A Reference Manual for "Road Transport Management"

Bibek Magar

April 23, 2024

Contents

Chapter 1

Transportation System

1.1 Introduction

Transportation is a crucial aspect of a nation's development and prosperity. However, with the advent of modern means of transportation, managing its operation has become more complex, so there is ample opportunity for engineers to devise new operational and management techniques for the smooth control of various modes of transportation of a country.

1.2 History of Transportation

Early Boats

The first mode of transportation was created in the effort to traverse water: boats. Those who colonized Australia roughly 60,000–40,000 years ago have been credited as the first people to cross the sea, though there is some evidence that seafaring trips were carried out as far back as 900,000 years ago.

The earliest known boats were simple logboats, also referred to as dugouts, which were made by hollowing out a tree trunk. Evidence for these floating vehicles comes from artifacts that date back to around 10,000–7,000 years ago. The Pesse canoe—a logboat—is the oldest boat unearthed and dates as far back as 7600 BCE. Rafts have been around nearly as long, with artifacts showing them in use for at least 8,000 years.

Horses and Wheeled Vehicles

Next, came horses. While it's difficult to pinpoint exactly when humans first began domesticating them as a means of getting around and transporting goods, experts generally go by the emergence of certain human biological and cultural markers that indicate when such practices started to take place.

Based on changes in teeth records, butchering activities, shifts in settlement patterns, and historic depictions, experts believe that domestication took place around 4000 BCE. Genetic evidence from horses, including changes in musculature and cognitive function, support this.

It was also roughly around this period that the wheel was invented. Archaeological records show that the first wheeled vehicles were in use around 3500 BCE, with evidence of the existence of such contraptions found in Mesopotamia, the Northern Caucuses, and Central Europe. The earliest well-dated artifact from that time period is the "Bronocice pot," a ceramic vase that depicts a four-wheeled wagon that featured two axles. It was unearthed in southern Poland.

Steam Engines

In 1769, the Watt steam engine changed everything. Boats were among the first to take advantage of steam-generated power; in 1783, a French inventor by the name of Claude de Jouffroy built the "Pyroscaphe," the world's first steamship. But despite successfully making trips up and down the river and carrying passengers as part of a demonstration, there wasn't enough interest to fund further development.

While other inventors tried to make steamships that were practical enough for mass transport, it was American Robert Fulton who furthered the technology to where it was commercially viable. In 1807, the Clermont completed a 150-mile trip from New York City to Albany that took 32 hours, with the average speed clocking in at about five miles per hour. Within a few years, Fulton and company would offer regular passenger and freight service between New Orleans, Louisiana, and Natchez, Mississippi.

Back in 1769, another Frenchman named Nicolas Joseph Cugnot attempted to adapt steam engine technology to a road vehicle—the result was the invention of the first automobile. However, the heavy engine added so much weight to the vehicle that it wasn't practical. It had a top speed of 2.5 miles per hour.

Another effort to repurpose the steam engine for a different means of personal transport resulted in the "Roper Steam Velocipede." Developed in 1867, the two-wheeled steam-powered bicycle is considered by many historians to be the world's first motorcycle.

Locomotives

One mode of land transport powered by a steam engine that did go mainstream was the locomotive. In 1801, British inventor Richard Trevithick unveiled the world's first road locomotive—called the “Puffing Devil”—and used it to give six passengers a ride to a nearby village. It was three years later that Trevithick first demonstrated a locomotive that ran on rails, and another one that hauled 10 tons of iron to the community of Penydarren, Wales, to a small village called Abercynon.

It took a fellow Brit—a civil and mechanical engineer named George Stephenson—to turn locomotives into a form of mass transport. In 1812, Matthew Murray of Holbeck designed and built the first commercially successful steam locomotive, “The Salamanca,” and Stephenson wanted to take the technology a step further. So in 1814, Stephenson designed the ”Blücher,” an eight-wagon locomotive capable of hauling 30 tons of coal uphill at a speed of four miles per hour.

By 1824, Stephenson improved the efficiency of his locomotive designs to where he was commissioned by the Stockton and Darlington Railway to build the first steam locomotive to carry passengers on a public rail line, the aptly named ”Locomotion No. 1.” Six years later, he opened the Liverpool and Manchester Railway, the first public inter-city railway line serviced by steam locomotives. His notable accomplishments also include establishing the standard for rail spacing for most of the railways in use today. No wonder he’s been hailed as ”Father of Railways.”

