggregate one, the latter suffers from inherent drawbacks that makes it impractical in some applications. The main disadvantage is the requirement of extensive input data. It should be pointed out, however, that the primary difference between aggregate and disaggregate models is the nature of the required data rather than different behavioral approaches. Indeed, both can be derived from the same theory of optimal individual firm behavior (Romeo, Lucia, & Danielis, 1999).

The state wide freight forecasting released by the National Cooperative Highway Research Program (NCHRP 606, 2008) gives a comprehensive review of current models used in practice.

Table 1 Freight Model Classes by Component (re-created from NCHRP 606, 2008)

Model classes	Model component					
	Direct factoring	Trip Generation	Trip Distribution	Modal Split	Traffic Assignment	Economic/Land use modeling
Direct facility flow factoring method	Of facility flows					
OD factoring method	Of OD tables					
Truck Model		Based on exogenously supplied zonal activity				
Four step activity model		Based on exogenously supplied zonal activity				
Economic activity model		Based on outputs of economic model				

Yet another classification of freight transportation studies concerns the geographical scope under consideration. Studies can be divided into three broad categories: international, intercity, and urban. This classification is particularly useful because the decision makers involved in each category are quite different, as are the forces that generate freight movements.

In addition, the technologies involved in each category are substantially different. For these reasons, this review classifies research first on the basis of geographical scope. It should be noted that this classification of existing models may not be appropriate in the future as increasing globalization will certainly modify both the decisions made throughout the supply chain and the legal and managerial considerations regarding geographical jurisdiction (Ehsan, et al., 2016).

Haralambides and Veenstra (1998) identify three main approaches to model international demand for shipping. The first approach follows the standard theory of international trade which allows the direct inclusion of transportation costs. The standard theory is concerned with the pattern of trade between two or more nations, based on the Ricardian Principle of Competitive advantage. The theory lies on a simple comparison among autarkic relative prices in different countries which trade all available final goods but fully mobile within the boundaries of each domestic economy. (Garrido, Amelia, & Regan, 2002)

2.3. FACTORS AFFECTING FREIGHT DEMAND

There are a variety of factors that influence freight demand. These factors are presented in two groups. The first group consists of factors that affect demand relatively directly. The second group consists of factors whose direct effects are on the costs of one or more transport modes and, in some cases, on services offered; these factors affect demand indirectly as a result of changes in transport costs and rates and the services offered.

2.3.1. FACTORS AFFECTING FREIGHT DEMAND DIRECTLY

A. The Influence of Economy

The demand for freight transportation is commonly referred to as a ‘derived demand’ that is, it derives from a more basic demand- in this case a location-specific demand for a product that results in a need to ship the product to this location. As derived demand, the most basic influence on total freight is the volume of goods produced and consumed. Expansion in the national economy, or the

economy of any region, results in increases in overall freight demand, while economic contractions result in reductions in freight demand.

At the national level, the size of the economy is most frequently measured in dollar terms as gross national product (GNP) or gross domestic product (GDP).

However, freight demand is more closely related to the goods-production component of GNP or GDP. (Cambridge Systematics, Sep 1996)

B. Industrial location patterns

Just as the economy determines the amount of goods transported, the spatial distribution of economic activities determines the distances they are transported. Thus, industrial location patterns are an essential factor in determining transport demand when it's measured in ton-miles or in any similar units that reflect length of haul. This influence of spatial distribution can be best measured through its actual effect on demand: as average length of haul by commodity or total ton-miles transported. Ideally, such a measure would be applied in mode-neutral form as a great-circle miles, though reducing actual origin-to-destination distances to great-circle miles usually entails more effort than is warranted. The real GDP of goods is a reasonable overall measure of the influence of the economy on freight demand, it measures goods production in dollars rather than tons or volume.

The spatial distribution of economic activity is also a major influence on the modes that are used: Many commodities are likely to be shipped by one mode (e.g., truck) when distances are relatively short and by another modes (e.g., rail or air) when distances are longer. Water transport is competitive for many low valued commodities being shipped domestically between points at or near appropriate ports, but is rarely competitive for transport between points that are not located near ports. Plants located on rail lines are likely to use rail for an appreciably greater share of their transport needs than similar plants that are not so located.

C. Globalization of business

Measure of world trade include value and volume of imports, exports, and in-transit shipments, by foreign country (or region) of origin or destination.

Changing patterns of world trade not only affect transport flows, they affect modes used. Products that are received by truck from domestic suppliers may be obtained containerized and doublestacked from overseas suppliers or, if their value is relatively high or delivery speed important, by air freight.

Hence the world is becoming globalized products manufactured from different materials originated from many countries and become assembled to give one product (Cambridge Systematics, Sep 1996).

D. International Trade Agreements

Global patterns of production and distribution affected by countries' import restrictions and tariffs, those of our trading partners, and by international trade agreements. Quotas not only have the obvious effects on volumes of goods shipped internationally, but in case of natural resources on which quotas are frequently adjusted to reflect changing supply conditions, and they encourage the use of foreign distribution warehouses that provide capabilities for responding quickly to quota changes. Most countries exporting to other countries qualify for 'most favored nation' status, and additional agreements exist with many countries.

E. Just-in-time Inventory Practice

Just-in-time (JIT) systems, originated by the Japanese during the 1950s and 1960s, have been embraced by U.S. manufacturers at the rapid pace during the past decade. Industries in which the U.S. manufacturers have successfully adopted JIT systems include the metal products, automotive, electronics, food, and beverage industries.

JIT systems focus on keeping inventories at minimum levels through a coordination of input deliveries with production schedules. Adopting a JIT system results in increasing the frequency with which inbound shipments are scheduled, decreasing the lead times for these shipments and their size, and increasing the importance of receiving these shipments on time. Firms adopting JIT systems frequently reduce the number of suppliers and transport companies with which they deal, and they require suppliers that are close enough to be able to deliver shipments reliably within the constraints of short lead times.

The effects on freight demand are to increase the number of individual shipments, decrease their length of haul, and, most importantly, increase the importance of on-time delivery. Some shift may occur to modes that are faster or can handle smaller shipment sizes (from rail to truck load, truck load to LTL, or LTL to air freight or parcel). Within modes, a shift is likely to carriers that are capable of delivering highly reliable service, and, as the emphasis on reliability increases, the total number of carriers used generally falls. Total VMT of trucks may rise as a result of diversion from rail and reduced shipment sizes for truckload shipments, but these effects are likely to be partly balanced by reductions in lengths of haul and diversion to air.

Appropriate measures of the use of JIT systems are the number of companies or plants that consider themselves to be using such systems, the total value of the product of plants using these systems, and the total volume (tons) of inbound shipments to these companies. These measures are imperfect (in part, because there is substantial variation in the actual inventory practices of companies that identify themselves as using JIT systems), and they are difficult to quantify.

F. Centralized Warehousing

As transportation systems have become efficient and more reliable, there has been a trend toward using fewer warehouses for the distribution of products. Reducing the number of warehouses reduces inventory requirements but increases the lengths of haul for many shipments from warehouses. This trend is in part the result of increasing use by manufacturing firms of third-party logistics operators that specialize in optimizing the distribution process. The trend results in increasing transport demand and associated costs in order to achieve a larger saving in inventory costs.

The trend toward centralized warehousing results in increased transport demand (measured in ton-miles, shipment-miles, or value of service) and, in some instances, a shift from truck to air delivery. Appropriate measures of this trend are: the number of companies using one or two warehouses (or otherwise reducing the number of warehouses they use); and the value or volume (tons) of products shipped from these warehouses.

G. Packaging Materials

The age of plastics has brought with it the use of Styrofoam, bubble packs, and other very lightweight materials as protective packaging for many manufactured products. The result has been a reduction in the average density of shipments of these products. Since low-density shipments cause trucks to “cube-out” before they “weight-out,” the increase in relatively low-density shipments has created a demand for larger truck trailers.

Although some historic estimates of shipment density exist, there do not appear to be any data on how the shipment density of manufactured products has been changing (and even the historic estimates tend to focus primarily on the density of natural resource shipments- shipmentss that are usually are quite dense and whose density is likely to vary very little).

H. Recycling

Increasing use of recycled materials affects the origin/ distribution patterns, lengths of haul, and modal usage of several commodities.

Processing plants that use virgin materials are usually located near a major source of supply of these materials, and they commonly ship their products long distances to their markets.

Recycling plants, on the other hand, usually are located near the markets they serve, which also provide them with substantial volumes of material for recycling. Plants producing products from a combination of raw and recycled materials are likely to be located near sources of supply for their more important inputs and may receive some inputs by rail from more distant sources of supply.

2.3.2. FACTORS THAT AFFECT DEMAND INDIRECTLY

A. Fuel Prices

Fuel constitutes a moderately significant and relatively volatile component of costs for all freight modes. Fuel consumption and fuel costs are highest for air freight and generally are lower for the slower, lower quality-of-service modes. A significance increase in real fuel prices is likely to result

in greater rate increases for the faster modes than for the slower ones and some corresponding shift of demand across modes.

In evaluating the effect of fuel price changes on modal demand, it's necessary to consider fuel requirements for competing services rather than modal averages. Typically rail-competitive intercity truckload operators require less fuel per ton-mile than much other truck transport, while rail double stack and other intermodal services (which have relatively high tare weights, high speed, and poor aerodynamics) require more fuel than much other rail transport. Thus, a significant increase in fuel prices is likely to result in less diversion from truck to rail intermodal service than a simple comparison of overall fuel efficiency for truck and rail operations would suggest (Center for Urban Transportation, March 1999)

B. User Charges

Most publicly provided transportation infrastructure is funded primarily through user charges. The major exception is the inland waterway system.

The operations of coastal ports are financed by variety of user charges (per container or per ton of cargo), dockage charges, lease revenue, equipment rental fees, gate fees (for trucks and rail cars), and franchise fees (for stevedore firms and other vendors). Facility construction is financed primarily by a combination of revenue bonds, general obligation bonds, and federal aid. The inclusion of general obligation bonds in the mix suggests that some port facilities are not fully supported by user charges but may require some financial support from state or local governments (Center for Urban Transportation, March 1999).

2.4. MODELLING AND DECISION MAKING

Before choosing a modelling framework one needs to identify the general decision-making approach adopted in the country, government or decision unit.

Previous editions of this text have characterized decision-making styles following the ideas of Nutt (1981); in practice, no decision-making style fits any of these categories exactly (Juan de Dios Ortuzar, Lewis G-Willumsen, 2011).

The substantive rationality view of the world assumes that we know what our objectives are and we can envisage all alternative ways of achieving them and, with some luck, quantify the costs and benefits associated to each approach. Some of the problems of applying this approach includes:

- Difficulties in actually specifying what the objectives are beyond generalities like reducing congestion or improving accessibility;
- The accusation of insensitivity to the aspirations of the public; people do not actually care about ‘optimized’ systems, decision makers needs to see progress that is sustained along lines that are difficult to identify: asking for speed but dissatisfaction with the associated noise and emissions.
- Its high costs; substantive rationality is expensive to implement, requires advanced models and many runs for alternative arrangements and sensitivity analyses; efforts to apply this approach often overrun in time and budget; and
- The alienation of decision makers who may not understand, nor accept, the analytical treatment of the problem.

The main alternative approach to substantive rationality is what Lindblom (1959) called muddling through. The name, misleadingly self-deprecating, is not meant to imply that intuitive and unstructured decision making is desirable. On the contrary, in Lindblom’s eyes, muddling through is a disciplined process but not one based on the substantive rational handling of defined objectives.

The approach uses a combination of high-level (often unquantifiable) objectives, intermediate goals and immediate actions or experiments.

Muddling through, or what Kay calls ‘oblique or indirect approach’, is characterized by:

- The use of high level objectives that are only loosely defined with no attempt to quantify them without identifying nor describing all the range of options available, most decision makers deal with a limited set without aspiring to exhaust the search.
- Abandoning any clear distinction between objective, goals and actions; see high-level objectives by adopting goals and implementing actions.
- Recognizing that the environment is uncertain and that without knowing the range of events that might take place in the future.

Table 2 Kay's Idea which identifies additional contrast between the two basic approaches:

Substantive rationality	Issue	Indirect Approach
Interactions with others are limited and their response depend on our actions alone	Interactions	The outcome of interactions with others depend on context and their interpretation of our intentions
The relationships between objectives, states, goals and actions are understandable	Complexity	Our understanding of the relationships between objectives, states, goals and actions is imperfect but can be improved by experience
The problem and context can be described by a well specified and estimated analytical model	Abstraction	Appropriate simplification of complex problems must rely on judgement and understanding of context
What happens is what we intended to happen	Intentionality	What happens is the result of complex processes whose totality nobody fully understands
Decisions are made on the basis of the fullest possible information	Information	Decisions are recommended and made acknowledging that only limited knowledge is or can be available
The best outcome is achieved through a conscious process of maximization	Adaptation	Good results are obtained through continual adaptation to constantly changing conditions
Rules and guidelines can be defined that allow people to find the correct solutions	Expertise	Experts can do things that others cannot – and can only learn with difficulty

2.5. MODEL DISCUSSION

It has been stated in the above session many researchers give a direction about model classification for freight transport but it can easily be seen from the above for an investigation of this research paper; Time series freight demand forecasting has been selected due to these reasons based on the report on quick response freight manual. (Cambridge Systematics, Sep 1996)

- ❖ Influence of economy
- ❖ Globalization of business
- ❖ Packaging materials
- ❖ Recycling
- ❖ Fuel prices and user charges

Again due to the following reasons, an Activity based freight demand forecasting is selected for discussion of this paper.