Submarines

Technically speaking, the first navigable submarine was invented in 1620 by Dutchman Cornelis Drebbel. Built for the English Royal Navy, Drebbel’s submarine could stay submerged for up to three hours and was propelled by oars. However, the submarine was never used in combat, and it wasn’t until the turn of the 20th century that designs leading to practical and widely used submersible vehicles were realized.

Along the way, there were important milestones such as the launch of the hand-powered, egg-shaped ”Turtle” in 1776, the first military submarine used in combat. There was also the French Navy submarine ”Plongeur,” the first mechanically powered submarine.

Finally, in 1888, the Spanish Navy launched the ”Peral,” the first electric, battery-powered submarine, which also so happened to be the first fully capable military submarine. Built by a Spanish engineer and sailor named Isaac Peral, it was equipped with a torpedo tube, two torpedoes, an air regeneration system, and the first fully reliable underwater navigation system, and it posted an underwater speed of 3.5 miles per hour.

Aircraft

The start of the twentieth century was truly the dawn of a new era in the history of transportation as two American brothers, Orville and Wilbur Wright, pulled off the first official powered flight in 1903. In essence, they invented the world's first airplane. Transport via aircraft took off from there with airplanes being put into service within a few short years during World War I. In 1919, British aviators John Alcock and Arthur Brown completed the first transatlantic flight, crossing from Canada to Ireland. The same year, passengers were able to fly internationally for the first time.

Around the same time that the Wright brothers were taking flight, French inventor Paul Cornu started developing a rotorcraft. And on November 13, 1907, his "Cornu" helicopter, made of little more than some tubing, an engine, and rotary wings, achieved a lift height of about one foot while staying airborne for about 20 seconds. With that, Cornu would lay claim to having piloted the first helicopter flight.

1.3 Transportation Sector and Modes of Transportation

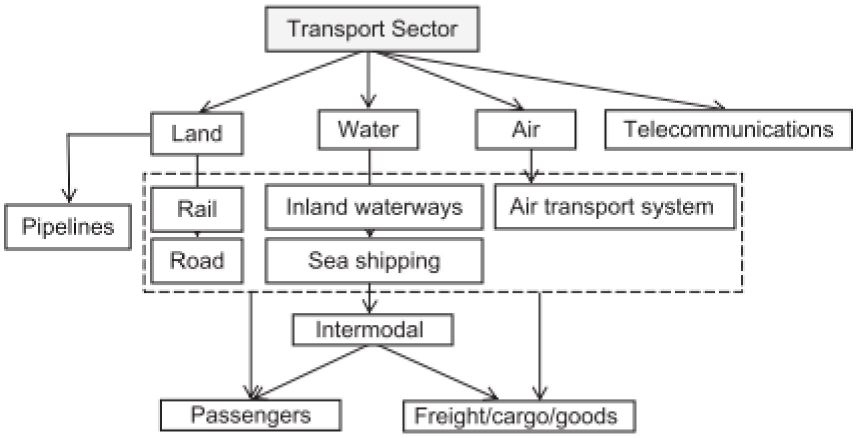
1.3.1 Transportation Sector

Transportation sector performs transportation services in order to satisfy demand for mobility of people and transport of freight shipments.A range of socioeconomic activities in a society, as well as land uses induce transportation demand. The demand is represented by the number of passengers and/or volumes of cargo to be transported between given origins and destinations during a given (specified) period of time. The supply component in the transportation sector consists of transport services provided by different transport modes and their particular systems. Airports, highways, streets, and ports should be able to meet transportation demand and offer acceptable level-of-service to the users. The transport supply component contributes to the economy of a region, country, and continent it serves. The demand and supply component are in permanent interaction.