- ❖ Industrial location patterns
- ❖ Globalization of business
- ❖ Just-in-time inventory practice
- ❖ Centralized warehousing

Due to all these reasons and together with the availability of data three models including Four-stage aggregate model selected here for discussion. Why four-stage is because it begins by considering the number of trips generated along the origins and destinations.

2.5.1. TIME SERIES MODEL

Forecasting methods can be classified as qualitative or quantitative. Qualitative methods generally involve the use of expert judgment to develop forecasts. Such methods are appropriate when historical data on the variable being forecast are either not applicable or unavailable. Quantitative forecast methods can be used when

1. Past information about the variable being forecast is available
2. The information quantified
3. It's reasonable to assume that the pattern of the past will continue into the future.

Table 3 Alternative forecasting approaches (Llvento, Titus Awokuse and Tom, 2004)

Categories	Application	Specific Techniques
Qualitative Techniques	Useful when historical data are scarce or non-existent	Delphi Technique Scenario writing Visionary forecast Historic Analysis
Casual Techniques	Useful when historical data are available for both the dependent (forecast) and the independent variables.	Regression models Econometric models Leading indicators
Time Series Techniques	Useful when historical data exists for forecast variable and the data exhibits a pattern	Moving average Auto regression models Seasonal regression models Exponential smoothening Trend projection

Casual forecasting methods are based on the assumption that the variable that we are forecasting has a cause and effect relationship with one or more other variables.

If the historical data are restricted to the past values of the variable to be forecast, the forecasting procedure is called the time series method and the historical data are referred to as time series.

A time series is a stochastic process which is to mean capable of predicting in a sequence of observations of random variable not only for transport engineering but also in day to day activity of a sector, organizations and the like (Joseph, Chow, Heon, Yang, & C.Regan, 2010).

Time series analysis provides tools for selecting a model that can be used to forecast of future events. Modeling the time series is a statistical problem time series models assume that observations vary according to the probability distribution about an underline function of time.

Forecasting studies have been conducted by transportation analysts for virtually each mode. However there have been relatively few railroad forecasting studies.

According to Pyndyck and Rubinfeld (1991) there are three general classes of forecasting with time series models. Each class involves amount of complexity and each assumes a different amount of complexity and each assumes a different amount of understanding concerning the structure one is trying to model.

One class of forecasting time series model presume to know little or nothing about causal relationships that affect the dependent variable that tries to forecast. Here the past behavior of the dependent variable examine to infer some-thing about its future behavior. The method employed to produce the forecast could be as simple as linear extrapolation or could involve the use of a complex stochastic mode. The decision to employ a time series model usually occurs if little is known about the determinants of the dependent variable, when ample historical data is available, and when the model is to be used for short term forecasting.

Another class of forecasting model is single equation regression models. In this class of models the dependent variable is explained by a single equation involving a number of independent variables. The forecasting equation is often time dependent so that one can forecast the response over time of the dependent variable to the forecast values of the independent variables.

The third class of forecasting time series model as suggested by Pyndyck and Rubinfeld (1991) is multi-equation simulation models. In this class of model the dependent variable may be a function of many explanatory variables which is related to each other as well as to the dependent variable through a group of equations. The construction of simulation model starts with the specification of a group of relationships, each of which is fitted to a variable data. Simulation involves solving these equations simultaneously (Michael W., Xiaohua, & Jerry, 1999).

There are three time series forecasts

Point forecast:- a single number or a ‘best guess’. It doesn’t provide information on the level of uncertainty around the point estimate/ forecast.

Interval forecast:- relative to point forecast, this is the range of forecasted values which is expected to include the actual observed value with some probability.

Density forecast: - this type of forecast provides information on the overall probability distribution of the future values of the time series of interest.

Characteristic of Time series

Any given time series can be divided into four categories: trend, seasonal components, cyclic components, and random fluctuations (Chatfield, 1996).

Trend:- the trend is a long term, persistent downward or upward change in the time series value. It represents the general tendency of a variable over an extended time period.

Seasonal components:- seasonal components of a time series refer to a regular change in the data values of a time series that occurs at the same time every year. The seasonal reputation may be exact (deterministic seasonality) or approximate (stochastic). Sources of seasonality are technologies, preferences, and institutions that are linked to specific times of the year. It may be appropriate to remove seasonality before modelling macroeconomic time series (seasonality adjusted series) since the emphasis is usually on non-seasonal fluctuations of macroeconomic series.

Cyclic components:- refer to periodic increases and decreases that are observed over more than a one-year period. It contrasts to seasonal components, these types of variations are also business cycles. These components cover a long time period and are not subject to a systematic pattern that is easily predictable. Cyclic variations can produce peak periods known as booms, and through periods known as recessions.

Random fluctuations:- refer to irregular variations in time series that are not due to any of the three time series components: trend, seasonality, and cyclic. This is also known as residual or error component is not predictable and is usually eliminated from the time series through data smoothing techniques. Stationary time series refers to a time series without trend, seasonal or cyclic components. A stationary time series contains only a random error component. Analysis of time series always assumes that the value of the variable, Y_t at time period t , is equal to the sum of the four components and represented by

$$Y_t = T_t + S_t + C_t + R_t \dots \dots \dots 1$$

According to Chatfield there are many models existed adjusted to a trend relating trend, seasonal component and cyclic component without random fluctuation because random components are non-justifiable.

Mathematical Model

The simplest model suggests that the time series is the constant value b with variations about b determined by a random variable ε_t .

$$X_t = b + \varepsilon_t \dots \dots \dots 2$$

The upper case X_t represents the random variable that is the unknown demand at the time t , while the lower case symbol x_t is the value that has actually been observed.

The random variations ε_t about the mean value is called the noise, & is assumed to have a mean value of zero & a given variance.

The noise variations and the linearity of the equations depend on the time periods and which are independent variable. It is also common to assume that the noise variations in two different time periods are independent. Specifically

$$E[\varepsilon_t] = 0, \text{Var}[\varepsilon_t] = \sigma\varepsilon^2, E[\varepsilon_t\varepsilon_w] = 0 \dots \dots 3 \text{ for } t \neq w$$

A more complex model includes a linear trend b_1 for the data

$$X_t = b_0 + b_1 t + \varepsilon_t \dots \dots \dots 4$$

Eq (1) and Eq (2) are special cases of a polynomial model

$$X_t = b_0 + b_1 t + b_2 t^2 + \dots \dots \dots b_n t^n + \varepsilon_t \dots \dots 5$$

A more complex model and seasonal variations might include transcendental functions. The cycle of the model is 4. The model might be used to represent the 4 seasons of the year.

$$X_t = b_0 + b_1 \sin \frac{2\Pi t}{4} + b_1 \cos \frac{2\Pi t}{4} + \varepsilon_t \dots \dots \dots 6$$

The time series is the function of only of time and parameters of the models.

$$X_t = f(b_0, b_1, b_2, \dots \dots b_n, t) + \varepsilon_t \dots \dots \dots 7$$

Since the value of f is constant at any given time and the expected value of ε_t is zero.

$$E[X_t] = f(b_0, b_1, b_2, \dots \dots b_n, t) \dots \dots \dots 8$$

$$\text{And } \text{Var}[X_t] = V[\varepsilon_t] = \sigma\varepsilon^2 \dots \dots \dots 9$$

The model supposes that there are two components of Variability for the time series; the mean value varies with time and the difference from the mean varies randomly.

Fitting Parameters of the Model

For time series, most methods recognize that recent data are more accurate than aged data. Influences governing the data are likely to change with time so a method should have the ability of de-emphasizing old data while forecasting new.

The time series model includes one more parameters. The estimated values of this parameters

$$b_1^{\wedge}, b_2^{\wedge}, \dots \dots \dots, b_n^{\wedge}$$

The procedure also provide estimates of the standard deviations of the noise, that is $\sigma_{\varepsilon}^{\wedge}$

According to statisticians all values in a given sample are equally valid but if the model is actually changing over time, it's better to use a method that gives less weight to old data and more weight to the new which is called moving average.

In general, the moving average estimator is the average of the last m observations.

$$b^{\wedge} = \sum_{i=k}^t \frac{x_i}{m}, \dots \dots \dots 10$$

Where $k=t-m+1$. The quantity m is the time range & the parameter of the method.

Forecasting from the model

In the analysis, let the current time be T, and assume that the demand data for periods 1 through T are known. Even if the time series actually followed the assumed model, the future value of the noise is unknowable.

Assuming the model is correct

$$X_{T+\tau} = E[X_{T+\tau} + \varepsilon_t] \text{ where } E[X_{T+\tau}] = f(b_0, b_1, b_2, \dots \dots b_n, T + \tau)$$

Estimating the mean value of $X_{T+\tau}$ through parameters from data for times 1 through T.

The expected value for the random variable as a function of τ will then be

$$X_{T+\tau}^{\wedge} = f(b_0^{\wedge}, b_1^{\wedge}, \dots \dots b_2^{\wedge}, T + \tau)$$

Using specific value of τ this formula provides the forecast for period $T+\tau$. For the moving average, statistician adopts the model $X_t = b + \varepsilon_t$

Because this model has a constant expected value over time, the forecast is the same for future periods.

$$X_{T+\tau}^{\wedge} = b^{\wedge}$$

Assuming the model is correct the forecast is average of m observations all with same mean and standard deviation σ_{ε} . Because the noise is normally distributed the forecast is also normally distributed with mean b and standard deviation $\frac{\sigma_{\varepsilon}}{\sqrt{m}}$.

Measuring Accuracy of the forecast

The error in a forecast is the difference between the realization and the forecast,

$$e_{\tau} = X_{T+\tau} - X_{T+\tau}^{\wedge}$$

Assuming the model is correct

$$\varepsilon_{\tau} = E[X_{T+\tau}] + \varepsilon_t - X_{T+\tau}^{\wedge}$$

Good characteristic of the forecast $X_{T+\tau}^{\wedge}$ is that it is unbiased. For unbiased estimate, the expected value of the forecast is the same as the expected value of the time series.

Because ε_t is assumed to have a sum of zero, an unbiased forecast implies $E[\varepsilon_{\tau}] = 0$

Moreover, the fact that the noise is independent from one period to the next means that the variance of the error is

$$Var[\varepsilon_{\tau}] = Var[E[X_{T+\tau}] - X_{T+\tau}^{\wedge}] + Var[\varepsilon_{T+\tau}]$$

$$\sigma_{\varepsilon}^2(t) = \sigma_E^2(\tau) + \sigma^2$$

The term has two parts

1. Due to the variance in the estimation of the mean $\sigma_E^2(\tau)$, &
2. Due to the variance of the noise σ^2

Due to the inherent in accuracy of the statistical methods used to estimate the model parameters and the possibility that the model is not exactly correct, the variance in the estimation of the means is an increasing function of τ .

For moving Averages,

$$\sigma_{\varepsilon}^2(\tau) = \frac{\sigma^2}{m} + \sigma^2 = \sigma^2 \left[1 + \frac{1}{m} \right]$$

The variance of the error is a decreasing function of time. For large m the model is to become correct.

The forecast error, $e_t = x_t - x_t^{\wedge}$

One common measure of forecasting error is the mean absolute deviation, MAD and sometimes MSE (Mean Squared Error) and also MAPE (Mean Absolute Percentage Error) are also used.

$$MAD = \frac{\sum_{i=1}^n |e_i|}{n}$$

Where n error observations are used to compute the mean.

The sample standard deviation of errors is also a useful measure

$$s_e = \frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n - p} = \frac{\sum_{i=1}^n (e_i)^2 - n\bar{e}^2}{n - p}$$

Where $\bar{e}$ is the average error and p is the number of parameters estimated for the model.

As n grows, the MAD provides a reasonable estimate of the sample standard deviation.

$$s_e \approx 1.25MAD$$

If the random noise comes from normal distribution an interval estimate of the forecast can be computed using the *Students t* distribution,

$$X_{T+\tau}^{\wedge} \pm \frac{t_{\alpha}}{2s_e(\tau)}$$

The parameter $t_{\alpha/2}$ is found in a student distribution table with $n-p$ degree of freedom.

Moving Average

This method assumes the time series follows a constant model. i.e Eq(1) given by $X_t = b + \varepsilon_t$

Estimate the single parameter of the model a average of the last m observations.

$$b^{\wedge} = \sum_{i=k}^t \frac{x_i}{m} \dots \dots \dots 12$$

Where $k = t - m + 1$ the forecast is the same as the estimate.

$$X_{T+\tau}^{\wedge} = b^{\wedge}$$

The moving average forecasts should not begin with m periods of data are available. For all three or more estimates the moving average lags behind the linear trend, with the lag increasing with m. because of the lag, the moving average underestimates the observations as the mean is increasing.

The *lag* in time & the *bias* introduced in the estimate are

$$lag = \frac{(m-1)}{2}, \quad bias = -\frac{a(m-1)}{2}$$

The moving average forecasts of τ periods into the future increases the effects.

$$lag = \tau + \frac{(m-1)}{2}, \quad bias = -a[\tau + (m-1)/2]$$

The moving average estimator is based on the assumptions of a constant mean.

Exponential smoothing for the constant model

Like that of the moving average, this method also assumes that the time series follows a constant model, $X_t = b + \varepsilon_t$. The parameter value b is estimated as the weighted average of the last observations & the last estimate.

$$b_T^{\wedge} = \alpha x_T + (1 - \alpha)b_{T-1}^{\wedge}$$

Where α is the parameter n the interval $[0, 1]$.

Or $b_T^{\wedge} = b_{T-1}^{\wedge} + \alpha(x_T - b_{T-1}^{\wedge})$ the new estimate is the old estimate plus proportion of the observed error.

Replacing $b_{T-1}^{\wedge}$ with its equivalent

$$b_T^{\wedge} = \alpha x_T + (1 - \alpha)(\alpha x_{T-1} + (1 - \alpha)b_{T-2}^{\wedge})$$

Continuing this, the estimate is really weighted sum of all past data.

$$b_T^{\wedge} = \alpha(x_T + (1 - \alpha)x_{T-1}^2 + (1 - \alpha)^2 x_{T-2} + \dots + (1 - \alpha)^{T-1} x_1)$$

Because α is a fraction, recent date has a great weight than more distant data. The larger values of α provide relatively greater weight to more recent data than the smaller values of α .

The lag characteristic, similar to one associate with the moving average estimate, can be shown

$$lag = \frac{1 - \alpha}{\alpha}, bias = \frac{-1(1 - \alpha)}{\alpha}$$

For large value of α , a grater lag in response to the trend.

Error in exponential smoothening are $e_\tau = x_{T+\tau} - x_{T+\tau}^{\wedge}$ assuming the model is correct

$$e_\tau = E[X_{T+\tau}] + \varepsilon_\tau - x_{T+\tau}^{\wedge}$$

$$E[e_\tau] = E[X_{T+\tau}] + E[\varepsilon_\tau] - E[\hat{X}_{T+\tau}]$$

$$E[\hat{X}_{T+\tau}] = \alpha b[1 + (1 - \alpha) + (1 - \alpha)^2 + \dots + (1 - \alpha)^{T-1}]$$

As the T goes to infinity, the series in the brackets goes to $\frac{1}{\alpha}$,

$$E[\hat{X}_{T+\tau}] = b \text{ and } E[e_\tau] = 0$$

Because the estimate at any time is independent of noise at a future time, the variance of the error is

$$Var[\hat{X}_{T+\tau}] + Var[\varepsilon_{T+t}]$$

$$\sigma_\varepsilon^2(\tau) = \sigma_E^2(\tau) + \sigma^2.$$

The variace of the error has two parts,

1. due to the variance in the estimate of the mean $\sigma_E^2(\tau)$
2. due to the variance of noise σ^2

For the exponential smoothng

$$\sigma_E^2(\tau) = \frac{\alpha\sigma^2}{2-\alpha}$$

Thus assuming the model is correct, the error of the estimate increases as increases. Setting the estimating error for the moving average and the exponential smoothing equal,

$$\sigma_E^2(\tau) = \frac{\sigma^2}{m} = \frac{\alpha\sigma^2}{2-\alpha}$$

Solving for in terms of m ,

$$\alpha = \frac{2}{m+1}$$

Analysis of the linear trend model

One way to overcome the problem of responding to trends in the time series is to use a model that explicitly includes a trend component,

$$X_t = a + bt + \varepsilon_t$$

Which is of the form eq(2) a linear model of the form

$$X_T = a_T + b_T(t - T) + \varepsilon_t$$

Estimate two parameters of a_T & b_T from the observations previous to time T. forecasts will be made by projecting the estimated model in the future.

$$X_{T+\tau} = a_T + \tau b_T \text{ for } \tau > 0$$

Regression Analysis

The formulas for determining the parameter estimates to minimize the least square differences between the line and observations. The last m observations used to estimate

$$x_{T-m+1}, x_{T-m-2}, \dots, x_T$$

Define the following sums

$$s_1(t) = \sum_{k=t-m+1}^t x_k$$

$$s_2(t) = \sum_{k=t-m+1}^t (k-t)x_k$$

The estimates of the parameters are determined by operations on these sums

$$a_t^\wedge = \frac{6s_2(T)}{m(m+1)} + \frac{2(2m-1)s_1(T)}{m(m+1)}$$

$$b_T^\wedge = \frac{12s_2(T)}{m(m^2-1)} + \frac{6s_1(T)}{m(m+1)}$$

The expressions of the sum are awkward for spread sheet computation.

$$s_1(t) = s_1(t-1) - x_{t-m} + x_t$$

$$s_2(t) = s_2(t-1) - s_1(t-1) + mx_{t-m}$$

Exponential smoothening Adjusted for the trend

There is also a variation of the exponential smoothing method that explicitly accounts for a trend component.

Assume the linear model again

$X_T = a_T + b_T(t - T) + \varepsilon_t$ The new method simultaneously estimates the constant and trend component using two parameters α & β .

$$\alpha_T^\wedge = \alpha x_T + (1 - \alpha)(\alpha_{T-1}^\wedge + b_{T-1}^\wedge)$$

$$b_T^\wedge = \beta(a_T^\wedge - \alpha_{T-1}^\wedge) + (1 - \beta)b_{T-1}^\wedge$$

Forecasts are made with the expression

$$x_{T+\tau}^\wedge = a_T^\wedge + b_T^\wedge \tau$$

$$\alpha = 1 - (1 - \delta)^2, \beta = \frac{\delta^2}{1 - (1 - \delta)^2}$$

Selecting method for time series model

Forecasting time series method based purely on historical data it's impossible to filter out all the noise. The problem is to set parameters that find an acceptable tradeoff between the fundamental process and the noise. If the process is changing very slowly, both the moving average and the

regression approach should be used with a long stream of data. For the exponential smoothing method, the value of α should be small to de-emphasize the most recent observations. Stochastic variations will be almost entirely filtered out.

If the process is changing rapidly with rapid changes in the linear trend each of the methods may be in trouble this is because it's difficult to separate changes in the process from random changes. The time ranges must be set small for moving average and regression methods, resulting in sensitivity to random effects.

Both the moving average and regression methods have the disadvantage that they are most accurate with respect to forecasts in the middle of the time range. Unfortunately, all interesting forecasts are in the future, outside the range of the data. With all methods, though, the accuracy of the results decreases with the distance into the future one wishes to forecast.

2.5.2. ACTIVITY-BASED MODEL

In a growing globalized context and consumption economy freight transport is of crucial importance. Activities of firms are expanding, even across borders. This causes an increase in logistics activities of firms as they become more dynamic. Public and private decision makers need to take these trends into consideration with regard to their decisions and a better projection of freight traffic flows becomes necessary. Being able to understand the drivers of freight flows makes it possible to forecast freight flows in the future and to calculate the impact of different policies on freight traffic. It will put policymakers in the position to get a better insight in the way the transport of goods comes about. Still freight demand modelling is lacking behind on the efforts made in passenger transport models.

Hence, there is a growing need for models that can predict freight flows more accurately. Here the category of activity-based models comes into play, as they are able to better represent the link with the economy, interactions between different actors and logistic elements inherent of freight movement. The development of a comprehensive and reliable freight transport model is needed (Tabitha, Karien,, Caris, Tom, & Gerrit K., 2011).

Today, most state-of-the-practice models in freight transport are still four-step models, where the focus is on individual trips. These models have as main disadvantage that they are looking at the

aggregated flows between zones and cannot model flows at a more detailed level. For that, they are missing out on the behavioral aspects behind transport and are having errors due to aggregation. Most importantly they are lacking elements of logistics decision making. The importance of incorporating logistics decisions and behavioral aspects in freight transportation model is widely recognized (Tatineni and Demetsky (2005), Tavasszy et al. (1998), MOTOS (2006) and Liedke (2009)). Some of the more recently developed four-step models are already incorporating logistic components (Tavasszy et al. (1998), SCENES Consortium (2000) and Yin et al. (2005). However, these models are on an aggregated level and are not taken into account aspects of the different agents.

Recent trends in freight modelling are moving to agent-based models, which are part of the group of activity-based models and focus on each freight agent separately. Therefore they are better able to model their individual operational decisions and their interactions concerning logistics and transports. Furthermore a disaggregated approach is applied by looking at trips and decisions on microscopic scale and no longer to aggregate flows between different zones. This enables the understanding and representation of roles that each actor plays in the freight transportation system, as also the interactions between actors. Besides, it is possible to incorporate changes in actors and their interactions over time. These elements are of fundamental importance in the development of more behavioral models for the freight system (Roorda et al., 2010). The disaggregated approach of these models, together with the representation of the different actors, enables better modelling possibilities for logistics decisions (Tabitha, Karien., Caris, Tom, & Gerrit K., 2011).

Trends like just-in-time (JIT) logistics are having an impact on the modes used, and size and frequency of shipments. Two of the main trends are stated in Hesse and Rodrigue (2004) are:

- Demand-side orientation of activities. While traditional delivery was primarily managed by the supply side, current supply chains are increasingly managed by demand.
- Logistics services are becoming complex and time sensitive. This has led to the point that many firms are now management to third-party logistics providers. These providers benefit from economies of scale and scope.

This leads to the need for agent-based models at a microscopic level. Roorda gives several reasons for implementing agent-based modelling. First of all, these are diverse actors involved in the production and distribution of goods, none of which may have full control or even knowledge of all decisions made throughout the supply chain. Secondly, interactions between firms are diverse and finally business models are changing over time. So a close follow up of all these interactions are requested to have a more realistic image of freight transport flows. Due to the modelling at a micro level, it is possible to look at individual instead of aggregated flows. This gives the opportunity to include individual firm characteristics and detailed representation of commodity groups. When looking at single movements of goods, more information of a shipment may be represented that would go lost in aggregated data.

Here in an activity-based freight model the groups of agents that are mostly used are shipper, receivers/customers, carriers/transporters and forwarders and may be expanded to include politics as in the Good Trip model.

The receiver or customer initiates the demand and chooses a supplier to deliver the required goods. After the shipper is chosen the receiver decides on the delivery moment, shipment size and whether he conducts the transport himself or not (Boerkamps et al., 2000).

In wiestjindawat et al., (2007) shippers play a major role in the section of carrier and vehicle choice. (Boerkamps et al., 2000) state that shippers are often responsible for transportation and therefore have to decide on mode choice vehicle type and vehicle size. Furthermore, they decide on groupings of goods types, product range to offer, location of facilities, and availability of distribution channels and whether or not to maintain own transport services. The carriers or transporter are responsible for the actual movement of the shipments and the tour planning problems. In Liedtke (2009) forwarders have the extra responsibility to build and coordinate transport chains. Freight movements are also indirectly influenced by politics (Boerkamps et al., 1999). Politics have an influence on the market structure, as they are responsible for an optimal spatial-economic organization. On the transport market they may make a difference by regulating the accessibility and mobility of the transport.

According to Boerkamps and Roorda many companies act as both shipper and receiver. Ultimately, a simple actor may fulfill all roles in the supply chain, that is, as receiver of goods delivers, as shipper and/ or as transporter of shipments. An actor may be active in different activity types, for example: consumer, supermarket, distribution center, production factory, etc. at the same time, they may own goods and may provide transportation services as well. (Tabitha, Karien,, Caris, Tom, & Gerrit K., 2011)

Roorda et al., (2010) established a new set of agents in his framework. The main agents are business establishments, firms and facilities (commodity, business service and logistic service).

- Business establishment: an organization at a specific location that produces, processes, or stores commodities, or provides business or logistics services. A business establishment may include several different facilities.
- Firm: an organization that owns or operates one or more business establishments. Within a logistics firm, business establishments are different locations may be integrated into a logistics network.
- Commodity production facility: one of the internal resources of business establishment. The function of a commodity production facility is to produce or process commodity inputs.
- Business service facilities: provide services instead of commodities.
- Logistics service facility: provides logistics services, including transportation and inventory.
- End consumers initiates demand for commodities (Tabitha, Karien,, Caris, Tom, & Gerrit K., 2011).

Another written literatures and journals used to study the model for passenger transport only but some tell many available freight transportation models have aggregate nature hence many current models are short in terms of logistics elements such as, use of intermediate handling facilities, determining optimum shipment size, mode choice and multi modal shipments.