1.3.2 Modes of Transportation

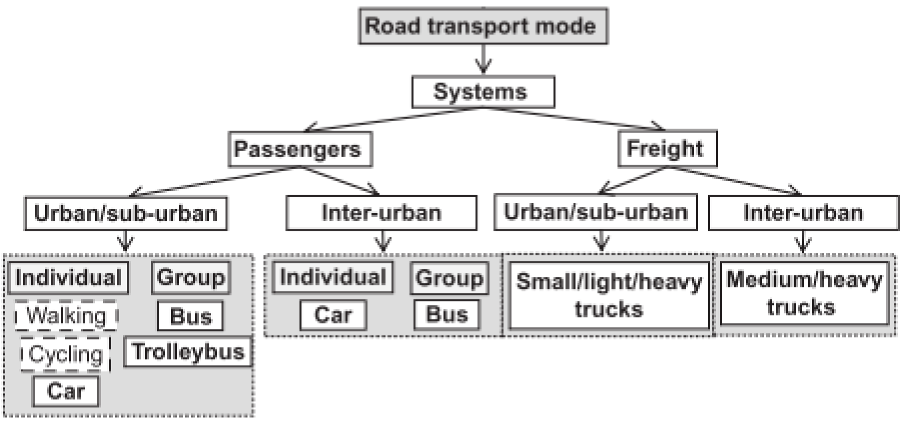
The transportation modes constituting the supply component are generally classified according to the way of performing their operations of transporting people and freight shipments. In general, the basic land-based transport modes include road, rail, and pipeline. The water-based mode includes inland waterways and sea shipping. The air transport is the air-based mode.

The specific mode is intermodal transport consisting of combinations of particular basic modes and their systems. In addition, the mode not carrying out physical entities but just information is telecommunications. Fig. shows a simplified scheme of the structure of transport sector, its modes and their systems.



Transport modes generally consist of two types of systems: the first is intended for serving passenger and the other for serving freight/cargo/goods demand. Each of these consists of subsystems, which will be called systems. The classification of systems within each mode is carried out at three levels: (i) type of the system (passengers, freight), spatial scale of operation (urban/suburban/regional, interurban), and carrier type (individual, group).

The Fig. below shows classification of systems of the road transport mode.



1.4 Factors influencing Transport Operations

There are mainly four components of road transportation: the driver, the pedestrian, the vehicle, and the road, which influence road transport management's design and operational considerations. To provide efficient and safe road transportation, a knowledge of the characteristics and limitations of each of these components is essential. It is also crucial to be aware of the interrelationships among these components to determine the effects they have on each other. Their characteristics are also of primary importance when traffic engineering measures such as traffic control devices are used in the highway mode.

1.4.1 Driver Characteristics

One problem that faces traffic and transportation engineers when they consider driver characteristics in the course of design is the varying skills and perceptual abilities of drivers on the highway, demonstrated by a wide range of capabilities to hear, see, evaluate, and react to information. Studies support that these abilities vary under different conditions, such as the influence of alcohol, fatigue, and time of the day. Therefore, it is essential that criteria used for design purposes be compatible with the capabilities and limitations of most drivers on the highway. The use of an average value, such as mean reaction time, may not be adequate for a large number of drivers. The 85th percentile and the 95th percentile have been used to select design criteria; in general, the higher the chosen percentile, the wider the range covered.

The Human Response Process

Actions taken by the drivers while driving result from their evaluation and reaction to the information they obtain from their certain stimuli. However, assessment and response must be carried out quickly, as the information received while driving the highways is dynamic. Furthermore, it has been suggested that most of the information received by a driver is visual, implying that the ability to see is of fundamental importance in the driving task. Therefore, highway and traffic engineers must have some rudimentary knowledge of visual perception and hearing perception.

Visual Reception : The principal characteristics of the eye are visual acuity, peripheral vision, color vision, glare vision and recovery, and depth perception.

1. Visual Acuity : It is the ability to see fine details of an object. It can be represented by the visual angle, which is the reciprocal of the smallest pattern detail in minutes of arc that can be resolved and given as:

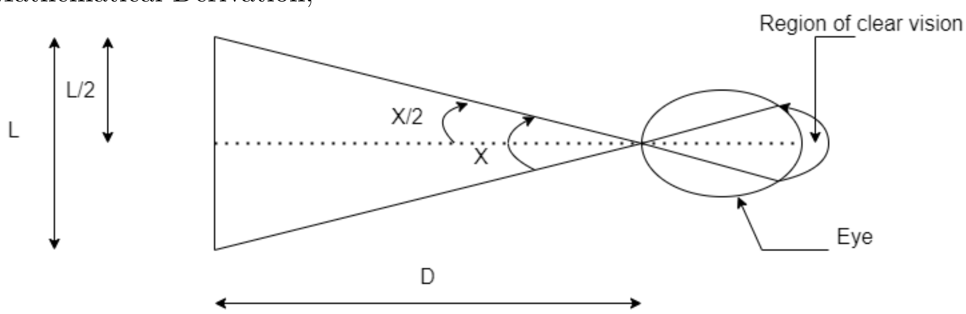
$$\phi = 2\arctan\left(\frac{L}{2D}\right) \tag{1.1}$$