An activity based mode usually called a behavioral freight uses a fresh approach in modelling freight demand by incorporating a modular structure and working at a disaggregate level to firm

as a basic decision making unit in the freight market. The large scale model simulates commodity movements at the dis aggregate level of firm to firm (Jose Castiglione & Gliele, 2015)

FAME was introduced as a freight activity-based modeling framework with five basic modules. FAME framework. In the first module, all firms in the study area are synthesized and their basic characteristics, including industry type, employee size, and location are identified. Based on each firm's characteristics, types and amounts of incoming and outgoing goods are determined, and trade relationships are formed for each pair of supplier-buyer firms in the second module. Third and fourth modules deal with logistics choices. In the third module, the shipment sizes are defined based on the acquired information on the firms' characteristics, and shipping modes are determined in the fourth module. Finally, in the last module, the impact of the goods movements' on transportation network is investigated.

Four categories of data are required for developing FAME: information on business establishments for the first module, aggregate freight flows (OD matrix) for the second module, detailed information on a sample of individual shipments and supply chains for the third and fourth tasks, and specifications of the transportation networks for the last module (Jose Castiglione & Gliele, 2015).

2.5.3. FOUR-STAGE AGGREGATE MODEL

This approach involves the following:

- Estimation of freight generations and attractions by zone.
- Destination of generated volumes to satisfy 'trip-end' generations and attractions constraint. The usual methods for this task are linear programming or gravity model.
- Assignment of origin-destination movements to modes and routes. (Juan de. Dios Ortuzar, Lewis G-Willumsen, 2011)

Freight Generations and Attractions

At the level of the individual firm, the number and type of freight trips can be regarded as the outcome of a series of decisions about products, markets, production locations, delivery times and frequencies, transport modes, types of vehicles and routes. At a collective level, the decisions of the individual firms result in transport and traffic flows at various geographic scales.

The relation between economic activity and transport flows is not a static one. It is influenced by internal and external factors such as company strategies, technological developments, government rules, etc. (Mirajam H.E., Wilhem, & Lori, 2002). The trip generation and attraction stage of classical transport model aims at predicting the total number of trips generated by (O_i) and attracted to (D_j) each zone of the study area.

- Direct survey of demand and supply may be undertaken for major flows for some homogeneous products like coffee, sugar, petroleum products, iron ore, coal, cement, fertilizers, grains, construction machineries, vehicles and others. This approaches is usually used for inter-urban and is not recommended for urban transport.
- The use of macro-economic models which relate the input-output nature based on regional rather than national data.
- Growth-factor methods are often used in forecasting future trip ends.
- Zonal multiple linear regression is often used to obtain more aggregate measures of freight generations and attractions.

Important variables in freight trip production and attraction

-  No of Employees;
-  No of sales
-  Roofed area of firm
-  Total area of firm

There are two approaches that usually used for estimating the trip generation and attraction:

1. Growth factor method
2. Regression method

1. Growth Factor Method

Growth factor modes tries to predict the number of trips produced or attracted by a commodity group as linear function of explanatory variables. The procedure involves applying growth factor

to base line traffic data or economic variable involved to project the future freight travel demand. The growth factor approach is classified in two types.

1. Based on historical traffic counts
2. Based on forecast of economic activity

The first approach involves the direct application of a growth factor, calculated based on historical traffic information, to the base line of traffic data.

The second approach recognizes that demand for freight transportation is underlying economic activities (e.g. Employment, population, income, etc.)

Growth factor based on Historical traffic trends

This trend assumes the availability of at least two years of historical data for the freight demand variable being forecast. Using two years of historical data, n AGF (Annual Growth Factor):

$$AGF = \left(\frac{T_2}{T_1}\right)^{1/(Y_2-Y_1)} \dots \dots \dots 13$$

Where T_1 is the freight demand in year Y_1

T_2 is the freight demand in year Y_2

The annual growth factor can be applied to predict future demand (T_3) for future year (Y_3) as

$$T_3 = T_2 AGF^{Y_3-Y_2} \dots \dots \dots 14$$

If more than two years of historical data are available for the variable to be forecast, it's suggested to have plotted and examined to ensure that they exhibit a relatively steady growth rate over time. If year to year changes appear erratic than assumption underlying the simple procedure- relatively constant growth rate over time should be in question.

Based on Economic Projection

Simple procedure for forecasting freight using projection of future demand or output for the goods being transported.

To simplify the approach for deriving forecasts of future freight traffic from economic forecasts, it can be assumed the demand for transport of specific commodity is directly proportional to an economic indicator variable that measures output or demand for the commodity.

The basic steps involved in the process are presented as follows.

- ✚ Select the commodity or industry groups that will be used in the analysis
- ✚ Obtain or estimate the distribution of base year freight traffic by commodity or industry group. If actual data on the distribution are not available, state or national sources may be used to estimate this distribution.
- ✚ Determine the annual growth for each commodity or industry group as follows

$$AGF = \left(\frac{I_2}{I_1}\right)^{1/(Y_2-Y_1)} \dots \dots \dots 15$$

Where I_1 is the value of economic indicator in year Y_1 and I_2 is the value of economic indicator in year Y_2 .

- ✚ Using annual growth factor and base-year traffic, calculate forecast-year traffic for each commodity or industry group.

$$T_f = T_b(AGF)^n \dots \dots \dots 16$$

Where n is the no of years in the forecast period.

- ✚ Aggregate the forecast across commodity group or industry group to produce the forecast of total freight demand the most desirable indicator variables are those that measured goods output or demand in physical units (tons, cubic feet, etc.)(Winston, C., 1981)

Strength of Growth factor Method

- ✚ Mostly used to predict the future number of external trips to an area; this is because error cannot be large and there is no simple way to predict them.

Weakness of Growth factor Method

- ✚ Growth factor estimation is difficult

- ✚ Overestimates the total number of trips by approximately 42 %

2. Zonal-based Multiple Regression

This method is made to find a linear relationship between the number of trips produced or attracted by zone and average socioeconomic characteristics of commodity in each zone.

The general form of trip generation is given by:

$$T_i = f(x_1, x_2, x_3, \dots x_i \dots x_k)$$

Where x_i 's are prediction factor or explanatory variable. The most common form of trip generation model is a linear function of the form:- $T_i = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \dots a_ix_i + \dots a_kx_k$

The linear equation will then have the form of $y = a + bx$ where y is trip rate usually within a week because the distance between Addis Ababa and Djibouti takes days

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

$$a = \bar{y} - b\bar{x}$$

Strength of zonal-based multiple regression

- ✚ Can easily use the data since it considers origins and destinations.

Weakness of zonal-based multiple regression

- ✚ The solution is tedious to obtain manually

Distribution Models

There are two aggregate technique mostly used in distributing the generated trips to different zones these are Growth factor method and Gravity model.

The purpose of the distribution model is to produce an origin-destination (OD) matrix for the forecast year. The input of this model is the OD matrix of the base year, the total incoming and outgoing flows per zone in the forecast year, and the level of service (log sums) per OD relation (all per commodity type).

The distribution model works by calculating growth factors per OD relation. Using a gravity model, temporary OD matrices are generated for both the base year and the forecast year. From

these temporary matrices the growth factors can be calculated, and then applied to the actual base year OD matrix (Gerard, et al., 2011).

According to the research which tells about distribution and modal split models for freight transport in Netherlands almost all the time Gravity model is used for freight transport because growth factor like to depend the observed trip pattern, do not consider changes in travel cost, it cannot explain unobserved trips and not suitable for policy studies like introduction of a mode. (Gerard, et al., 2011)

Gravity Model

The gravity model has the following relation:

$$T^k_{ij} = A^k_i B^k_j O^k_i D^k_j e^{(-\beta^k C^k_{ij})}$$

Where k is a commodity type index; T^k_{ij} are tones of product k moved from i to j ; A^k_i, B^k_j are balancing factors; O^k_i, D^k_j are supply and demand for product k at zone i (or j); β^k are calibration parameters, one per product k ; C^k_{ij} are generalized transport cost per ton of product k between zones i and j (Ertander & Stewart, 1990).

The idea of using generalized cost function formulation for freight demand is apparently due to Kresge and Roberts (Ortuzar & Willumsen, 2011).

$$C_{ij} = f_{ij} + b_1 s_{ij} + b_2 \sigma s_{ij} + b_3 \omega_{ij} + b_4 p_{ij}$$

where f_{ij} is the out-of-pocket charge for using a service from i to j ; s_{ij} is door-to-door travel time between i and j ; σs_{ij} is the variability of travel time s ; ω_{ij} is the waiting time or delay from request for service to actual delivery – it may be a long time for maritime transport, for example, and p_{ij} is the probability of loss or damage to goods in transit.

Mode Choice

The modern approach to mode split analysis is to develop disaggregate demand algorithms known as discrete choice models. In contrast to aggregate demand (first generation) models which are based on observed relations for group of travelers, or average patterns at the traffic analysis zone

(TAZ) level, discrete choice models tend to be more efficient in terms of information usage and thus usually require smaller sample sizes.

Discrete choice models yield the probability of individuals choosing a given transportation mode, based on a relative measure of its attractiveness to that of the other modes in option set. This measure of attractiveness is provided through the development of utility function of each of modes. Because of their probabilistic nature and their dependence on utility functions, discrete choice models are also known as random utility models

In preparing a mode choice shipper's decision plays a vital role this is to shipper's extent to which carrier should be used to deliver the goods to their destination. When modelled in a very aggregate level, modal choice is often treated using multinomial logit (MNL) formulation based on generalized cost.

$$c_{ij} = a_1 t_{ij}^v + a_2 t_{ij}^w + a_3 t_{ij}^t + a_4 t_{nij} + a_5 F_{ij} + a_6 \phi_j + \delta$$

Where t_{ij}^v is the in-vehicle travel time between i and j , t_{ij}^w is the walking time to and from stops, t_{ij}^t is the waiting time at stops, F_{ij} is the fare charged to travel between i and j , ϕ_j is the parking cost, and δ is a parameter representing comfort and convenience. If the travel cost is low, then that mode has more probability of being chosen.

If there are two types of modes (1,2) and portions of the trips are done by mode 1 and the other by mode 2 by comparing the cost of traveling from zone i to j by using mode 1 and vice versa one can easily compare and choose the better mode (Sillano & Ortuzar, 2005).

The relation becomes

$$P_{ij}^1 = \frac{T_{ij}^1}{T_{ij}} = \frac{e^{-\beta c_{ij}^1}}{e^{-\beta c_{ij}^1} + e^{-\beta c_{ij}^2}}$$

Assignment

Assignment is the choice of the best route to take the goods from origin and destination. Trains are sometimes scheduled according to a semi-variable time table (Roll-on Roll-off, mail). Here the algorithm from time table based public transport assignment can also be applied to rail freight assignment. More often, freight trains operate in response to demand. In this case, a time table

doesn't exist, not even a line network with headways or frequencies. What is then required is a train formulation algorithm to build the trains journeys and their implied time table. There are different approaches which are used in assignment these include all-or-nothing assignment, incremental assignment, capacity restraint assignment, user equilibrium assignment and others and for freight transportation it has believed user equilibrium assignment is the best one (Ortuzar & Willumsen, 2011).

Equilibrium

This model considers explicitly the decisions of both shippers and carriers for an inter-modal freight network with nonlinear costs and delay functions that vary with commodity volumes.

Freight Network Equilibrium Model (FNEM) treats shippers and carriers sequentially; shippers are assumed to be user optimized trying to minimize the delivered price of commodity sent.

The user equilibrium assignment is based on wardrop's first principle, which states that no decision unilaterally reduce his/ her travel costs by shifting for a given O-D pair as:

$$f_k(c_k - u) = 0: \forall k \text{ And } c_k - u \geq 0: \forall k$$

Where f_k the flow on path k , c_k is the travel cost on path k , and u is the minimum cost.