Where,

L = diameter of the target (letter or symbol)

D = distance from the eye to target in the same units as L

Mathematical Derivation,



Using trigonometry,

$$\tan\left(\frac{X}{2}\right) = \frac{\frac{L}{2}}{D}$$
$$\tan\left(\frac{X}{2}\right) = \frac{L}{2D}$$
$$X = 2\arctan\left(\frac{L}{2D}\right)$$

The ability to resolve a pattern detail with a visual acuity of one minute of arc (1/60 of a degree) is considered as normal vision of acuity (20/20 vision).

**20/20 vision is a term used to express normal visual acuity measured at a distance of 20 feet. If you have 20/20 vision, you can see clearly at 20 ft what should normally be seen at that distance. If you have 20/100 vision, it means you must be as close as 20ft to see what a person with normal vision can see at 100ft.*

2. Peripheral Vision : It is the ability of people to see objects beyond the cone of clearest vision. Although objects can be seen within this zone, details and color are not clear. The cone for peripheral vision could be one subtending up to 160 degrees; this value is affected by the speed of the vehicle. Age also influences peripheral vision. For instance, at about age 60, a significant change occurs in a person's peripheral vision.

3. Color Vision : It is the ability to differentiate one color from another, but deficiency in this ability, usually referred to as color blindness, is not of great significance in highway driving because other ways of recognizing traffic information devices (e.g., shape) can compensate for it. Combinations of black and white and black and yellow have been shown to be those to which the eye is most sensitive.

4. Glare Vision and Recovery : There are two types of glare vision: direct and specular. Rowland and others have indicated that direct glare occurs when relatively bright light appears in the individual's field of vision and specular glare occurs when the image reflected by the relatively bright light appears in the field of vision. Both types of glare result in a decrease of visibility and cause discomfort to the eyes. It is also known that age has a significant effect on the sensitivity to glare, and that at about age 40, a significant change occurs in a person's sensitivity to glare.

The time required by a person to recover from the effects of glare after passing the light source is known as glare recovery. Studies have shown that this time is about 3 seconds when moving from dark to light and can be 6 seconds or more when moving from light to dark. Glare vision is of great importance during night driving; it contributes to the problem of serving older people, who see much more poorly at night. This phenomenon should be taken into account in the design and location of street lighting so that glare effects are reduced to a minimum.

Glare effects can be minimized by reducing luminaire brightness and by increasing the background brightness in a driver's field of view. Specific actions taken to achieve this in lighting design include using higher mounting heights, positioning lighting supports farther away from the highway, and restricting the light from the luminaire to obtain minimum interference with the visibility of the driver.

5. Depth Perception : It affects the ability of a person to estimate speed and distance. It is particularly important on two-lane highways during passing maneuvers, when head-on crashes may result from a lack of proper judgment of speed and distance.

The ability of the human eye to differentiate between objects is fundamental to this phenomenon. It should be noted, however, that the human eye is not very good at estimating absolute values of speed, distance, size, and acceleration. This is why traffic control devices are standard in size, shape, and color. Standardization not only aids in distance estimation but also helps the color-blind driver to identify signs.

Perception-Reaction Process

The process through which a driver, cyclist, or pedestrian evaluates and reacts to a stimulus can be divided into four subprocesses:

1. Perception: the driver sees a control device, warning sign, or object on the road

2. Identification: the driver identifies the object or control device and thus understands the stimulus

3. Emotion: the driver decides what action to take in response to the stimulus; for example, to step on the brake pedal, to pass, to swerve, or to change lanes

4. Reaction or Volition: the driver actually executes the action decided on during the emotion sub-process

Time elapses during each of these subprocesses. The time that elapses from the start of perception to the end of reaction is the total time required for perception, identification, emotion, and volition, sometimes referred to as PIEV time or (more commonly) as perception-reaction time.

Perception-reaction time is an important factor in the determination of braking distances, which in turn dictates the minimum sight distance required on a highway and the length of the yellow phase at a signalized intersection. Perception-reaction time varies among individuals and may, in fact, vary for the same person as the occasion changes. These changes in perception-reaction time depend on how complicated the situation is, the existing environmental conditions, age, whether the person is tired or under the influence of drugs and/or alcohol, and whether the stimulus is expected or unexpected.

Triggs and Harris described this phenomenon in detail. They noted that the 85th-percentile time to brake, obtained from several situations, varied from 1.26 to over 3 seconds. The reaction time selected for design purposes should, however, be large enough to include reaction times for most drivers using the highways. Recommendations made by the American Association of State Highway and Transportation Officials (AASHTO) stipulate 2.5 seconds for stopping-sight distances. This encompasses the decision times for about 90 percent of drivers under most highway conditions. Note, however, that a reaction time of 2.5 second may not be adequate for unexpected conditions or for some very complex conditions, such as those at multiphase at-grade intersections and ramp terminals. For example, when signals are unexpected, reaction times can increase by 35 percent.

Numerical Example:

A driver with a perception-reaction time of 2.5 sec is driving at 65 mi/h when she observes that an accident has blocked the road ahead. Determine the distance the vehicle would move before the driver could activate the brakes. The vehicle will continue to move at 65 mi/h during the perception-reaction time of 2.5 sec.

Solution:

- Convert mi/h to ft/sec:

$$65 \text{ mi/h} = \left(65 \times \frac{5280}{3600}\right) \text{ft/sec} = 65 \times 1.47 = 95.55 \text{ ft/sec}$$

- Find the distance traveled:

$$\begin{aligned} D &= vt \\ &= 95.55 \times 2.5 \\ &= 238.9 \text{ ft} \end{aligned}$$

where v = velocity and t = time.

1.4.2 Vehicle Characteristics

Criteria for the geometric design of highways are partly based on the static, kinematic, and dynamic characteristics of vehicles. Static characteristics include the weight and size of the vehicle, while kinematic characteristics involve the motion of the vehicle without considering the forces that cause the motion. Dynamic characteristics involve the forces that cause the motion of the vehicle. Since nearly all highways carry both passenger-automobile and truck traffic, it is essential that design criteria take into account the characteristics of different types of vehicles. A thorough knowledge of these characteristics will aid the highway and/or traffic engineer in designing highways and traffic-control systems that allow the safe and smooth operation of a moving vehicle, particularly during the basic maneuvers of passing, stopping, and turning. Therefore, designing a highway involves the selection of a design vehicle, whose characteristics will encompass those of nearly all vehicles expected to use the highway. The characteristics of the design vehicle are then used to determine criteria for geometric design, intersection design, and sight-distance requirements.

Static characteristics

The size of the design vehicle for a highway is an important factor in the determination of design standards for several physical components of the highway. These include lane width, shoulder width, length and width of parking bays, and lengths of vertical curves. The axle weights of the vehicles expected on the highway are important when pavement depths and maximum grades are being determined.

For many years, each state prescribed by law the size and weight limits for trucks using its highways, and in some cases local authorities also imposed more severe restrictions on some roads. Table 3.1 shows some features of static characteristics for which limits were prescribed. A range of maximum allowable values is given for each feature.

Type	*Allowable Lengths (ft)*
Bus	35–60
Single truck	35–60
Trailer, semi/full	35–48
Semitrailer	55–85
Truck trailer	55–85
Tractor semitrailer trailer	55–85
Truck trailer trailer	65–80
Tractor semitrailer, trailer, trailer	60–105
Type	*Allowable Weights (lb)*
Single-axle	18,000–24,000
Tandem-axle	32,000–40,000
State maximum gross vehicle weight	73,280–164,000
Interstate maximum gross vehicle weight	73,280–164,000

SOURCE: Adapted from *State Maximum Sizes and Weights for Motor Vehicles*, Motor Vehicle Manufacturers' Association of the United States, Detroit, MI, May 1982.