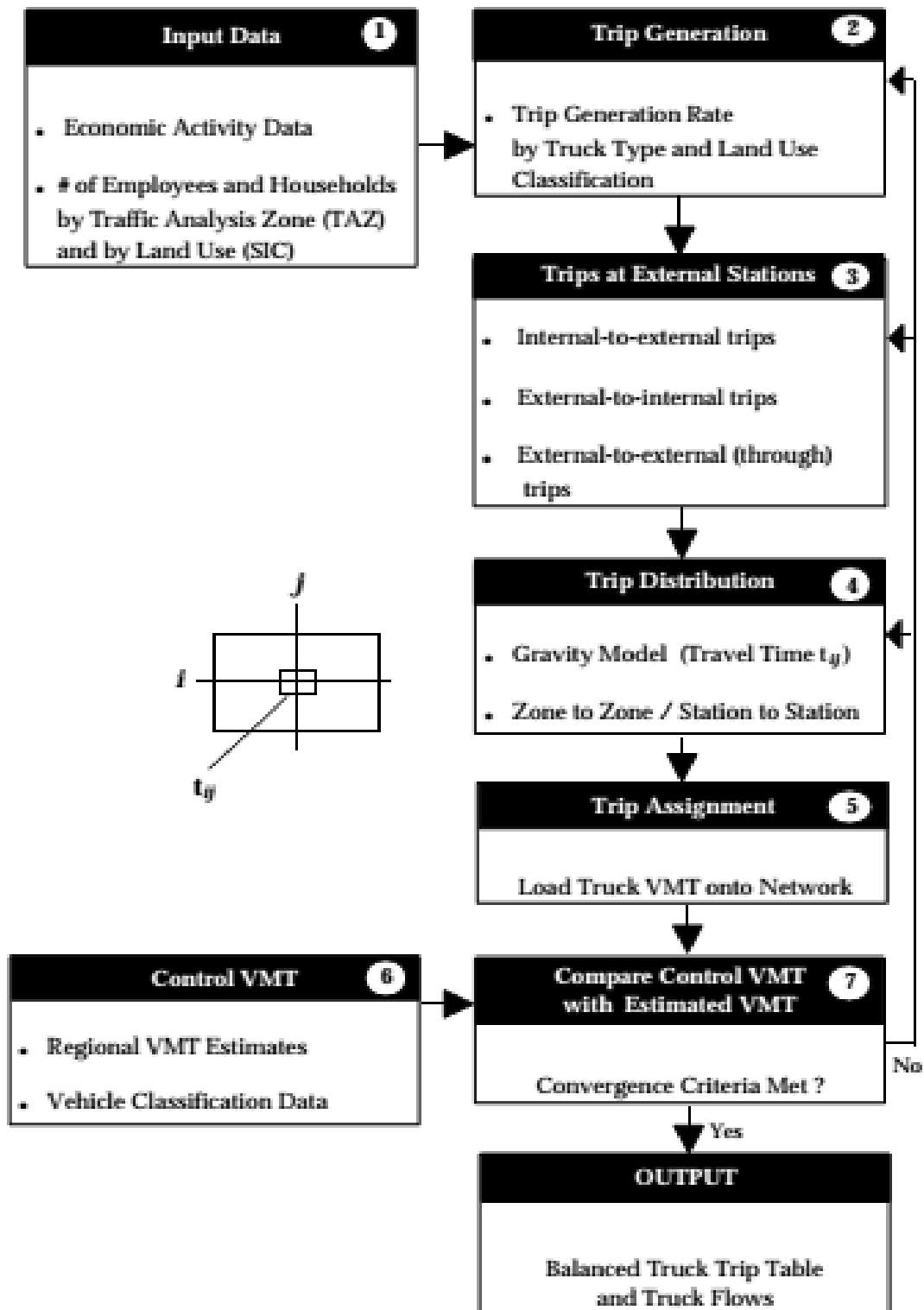


Figure 2 Four stage Aggregate freight demand model (Mc Fadden, D. & Reid, F.A., 1975)

CHAPTER THREE

METHODOLOGY

3.1. INTRODUCTION

‘Modeling freight demand and shipper behavior’ which is a sample of literature review given by Garrido et.al address that, there are many available projection models for freight transport. Most reviews of this paper describes classification models fall into a category of these models which include a time series model, an activity based model, econometric model, four stage aggregate model and others which helps for freight transport. This paper tries to design a method for forecasting freight demand for Addis Ababa- Djibouti rail line using one of the models by comparing the selected and discussed models based on literature and tries to adopt the model in new or other approach.

3.2. METHOD DESIGN

In order to come up with the best approach by adopting the existing forecasting model, for Addis Ababa-Djibouti rail line, a model has been selected from published literature reviews.

From (Joseph, Chow, Heon, Yang, & C.Regan, 2010) , (Michael W., Xiaohua, & Jerry, 1999)and (Chatfield, 1996) literature time series analysis has the following strengths and weaknesses.

Strength of Time Series Model

- Needs available historical data , once the data is there they provide quick answer
- The models can handle more than simple trends (growth & decline) as they can also consider cycles in the data (annual, weekly, daily).
- Can make relation to many variables like traffic Vs time, economic growth Vs time, commodity flow Vs time and others.
- The decisions to employ a time series model usually occurs if little is known about the determinants of the dependent variable.
- The forecasting equation is dependent so that one can forecast the response over time of the dependent variable to the forecast values of the independent variable.

Weakness of Time Series Model

- Used to forecast the future as long as the future is expected to behave like the past.
- Appropriate when forecast is short term and there is insufficient time and resources to build and calibrate a behavioral model.
- The seasonality of reputation may be exact (deterministic seasonality) or approximate (stochastic). Sources of seasonality are technologies, preferences, and institutions that are linked to specific times of the year.

Another model reviewed for this study is an Activity-based forecasting mode and it has its own strengths and weaknesses as it has been taken from different journal articles.

Strength of Activity Based forecasting model

- Can allocate better networks if firm types are registered well
- Appreciate partnership for any pair of supplier and buyer because it scores the appropriateness of all possible suppliers for the given firm type (Wiestjindawat et.al (2007).
- Can predict location, time, mode and route of transportation, similar characteristics of activity.
- Appreciate peoples to have private cars because owners can make industry to industry and farm to processing industry good movement.
- Mostly used in shipper, receivers/customers, carriers/transporters and forwarders and may be expanded to include politics as a good trip model.(Roorda et.al)

Weakness of Activity-Based forecasting model

- It doesn't consider the nature of goods that can be transported together. It considers only firm location, industry type and establishment size.
- Can easily be affected by external markets since it considers the behavior.
- Cannot give credit for season ability of goods since some manufacturers produce materials according to the demand and also recognize the time in need Wiestjindawat et.al (2007).

Many research paper wrote about classical four-stage aggregate model by pointing the strengths and weaknesses.

Strength of Four stage Aggregate Model

- It begins with zoning which mean generate the number of trips or activities.
- Can incorporate with other transport mode. It doesn't only forecast rather give solution.
- Can correlate outcome of a series of decision about products, markets, production locations, delivery times and frequency of transport modes, types of vehicles and routes (Mirajam H.E., Wilhem, & Lori, 2002).
- Can make relation with the direct application of a growth factor calculated based on historic traffic information, to base line of traffic data and also recognizes that demand for freight transportation is underlying economic activities (e.g. Employment, population, income, etc.)

Weakness of Four stage Aggregate Model

- Weak to relate with time fluctuation and other situation rather tells counts by house hold size and commodity group of current status.
- It only counts the trip and commodity groups. It doesn't recognize items of different types (which cannot be transported together)

Among the models a four stage aggregate freight forecasting model has been selected for this thesis work due to:-

- ✚ Availability of data which means time series and activity based needs detail data.
- ✚ Unlike time series and activity based models four stage aggregate model begins with zoning in order to generate a trip.
- ✚ Four stage doesn't consider commodity group and behavior rather it only consider the trip between origin and destination.
- ✚ Unlike time series and activity based models forecast four stage aggregate model doesn't only forecast rather give a solution which means it can allocate other transport mode.
- ✚ Four stage aggregate model can distribute the trip and can also assign but time series and activity based cannot do this.

- ✚ Four stage is not seasonal.
- ✚ Since grouping of commodities is not common in Ethiopia rather consideration of general cargo is available, four stage aggregate model is important because it considers the overall trip between origin and destination.

Previous studies used to calculate the number of trips generated by relating annual growth factor with GDP but this thesis work aims to forecast by taking a traffic congestion as a main problem in order to generate a trip along the path and relate it with an annual growth factor of the commodity tonnage through the line. Taking an Annual Growth for the commodity tonnage and traffic count rate of increase as a main discussion of issue. Because an annual growth rate is the rate at which the idea of investigation increase or decrease so that all freight factors which is the rate at which an economy, commodity tonnage, oil price, weather condition, geographical and other changes are all matters the traffic count.

3.3. DESCRIPTION ABOUT CASE STUDY AREA

As of the National Railway Network of Ethiopia (NRNE) the line starts from Sebeta in its west, go southward via Addis Ababa , Akaki , Dukem, Debrezeit, and Mojo then continues north eastwards ,via Welenchiti, Metehara, Awash, Meiso, Diredawa,Dewanle, Djibouti, finally Djibouti port with a length of 656 from Addis Ababa to Dewanle.

In this line there is a great import and export traffic congestion which took a great share of economy (Seada, 2016).

So studying about forecasting model will be helpful to develop a good model for the future This paper tries to work on trips generated on this line, and mode choice only along this path because the demand or all the data tells between origin and destination of course one way only.

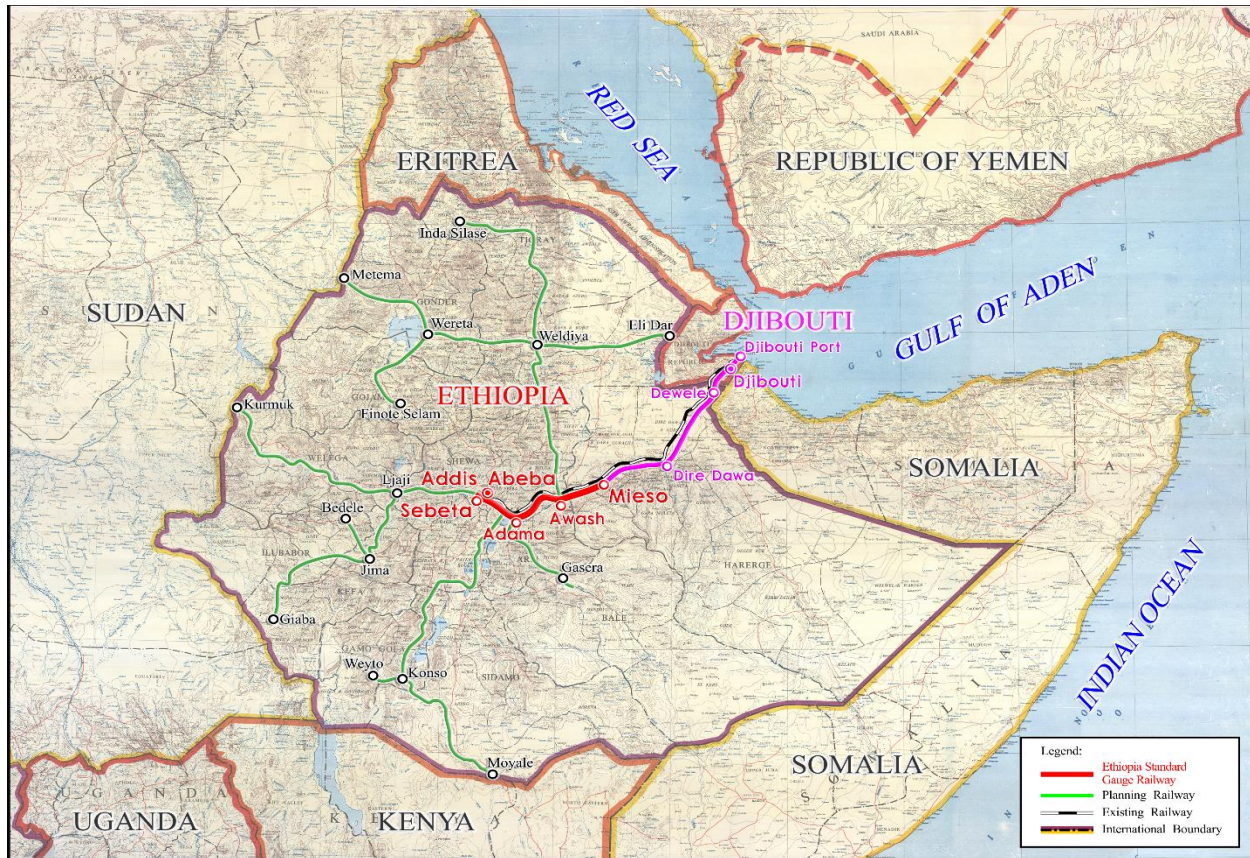


Figure 3 Ethiopia Standard Gauge Location Map (National Railway Network of Ethiopia)

3.4. DATA COLLECTION

This section shows different tables and figures which are collected from different sectors which will be helpful for analysis purpose most of the data that has been gathered are primary data which are found from the organizations and also an interviews of employee's taken from the institutions.

TABLE 4 Recent Historical data from traffic counts report from Addis Ababa-Djibouti

Year	Traffic counts from Addis Ababa to Djibouti	Traffic counts from Djibouti to Addis Ababa	Total trips (In and out)
2002	15,889	16,925	32,814
2003	16,264	18,205	34,469
2004	19,344	20,961	40,305
2005	17,552	15,634	33,186
2006	23,091	15,642	38,733
2007	23,686	24,566	48,252

Source: Ethiopian Transport Authority (2015).

Here all the data given here doesn't describe about the type of commodity transported except the petroleum and oil groups because the truck is other than those import and export commodity groups.

Due to this problem forecasting with an activity based or behavioral model is difficult.

Table 5 Value of export by major commodity groups measured in million birr

Export item	Value of export expressed as of million birr					Annual Average	Annual Average Growth rate
	2001	2002	2003	2004	2005		
Coffee	3,973	7,005	13,792	14,571	13,735	10,615	43.3
Oil seeds	3,864	4,728	5,349	8,257	8,181	6,076	22.2
Pulses	957	1,695	2,264	2,791	4,296	2,401	47.0
Live Animals	546	1,191	2,421	3,602	3,054	2,163	63.7
Meat/Hides	277	446	1,041	1,372	1,365	900	56.5
Fruits/Vegetables	123	422	520	779	807	530	79.7
Others	5,660	10,894	19,724	23,656	25,214	17,029	45.3
Total	15,400	26,381	45,111	55,028	56,652	39,714	41.8

Source: Ethiopian Customs and Revenue Authority

As it can be seen from the table expenses of each commodity group that are going to be transported through the Addis Ababa-Djibouti rail line has increased from year to year with an annual average growth rate of maximum 79.7 percent for vegetables and oils and minimum of 22.2 percent for oil seeds this may be due to:

- ✚ International trading system.
- ✚ GDP between importing and exporting countries.
- ✚ Politics between importing and exporting countries.
- ✚ Oil prices.
- ✚ Others

Table 6 Volume of major export by type of commodity in million ton (MT)

Commodity/Year	1996 E.C	1997	1998	1999	2000	2001	2002	2003	2004	% of items In 2002
Coffee	156	161	148	230	191	246	320	368	403	24.24
Pulses	73	122	110	260	287	347	417	479	513	31.59
Oilseeds	106	171	266	235	152	283	355	383	408	26.89
Meat/Hides	13	23	23	22	18	24	27	29	34	2.04
Fruit/Veg.	37	38	35	41	40	44	48	53	59	3.63
Others	38	56	62	80	114	131	153	178	205	11.61
Total	423	571	644	713	762	1075	1320	1490	1622	100

Source: Ethiopian Customs and Revenue Authority

The tonnage of the commodity groups has increased with an increasing rate highly for pulses even though it's annual average growth rate of the expense is not much like that of fruits and vegetables and its commodity for Meat/Hides is lower about 2.04 % while the expense rating gathered is not as low as oil seeds.

Table 7 Value of import by type of commodity groups (in billion birr)

Commodity Groups	2002	2003	2004	Annual Average	Annual average G.R	Annual Average % share
Industrial supplies	36.5	58.6	76.5	45	36.3	27.49
Capital Goods and accessories	36.5	40.24	63.1	39.7	37.9	22.55
Fuels and lubricants	22.96	36.6	41.2	27.3	26.3	16.55
Food and Beverage	11.9	27.3	32.5	19.12	38.9	11.7
Transport Equipment and Accessories	18.8	23.55	34.5	19	52	11.63
Fertilizer	13.8	19.66	28.5	15.92	43	9.73
Others	0.335	0.66	1.07	0.43	341.7	0.22
Total	140.8	206.6	277.3	163.8	36.01	100.00

Source: Ethiopian Customs and Revenue Authority.

According to ECRA the value of import of the commodity group annual average growth rate ranges 26.3 for fuels and lubricants even though the rate in international market is fluctuating highly and even the need of organizations and country for this goods is high.

Table 8 Volume of major import by commodity group in million ton (MT)

Commodity/Year	2001	2002	2003	2004	2005	% of import items in year 2002
Gen.Cargo	3,099	3,639	4,087	4,964	5,312	53.08
Petroleum	1,963	2,131	2,375	2,648	2,710	31.08
Fertilizer	513	566	618	679	699	8.25
Food Aid	491	520	543	569	574	7.59
Total	6,066	6,856	7,623	8,860	9,295	100

Source: Ethiopian Customs and Revenue Authority

3.5. ASSUMPTIONS AND CONSIDERATIONS OF THE DATA

According to the data gathered for this papers it has been told by interviewee's and seen from the documented data from organizations and the following assumptions and considerations has been taken for analysis purpose.

- ✓ Researches shows and advise to take credit for commodity group and tonnage (FHWA, 1967)
- ✓ According to ESLE 98% of the customers via Djibouti port tries to import good not export. Exporter's data is not available because ESLE has a data of only the shipping system which is to mean exporters may use different country's ship waiting at the port.
- ✓ According to ECX the data that has given tells the trading system in the country and some organization buying another may be for an export of purchased goods which doesn't tell they may use different system of delivery mode.
- ✓ Some countries which may import Ethiopian products use mostly their own shipping vessels rather than Ethiopian ships.
- ✓ Assume the rates will continue with the nearest rate that has been recorded.
- ✓ Assumption has been taken through origin and destination that is Djibouti as origin and Addis Ababa for destination or vice versa. The paper doesn't tell or considers the dry ports along the path or doesn't make network through the line.
- ✓ Assume producers producing an item do not import the same item from other countries. For example consider as if Ethiopia do not buy coffee of other countries.
- ✓ Buyers for a commodity do not export the same commodity to others.
- ✓ Assume that the number of house hold size for the passengers will substitute the number of commodity groups' tonnage for freight because the purpose is enabling to eliminate the problem of the study.
- ✓ An average of 30 ton is carried by a truck.

Due to all this reasons the paper begun with trip generation from the data of traffic counts because the preliminary purpose of the research is resolving this problem.

CHAPTER FOUR

DATA ANALYSIS AND PRESENTATION

4.1. DATA ANALYSIS

In order to generate and attract and work for the number of trips generated between the two zones growth factor method is often used in forecasting future trip ends and also a method called Zonal multiple linear regression is used in order to obtain more aggregate measure of freight generation and attraction.

The analysis in both kinds of methods used here relate truck count with each commodity share has been calculated and with this truck share of each commodity groups a truck count forecasted by relating this figure with the economic expense given to each of the commodities.

Applying Growth factor method for Addis Ababa Djibouti rail line

From table 10 summarizes total trips from export traffic counts i.e from Addis Ababa to Djibouti in 2002 E.C. is 15,889 and also in 2007 is about 23,686 so the Annual growth factor considered from the total number of trips with traffic counts can be calculated from the equation.

$$AGF = \left(\frac{T_2}{T_1}\right)^{1/(Y_2-Y_1)}$$

$$AGF = (23,686/15,889)^{(1/(2007-2002))} = 1.0832$$

Traffic counts for the coming 15 years can be forecasted and summarized as follows from Eqn 4

$$T_{2022} = T_{2007} AGF^{2022-2007}$$

But from table 4; $T_{2007}=23,686$ therefore $T_{2022}=23,686*(1.0832)^{15}=78,546$

The current traffic counts on the base year 2007 E.C that is 23,686; exported item category of that of coffee, pulses, oil seeds, meat/hides, fruits and vegetables and others.

Based on this pulses take a great share of truck traffic as seen in table 6 it takes a great share of commodity tonnage even though it doesn't take great expense gathered in exporting It.:

Table 9 Calculated Share of Exported commodity groups in 2007 E.C on the basis of 2002 E.C

Commodity Category	% of Item for given commodity	Calculated Truck Traffic(2007 E.C)
Coffee	24.24%	5,742
Pulses	31.59%	7,483
Oilseeds	26.89%	6,369
Meat/Hides	2.04%	484
Fruit/Vegetables	3.63%	860
Others	11.61%	2,748
Total	100%	23,686

From table 5 selecting like previous 2002 E.C and 2005 and calculating the annual growth factor like in above equation

Table 10 Calculated Annual Growth on the basis of exported economic variable given in million birr taken from table 5

Export Item	Value of expressed as of million for 2002 E.C	For 2005	Annual Growth Factor
Coffee	7,005	13,735	1.2516
Pulses	4,728	8,181	1.2
Oilseeds	1,695	4,296	1.3634
Meat/Hides	446	1,365	1.4519
Fruit/Vegetables	422	807	1.2413
Others	12,085	28,268	1.3274
Total	26,381	56,552	

The calculation for the above table is done by taking the values expressed in million for the commodity groups given in year 2005 E.C as T_2 and that of 2002 E.C as T_1

$$AGF \text{ for each commodity} = \left(\frac{T_2}{T_1}\right)^{1/(Y_2-Y_1)}$$

Each commodity group are those items to be transported in the export items category.

Once the growth factor is calculated validation of number of trucks for each commodity group is such a simple task.

Here an introduction of a hybrid model is used to calculate for the coming years this is done by analyzing problem is traffic congestion which is easily influence by a question of exporting this item the demand asked by importing countries of Ethiopia's exporting item which is all then be influenced by international economic growth.

Table 11 Expected or forecasted truck traffic counts after 15 years relating historical projection and Economic projection for exported goods

Commodity Category	Yearly truck traffic(2007 E.C)	AGF	AGF^{15}	Truck traffic 2022 (Expected)
Coffee	5,742	1.2516	28.973	166,363
Pulses	7,483	1.2	15.4	115,238
Oilseeds	6,369	1.3634	104.56	665,943
Meat/Hides	484	1.4519	268.57	129,989
Fruits and Vegetables	860	1.2314	22.7	19,522
Others	2,748	1.3274	70	192,360
Total	23,686			1,289,415

Here a great truck traffic count is seen in oilseeds even though the tonnage is not as much to that of pulses this is because oil seeds has a better highly rated with an expense got from an export almost 1.3634 as seen in table 10 but not as much rate to that of meat/hides which is about 1.4519.

Table 12 Export forecasted for major commodity groups with truck traffic counts.

Year/ Commodity	Forecasted truck traffic counts for export commodity Groups					
	Coffee With AGF 1.2516	Pulses With AGF 1.2	Oilseeds With AGF 1.3634	Meat/Hides With AGF 1.4519	Fruits/Vegetables With AGF 1.2314	Other With AGF 1.3274
2007 E.C	5,742	7,843	6,369	484	860	2,748
2008 E.C	7,187	9,411	8,684	703	1,059	3,647
2009 E.C	8,995	11,294	11,839	1,020	1,304	4,842
2010 E.C	11,258	13,553	16,141	1,481	1,606	6,427
2011 E.C	14,090	16,263	22,007	2,151	1,977	8,531
2012 E.C	17,636	19,516	30,005	3,123	2,435	11,324
2013 E.C	22,072	23,419	40,908	4,534	2,998	15,032
2014 E.C	27,626	28,103	55,774	6,583	3,692	19,954
2015 E.C	34,577	33,724	76,043	9,557	4,546	26,486
2016 E.C	43,277	40,468	103,677	13,877	5,599	35,158
2017 E.C	54,165	48,562	141,353	20,147	6,894	46,669
2018 E.C	67,793	58,274	192,720	29,252	8,490	61,948
2019 E.C	84,850	69,929	262,755	42,470	10,484	82,230
2020 E.C	106,198	83,915	358,240	61,663	12,873	109,151
2021 E.C	132,917	100,698	488,424	89,528	15,852	144,888
2022 E.C	166,363	115,238	665,943	129,989	19,522	192,360

Again from table 4- total number of trips as seen from import traffic counts i.e. from Djibouti to Addis Ababa is 16,925 in year 2002 E.C and 24,566 in 2007 E.C so the annual growth factor can be calculated as

$$AGF = \left(\frac{T_2}{T_1}\right)^{1/(Y_2-Y_1)}$$

$$AGF = (24566/16925)^{1/(2007-2002)} = 1.077$$

From Eq 4 total traffic counts (import) for the coming 15 years is then be:

$$T_{2022} = T_{2007} AGF^{2022-2007}$$

Again from the table $T_{2007}=24,566$ therefore $T_{2022}=24,566*(1.077)^{15}$

$$T_{2022}=74,743 \text{ annual trips}$$

The current traffic counts on the base year 2007 E.C that is 24,566 imported item category based on this will have the following share.

Table 13 Calculated share of Imported commodity groups in 2007 E.C on the basis of 2002 E.C data set

Commodity Category	% of item for the given commodity group in year 2002 E.C	Calculated Truck traffic In year 2007 E.C
Gen.Cargo	53.08%	13,140
Petroleum	31.08%	7,635
Fertilizer	8.25%	2,027
Food Aid	7.59%	1,764
Total	100%	24,566

From table 7 selecting like previous 2002 E.C and 2004 E.C and calculate the annual growth factor like in the above Equation for the truck traffic count as if sharing with the commodity tonnage in the past five years.

From this it can easily be seen General cargo takes higher percent about 53.08% and this can tell us an activity based freight demand forecasting model is not expressive enough in case of forecasting the demand through Addis Ababa-Djibouti rail line this is because here in Ethiopia grouping of commodity groups is not common.

Table 14 calculated annual growth on the bases of imported economic variable given in billion birr taken from table 7.

Import Item	Value of item expressed as of billion birr in year 2002 E.C	For 2004 E.C	Annual Growth Factor
Industrial Supplies	36.5	76.5	1.4477
Capital goods and Accessories	36.5	63.1	1.315
Fuels and Lubricants	22.96	41.2	1.34
Food and Beverage	11.9	32.5	1.653
Transport Equipment and Accessories	18.8	34.5	1.355
Fertilizer	13.8	28.5	1.437
Others	0.335	1.07	1.788
Total	140.8	277.3	

The calculation is done in the same procedure like it has been done for an export.

$$AGF \text{ for each commodity} = \left(\frac{T_2}{T_1}\right)^{1/(Y_2-Y_1)}$$

Each commodity group stated are individual like by taking industrial supplies, capital goods and accessories, fuels and lubricants, food and beverage, transport equipment and accessories, fertilizer and others.

Table 15 Expected or forecasted truck traffic counts after 15 years relating historical and economic projection for imported goods.

Imported Commodity Group	Yearly truck traffic counts (2007 E.C.)	AGF	AGF ¹⁵	Truck traffic 2022
Gen. Cargo	13,140	1.333	74.551	979,601
Petroleum	7,635	1.34	80.644	615,717
Fertilizer	2,027	1.437	230.066	466,344
Food Aid	1,764	0.894	0.1863	329
Total	24,566			2,061,991

Considering the industrial supplies, capital goods, transport equipment and accessories, food and beverages and also by considering half of others as a Food and taking the cumulative as General Cargo and taking all this commodity groups and taking average approximately be $AGF \approx 1.333$

Here the average general cargo growth rate is then be 1.33

Table 16 Import Forecasted for major commodity groups with truck traffic

Year/ Commodity	Forecasted truck traffic counts for imported commodity groups			
	Gen. Cargo with AGF 1.333	Petroleum with AGF 1.34	Fertilizer with AGF 1.437	Food Aid with AGF 0.894
2007 E.C	13,140	7,635	2,027	1,764
2008 E.C	17,516	10,231	2,913	1,577
2009 E.C	23,348	13,709	4,186	1,410
2010 E.C	31,123	18,371	6,015	1,260
2011 E.C	41,487	24,617	8,643	1,127
2012 E.C	55,303	32,986	12,420	1,007
2013 E.C	73,719	44,202	17,848	901
2014 E.C	98,267	59,230	25,648	805
2015 E.C	130,990	79,368	36,856	720
2016 E.C	174,609	106,354	52,962	643
2017 E.C	232,276	142,514	76,106	575
2018 E.C	310,261	190,968	109,365	514
2019 E.C	413,578	255,897	157,157	460
2020 E.C	551,299	342,903	225,835	411
2021 E.C	734,882	459,490	324,525	367
2022 E.C	979,601	615,717	466,344	329

Forecasting using Regression Method

Here determination of an explanatory variables is such an important thing that should be considered. Hence many of this analysis fall for computational purpose of passenger demand forecasting there is one similarity that should take in to account; they both look for the calculation of the demand. In accessing trip generation passengers use House hold size and tries to relate with trips with a day. Similarly let the house hold size represent the volume in million tons of those import and export commodity groups.

First from the above table 6 and 8 let us prepare the commodity groups with their tonnage and forecast for the tonnage and share this tonnage with the traffic counts and prepare regression analysis. Calculating for the annual growth of the each commodity groups from the above table.

$$AGF = (\text{tonnage from the base year of 2004} / \text{tonnage from the base year of 2002})^{1/(2004-2002)}$$

And in order to calculate for the commodity group's tonnage

$$\text{Tonnage for each item to be forecasted} = \text{base year tonnage} * AGF^{\frac{1}{Y_2 - Y_1}}$$

$$T_{\text{forecasted}} = T_{2004} * (AGF)^{1/3} = T_{2004} * (AGF)^{0.333}$$

Where the tonnage to be forecasted is the forecast for each commodity group on year 2007 E.C and the base year considered here is the recent historical data of the commodity tonnage in year 2004 E.C

Table 17 Summary of imported and exported commodity groups tonnage on the basis of the recent data and calculating for 2007 E.C tonnage

Commodity/Year	2002 E.C	2004 E.C	AGF	AGF ³	2007 E.C	% share of item in year 2007 E.C
Coffee	320	403	1.122	1.4125	570	3.69%
Pulses	417	513	1.11	1.368	702	4.55%
Oilseeds	355	408	1.08	1.26	514	3.33%
Meat/Hides	27	34	1.123	1.4163	49	0.32%
Fruits/Vegetables	48	59	1.11	1.368	81	0.52%
Others	153	205	1.16	1.561	320	2.07%
General cargo	3,639	4,964	1.17	1.602	7,953	51.51%
Petroleum	2,131	2,648	1.118	1.4	3,707	24%
Fertilizer	566	679	1.09	1.3	883	5.72%
Food Aid	520	569	1.05	1.16	660	4.29%
Total	8,176	10,482			15,439	100%

And for this all major commodity groups forecasting is done with their respective AGF; for the commodity groups shown below representation is done with their AGF; let A-Coffee with AGF=1.122; B-Pulses with AGF=1.11; C-Oilseeds with AGF=1.08; D-Meat/Hides with AGF=1.123; E-Fruits and Vegetables with AGF=1.11; F-others with 1.16; G-General cargo with AGF=1.17; H- Petroleum with AGF=1.118; I –Fertilizer with AGF=1.09; J-Food Aid with AGF=1.05.

Table 18 Forecasted import and export for major commodity groups

Year/ Commodity	Expressed in million ton (Approximated value)									
	A	B	C	D	E	F	G	H	I	J
2007 E.C	570	702	514	49	81	320	7,953	3,707	883	660
2008	639	779	555	55	90	371	9,305	4,144	962	693
2009	718	865	600	62	100	431	10,886	4,633	1,049	728
2010	805	960	647	69	111	499	12,738	5,180	1,144	764
2011	903	1,066	699	78	123	579	14,903	5,791	1,246	802
2012	1,014	1,183	755	87	136	672	17,437	6,475	1,359	842
2013	1,137	1,313	816	98	152	780	20,401	7,239	1,481	884
2014	1,276	1,457	881	110	168	904	23,869	8,093	1,614	929
2015	1,432	1,618	951	124	187	1,049	27,927	9,048	1,760	975
2016	1,606	1,796	1,027	139	207	1,217	32,674	10,116	1,918	1,024
2017	1,802	1,993	1,110	156	230	1,412	38,229	11,309	2,091	1,075
2018	2,022	2,213	1,198	176	255	1,638	44,728	12,644	2,279	1,129
2019	2,269	2,456	1,294	197	283	1,900	52,331	14,136	2,484	1,185
2020	2,546	2,726	1,398	231	315	2,203	61,228	15,804	2,707	1,245
2021	2,856	3,026	1,509	249	349	2,556	71,636	17,669	2,951	1,307
2022	3205	3,359	1,631	280	388	2,965	83,815	19,754	3,217	1,373

Once the share of each commodity group has determined in 2007 E.C, and since there is an estimated or calculated value of truck traffic for imports and exports of those commodity group

developing a table for this is an easy task so linear integrity and development of an equation is then can be followed.

Table 19 regression table showing commodity share in million ton with respect to annual truck traffic flow for imports and exports (2007 E.C)

X	% Commodity share (VMT) (x)									
	A 3.69	B 4.55	C 3.33	D 0.32	E 0.52	F 2.07	G 51.51	H 24	I 2.72	J 4.29
Annual Truck traffic flow Import(y)	0	0	0	0	0	0	13,140	7,635	2,027	1,764
For an export(y)	5,742	7,843	6,349	484	860	2,748	0	0	0	0
$\sum y$	5,742	7,843	6,349	484	860	2,748	13,140	7,635	2,027	1,764

The equation will have a linear form with the equation $y = a + bx$ where y tells the annual trip rate, and x tells the tonnage of imported and exported commodities.

From this relation the next step is determination of the variables a and b.

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - \sum x \sum x}$$

$$b = \frac{20 * 9574.6864 - 2 * 48,592}{20 * 0.661 - 2 * 2}$$

$$b = (94,309.728 / 9.22) = 10,228.742$$

$$a = 2,429.6 - 10,228.742 * 0.1$$

$$a = 1,406.7258$$

$$y = 1,406.7258 + 10,228.742 * x$$

Taking the average share of commodity group as 1.0

$$y = 1,406.7258 + 10,228.742 * 1.0$$

$$y = 11,635.4678 \text{ expected annual trips.}$$

Even though the distribution is not well defined taking the average of the annual growth factor for each commodity groups.

$$(1.122+1.11+1.08+1.123+1.11+1.16+1.17+1.118+1.09+1.05)/10=1.1133$$

Table 20 Expected annual forecasted trips by regression analysis

Year/expected forecast	Expected forecast relating with AGF 1.113
2007 E.C	11,636
2008 E.C	12,950
2009 E.C	14,415
2010 E.C	16,043
2011 E.C	17,856
2012 E.C	19,874
2013 E.C	22,120
2014 E.C	24,619
2015 E.C	27,401
2016 E.C	30,498
2017 E.C	33,944
2018 E.C	37,780
2019 E.C	42,049
2020 E.C	46,800
2021 E.C	52,088
2022 E.C	57,974

4.2. COMPARISON OF THE GROWTH FACTOR MODEL AND REGRESSION MODEL

As seen in the above, forecasting using growth factor and regression has been done with the sample data that has been taken from different sectors and according to the result forecasting using regression analysis seems better than that of growth factor.

In both cases they use past historical data and need to forecast for the future. The result forecasting with growth factor seem over estimate, it can easily be seen from the tables that the growth factor method calculate and begin with annual growth factor which comes from the traffic counts along the path and the problem is in getting past historical data which is almost a time series model, the future thinks as if the past historical data will continue as it is which is the demand of the import-export will be in the same manner, the economic growth of the country will be in the same situation and the like.

In forecasting using regression analysis actually there are many regression parts which Zonal-based multiple regression analysis is advisable for freight transport the traffic count estimation seems quite good because it tries to relate the major commodity groups expressed in terms of tonnage which hopefully will allocate the demand of each products will grow independently and come to be forecasted as an indicator of the study.

The analysis in linear regression method will actually be more realistic if the variables that relate them is more a function of many variables that is if zoning with a commodity group, type of the firm and growth factor for analytic purpose.

In real cases the elasticity's of freight demand forecasting with respect to economic activity (GDP) vary across different commodity group which indicates different sensitivity variations in agricultural and industrial processing companies as different commodities have different transport demand and studying each estimation reflects particular circumstances for each commodity group.

Generally, if rich data is available regression model can be able to generate best results in forecasting, while growth rate model are more accurate for different sensitive elasticity and for similar growth factor of the data that has been taken.

MODAL SPLIT

Modal split represents the ratio of different transport mode in the total journey from origin (O) to the destination (D).

In analyzing a modal split or a mode choice; multi nominal mode choice or MNL is used for freight transport. The equation for this can be written as;

$$P_{ij}^1 = \frac{e^{-\beta c_{ij}^1}}{\sum e^{-\beta c_{ij}^m}}$$

Cost of travel by car

$$c_{ij} = a_1 t_{ij}^v + a_2 t_{ij}^w + a_3 t_{ij}^t + a_4 t_{nij} + a_5 F_{ij} + a_6 \phi_j + \delta$$

Where t_{ij}^v is the in vehicle travel time between i and j, t_{ij}^w is the walking time to and from stops, t_{ij}^t is the waiting time at stops, F_{ij} is the fair charged between i and j, ϕ_j is the parking cost, δ is a parameter representing comfort and convenience. If the travel cost is low, then that mode has more probability of being chosen.

Let there exists two modes ($m=1,2$) then the portion of trips by mode 1 from zone i to zone j is (P_{ij}^1) let c_{ij}^1 be the cost of traveling from zone i to zone j using mode 1, c_{ij}^2 be the cost of traveling from zone i to zone j by mode 2, there re three cases:

1. If $c_{ij}^2 - c_{ij}^1$ is positive, then mode 1 is chosen.
2. If $c_{ij}^2 - c_{ij}^1$ is negative, then mode 2 is chosen.
3. If $c_{ij}^2 = c_{ij}^1=0$, then both modes have equal probability.

In choosing a mode for Addis Ababa-Djibouti freight transport demand hence existing trips or journeys are with vehicle transport and in the allocation of this mode, which is in introducing this mode let the trip has the following multinomial characteristics.

Let the number of annual trips from origin to destination or vice versa be taken 57,974 which is from the regression analysis. And assumption has been taken there is no walking time in between Addis Ababa to Djibouti, the waiting time at the stops be waiting time at the port and dry ports which is 15 days and an assumption is also made the stay at port, the fair charged between origin and destination, the parking cost become almost half in ration with each other (vehicle and train).

Table 21 Multinomial Trip Characteristics

	t_{ij}^v	t_{ij}^w	t_{ij}^t	F_{ij}	ϕ_{ij}
Coefficient	0.03	0.04	0.07	0.2	0.2
Vehicle	2.25 days	-	14 days	2	8
Train	1 day	-	7 days	1	5

Considering introduction to the train

Cost of travel by car (c_{car}) = $0.03*2.25+0.07*15+0.2*10+0.2*10=3.0475$

Cost of travel by train $c_{train}=0.03*1.25+0.07*7+0.2*1+0.2*5=1.72$

Probability of choosing a car = $\frac{e^{-3.04175}}{e^{-3.0475}+e^{-1.72}}=0.21$

Probability of choosing a train = $\frac{e^{-1.72}}{e^{-3.0475}+e^{-1.72}}=0.79$

Trips carried by each mode in 2022 E.C will then be estimated as:

$$T_{ij}^{car} = 12,175 \text{ trips}$$

$$T_{ij}^{train} = 45,799 \text{ trips}$$

As it can easily be seen from the figure car share takes 12,175 trips and train takes 45,799 trips which indicates that train attract more trips.

4.3. COMPARISON OF RESULT WITH THE MODEL FOECASTED BY ERC

Table 22 Total potential freight flows (expressed in ton)

Total	Addis Ababa-Debrezeit	Debrezeit-Nazerit	Nazerit-Meiso	Meiso-Error	Error-Diredawa	Diredawa-Dewale	Sum Total
2022	1,751,870	1,198,986	385,785	89,446	74,288	82,171	3,582,546

Source: Ethiopian Railway Corporation (ERC)

And from this an average of truck trip has been calculated that on average a single truck may handle 30 tons so that the number of trips may then be:

Number of trips= $(3,582,546/30) = 119,419$ trips

Four stage aggregate model has been conducted by ERC in order to forecast the number of trips between Addis Ababa-Djibouti and found the above figure.

This thesis also uses a four stage aggregate model as a basics and hybrid this model with its own parameters. Here the traffic count is the basic problem and with this count, tonnage share of each commodity on a desired base year and annual growth factor of economic expense given for all commodity groups and make relations.

Using this hybrid model the result has been found to be 2,061,991 trips for an import items and 1,289,415 trips for export items which is an over estimation.

An expected trip from regression analysis relating with AGF in 2022 E.C. is 57,974 trips. So a trip generated with a regression analysis relating with AGF which is a hybrid model give comparable trip the ERC forecast.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1. CONCLUSION

Traffic congestion problem for all transport mechanism including freight begins with an appropriate planning of demand forecasting models. In order to solve this problem different countries use different kinds of forecasting approach; among these models the major ones include time series model, econometric model, and aggregate and disaggregate model, behavioral or activity based model and others and even they also try to find out a new approach or hybrid among the approaches.

In this thesis the problem or the need to study about models and forecasting using four stage aggregate model begins with and planned to solve traffic congestion and to shorten the travel time from origin to destination. Throughout this thesis it has been argued among major forecasting models, stated with their strengths and weakness and from those models four stage aggregate model has been selected due to stated preference (that is aggregation across options, such as origin and destination choice model on zonal data). For example activity based model or behavioral model is ignored because there is poor grouping of commodities in Ethiopia, and time series model is ignored because it is seasonal and can easily be affected reliability of data, influence of economy over all the world, advisable to work out for short terms rather here four stage aggregate model (hybrid) relating the growth rate of economic activity and traffic count seems advantageous and with this model forecasting has been done.

As stated above, consideration of traffic congestion and time delay as a problem trip generated with annual growth factor and linear regression analysis for the coming 15 years and mode split has also be done.

When using the growth rate in order to generate a trip, it estimates over about 3,351,406 trips and when applying with a linear regression analysis it estimates nearly half of the one forecasted by ERC and found 57,974 and 119,419 trips. According to this analysis from the total trips generated

used trains took about 79 percent of mode share and the remaining is given for truck. In this thesis the new thing is forecasting with annual growth factor relate not only the economic background but it also try to relate the commodity tonnage share, economic growth rate and also the traffic counts.

Result comparison between the forecasted result by ERC (four-stage aggregate model) and the model used in this thesis work (four-stage aggregate but hybrid model) and found comparable result.

Generally as it can be tried to show here in the above sessions forecasting with four stage aggregate model by incorporating traffic count, commodity share and expense annual growth rate given for each imported and exported commodity is advantageous in Ethiopia this is because the need is resolving the traffic problem and the travel time between origin and destination.

5.2. RECOMMENDATION

In studying freight demand model especially in the case of Addis Ababa-Djibouti rail line. The following g things are recommended:

- ✚ Since Ethiopia is a land locked country and many commodity groups are moved through Djibouti port therefore taking past historical traffic count data as an issue for trip generation purpose like as it is used in this thesis is advantageous because the problem begins from traffic congestion and the time used to move goods from origin-destination. Therefore ERC should take this model in to account.
- ✚ The commodity tonnage share increase and decrease tells the demand of that item therefore ERC should also consider the annual growth of the tonnage rate.
- ✚ ERC should consider instead of using collective grouping of commodity group as used in ETA for the traffic count, using of each commodity (for example petroleum, coffee, meet/hides, pulses and other) can give better forecast and even be helpful to allocate other mode of transport because items of the same property can easily be handled together.

5.3. FUTURE WORK

As part of this thesis the trip generated with annual growth factor and regression has been made but for further expansion of a better neural network between regions or along the path (a trip other than Addis Ababa-Djibouti as origin-destination) distribution has been made. Therefore a real data of a traffic count along the path which doesn't tell about origins and destinations only may give more accurate forecast to estimate the demand along the corridor.

Bibliography

1. Cambridge Systematics, I. (Sep 1996). *Quick Response Freight Manual*. USA: University of Wisconsin-Milwaukee.
- 2.Center for Urban Transportation, S. (March 1999). *Statewide Travel Forecasting*. USA: University of Wisconsin-Milwaukee.
- 3.Chatfield, C. (1996). *The Analysis of Time Series*. Boca Raton, FL: Champ and Hall.
- 4.Ehsan, D., Virginia, P., Sisiopiku, Michael, D., Anderson, Mehrnaz, D., & Andrew, S. (2016). Comparison of Freight Demand Forecasting Models. *International Journal of Traffic and Transportation Engineering*, 19-26.
- 5.Ertander, S., & Stewart, N. (1990). The Gravity model in transportation analysis: theory and extensions.
- 6.Garrido, Amelia, C., & Regan. (2002). Modeling freight demand and shipper behaviour: State of Art, future directions. *Forthcoming in the Leading Edge of Travel Behaviour Research*, 3-8.
- 7.Gerard, d. j., Arnand, B., Lori, T., Robert, V., bok, M. d., & Nora, S. (2011). Distribution and Modal split models for freight transport. *Netherlands organization for applied scientific research*.
- 8.Gor.Ravi, M. (2009). Forecasting Techniques. In M. Gor.Ravi, *Industrial Statistics and Operational Management* (pp. 142-143). New Delhi: ICAFAI business school.
- 9.Ilveto, Titus, A., & Tom. (2004). Introduction to time series forecasting.
- 10.Jose Castiglione, M. B., & Glielie, J. (2015). Activity based travel demand models. 3-7.
- 11.Joseph, Y., Chow, Heon, C., Yang, & C.Regan, A. (2010). State-of-the-art of Freight Forecast Modeling. *Transportation*.
- 12.Michael W., B., Xiaohua, L., & Jerry, N. (1999). Time series forecasting of quarterly railroad grain car loadings. *Transportation research paper*, 43-57.

13. Mirajam H.E., I., Wilhem, J., & Lori, A. (2002). Freight trip generation by firms. *European Congress of the Regional Science Association*.
14. Nigussie, S. S. (November 10, 2004). *The Ethiopian foreign trade statistics*. Addis Ababa.
15. Ortuzar, J. d., & Willumsen, L. (2011). *Modelling Transport*. London: John Wiley and sons Ltd.
16. Richard, P. (2008). Retrieved from ([http://www.google.com/ethio-franco railway and its history/](http://www.google.com/ethio-franco_railway_and_its_history/) image/html: www.google.com)
17. Richard, P. (n.d.). *Journal of Franco-Ethiopian Railway and Its History*.
18. Romeo, Lucia, R., & Danielis. (1999). Analysis Freight Transport Demand Using Stated Preference data a Survey and a research Project for the Friuli-Venezia Giulia Region. *Transporti Europei*, 2-4.
19. Seada. (2016, April 21). how many trucks are using this lines, . (A. Arefeayne, Interviewer)
20. Sillano, M., & Ortuzar, J. (2005). Willing to pay estimation with mixed logit model, some new evidence. *Environment Planning*.
21. Tabitha, M., K. R., Caris, A., Tom, B., & Gerrit K., J. (2011). Creating an Innovative Activity-Based Freight Transportation Framework. 1-6.
22. White Legg, J. (1993). *Transport for sustainable Future: the case for europe*. London : Belhaven press.

Appendix 1 Summary of trade volume in tons: (Ethiopia Commodity Exchange).2016

Year	Trade Volume in tons					
	Coffee	Green Mung Bean	Haricot Beans	Maize	Sesame	Wheat
2012	255,132		80,030	379	310,204	
2013	215,120		75,736		197,382	32
2014	242,652	146	50,747		281,974	40
2014	252,886		41,435		333,216	
2016*	43,464					

**Appendix 2 Total cargo lifted by shipping sector by trade type and vessel type:
(Ethiopian Shipping and Logistic Enterprise.2015**

Year	Import	Export	Cross Trade	Own Vessel	Chattered Vessel
2000	1,867,692	3,527	353,429	763,498	1,461,150
2001	1,953,492	4,974	213,473	699,793	1,472,146
2002	2,170,353	1,138	305,887	841,375	1,636,003
2003	1,739,571	1,222	466,987	787,906	1,419,874
2004	3,062,330	492	202,062	686,404	2,578,480
2005	3,018,969	1,082	75,000	655,498	2,439,553
2006	2,767,053	217,132	44,5704	1,767,649	1,662,240
2007	3,340,135	10,706	306,214	1,315,517	2,341,538

Appendix 3. Estimate of transit and transport cost of import and export percent of import and export value: the management of commercial road transport in Ethiopia (ERA). 2014

Year	Volume(Ton)		Value (birr in million)		Transit and Transport cost (birr in million)		Total cost (birr in million)		%	
	Import	Export	Import	Export	Import	Export	Import	Export	Import	Export
2001	3373380	306522	1.371	2.685	3234397	1.054	1.695	2.791	19.08	3.78
2002	2565303	524535	1.199	3.171	2459613	1.804	1.446	3.351	17.01	5.39
2003	5074223	413195	2.091	3.574	4865165	1.421	2.577	3.716	18.88	3.83
2004	3823023	518805	2.329	4.207	3665515	1.785	2.695	4.385	13.60	4.07
2005	4532482	690938	3.123	6.645	4345753	2.377	3.562	6.883	12.20	3.45

Note: At constant transit and transport cost of birr 344.00/ ton for an export and at constant transit and transport cost for 20' containers birr 958.00 for import.

Appendix 4 Number of Vehicles along Addis Ababa-Djibouti path with their life period from the day in use and carrying tonnage and also level given by the authority: Ethiopian Transport Authority.2016

Level of the vehicle	Age≤10		Age 10.1-20		Age >20		Total
	30-40 Ton	20-29.9 Ton	30-40 Ton	20-29.9 Ton	30-40 ton	20-29.9 Ton	
1A	4124	150	157	215	145	64	4855
2A	-	642	1497	133	207	-	2479
3A	331	348	469	544	234	127	2053
3B	-	114	-	-	92	313	519
4	120	21	19	74	-	-	234
Total	4575	1275	2142	966	678	504	10140

Appendix 5. GDP by Economic activity at constant prices and growth rate of GDP by Economic activity at constant prices; Ministry of Finance and Economic Development.2015

Industry/Year	2000	2001	2002	2003	2004
Agriculture, Hunting and Forestry	101,956,604 (7.5)	180,984,518 (6.4)	194,797,280 (7.6)	212,360,824 (9.0)	222,803,730 (4.9)
Fishing	159,289 (34.0)	201,543 (26.5)	204,874 (1.7)	217,024 (5.9)	263,203 (21.3)
Mining and quarrying	2,655,503 (21.4)	2,994,775 (12.8)	4,317,580 (44.2)	6,806,661 (57.7)	7,675,101 (12.7)
Manufacturing	13,215,282 (10.3)	14,414,606 (9.1)	16,066,186 (11.6)	18,037,324 (12.1)	20,503,794 (13.7)
Services	228,750,454 (14.325)	183,056,016 (12.967)	206,464,039 (11.208)	232,445,631 (12.325)	258,905,668 (12.658)
Total	346,737,132 (11.4)	381,651,458 (10.1)	421,849,959 (10.5)	469,867,464 (11.4)	510,151,496 (8.6)
Less FISIM	2,393,696 (30.00)	2,693,718 (12.5)	2,898,065 (7.6)	3,219,328 (11.1)	3,618,229 (12.4)
GDP at constant market prices	371,716,667 (10.8)	404,436,976 (8.8)	455,196,015 (12.6)	506,079,135 (11.2)	548,921,587 (8.5)