Fig: Range of State Limits on Vehicle Lengths by Type and Maximum Weight of Vehicle

Since the passage of the Surface Transportation Assistance Act of 1982, the maximum allowable truck sizes and weights on Interstate and other qualifying federal-aided highways are at most:

- 80,000 lb gross weight, with axle loads of up to 20,000 lb for single axles and 34,000 lb for tandem (double) axles
- 102 in. width for all trucks
- 48 ft length for semitrailers and trailers

- 28 ft length for each twin trailer

(Note: Those states that had higher weight limits before this law was enacted are allowed to retain them for intrastate travel.)

The federal regulations also stipulate that the overall maximum gross weight for a group of two or more consecutive axles should be determined from the equation:

$$W = 500\left[\frac{LN}{N-1} + 12N + 36\right] \tag{1.2}$$

Where,

W = overall gross weight (calculated to the nearest 500 lb)
L = the extreme of any group of two or more consecutive axles (ft)
N = number of axles in the group under consideration

The regulations also stipulate that a gross load of 34,000 lb may be carried by two consecutive sets of tandem axles if the overall distance between the first and last axles of the consecutive sets is 36 ft or more.

Numerical Example:

A 5-axle truck traveling on an interstate highway has the following axle characteristics:

Distance between the front single axle and the first set of tandem axles = 20 ft
Distance between the first set of tandem axle and the back set of tandem axles = 48 ft

If the overall gross weight of the truck is 79,500 lb, determine whether this truck satisfies federal weight regulations.

Solution: Although the overall gross weight is less than the maximum allowable of 80,000 lb, the allowable gross weight based on the axle configuration should be checked.

Use Eq. 3.2.

$$W = 500\left[\frac{LN}{N-1} + 12N + 36\right]$$

$$W = 500\left[\frac{48 \times 4}{4-1} + 12 \times 4 + 36\right]$$

$$= 74{,}000 \text{ lb}$$

which is less than the allowable of 80,000 lb. The truck therefore satisfies the federal truck weight regulations.

Because the static characteristics of the predominant vehicles are used to establish certain geometric parameters of the road, vehicles have been classified so that they represent static characteristics within a particular class. AASHTO has classified automobiles into four general classes of vehicles: passenger cars, buses, trucks, and recreational vehicles. Vehicles included in the passenger-car class are all sizes of passenger cars, sport/utility vehicles, minivans, vans, and pick-up trucks. Those in the bus class include intercity (motor coaches), city transit, school, and articulated buses. Vehicles in the truck class are single-unit trucks, truck tractor–semitrailer combinations, and trucks or truck tractors with semitrailers in combination with full trailers. Vehicles in the recreational-vehicle class are motor homes, cars with camper trailers, cars with boat trailers, motor homes with boat trailers, and motor homes pulling cars. The largest and most frequent vehicle expected to use the facility is selected as the design vehicle, and the following guidelines are provided for selection:

1. For parking lots or a series of parking lots, passenger-car class could be considered.

2. For intersections on residential streets and park roads, a single-unit truck class could be considered.

3. For intersections of state highways with city streets on which buses travel, but on which relatively few large trucks travel, a city transit bus class could be considered.

4. For intersections of highways with low-volume county highways and township/local roads with traffic volume under 400 Average Daily Traffic (ADT), a large school bus (84 passengers) or a conventional school bus (65 passengers) could be considered.

5. For other intersections of state highways and industrial streets with high volumes of traffic and/or providing large truck access to local facilities, the minimum design vehicle is the WB20 (WB65 or WB67)

6. For intersections of freeway ramp terminals with arterial crossroads the minimum design vehicle is the WB20 (WB65 or WB67)

The characteristic of vehicle categories that influence design of intersections when speeds are 15 km/h or less are (1) the minimum centerline radius (CTR); (2) the out-to-out track width; (3) the wheel base; and (4) the path of the inner rear tire of the vehicle as it makes a turn at the intersection. When turns are made at speeds of 10 km/h or less, the turning radius and turning path depend mainly on the size of the vehicle making the turn. These have therefore been established for each design vehicle.

Design Vehicle Type	Symbol	Dimensions (ft)											
		Overall			Overhang								
		Height	Width	Length	Front	Rear	WB_1	WB_2	S	T	WB_3	WB_4	Typical Kingpin to Center of Rear Axle
Passenger Car	P	4.25	7	19	3	5	11	—	—	—	—	—	—
Single-Unit Truck	SU	11–13.5	8.0	30	4	6	20	—	—	—	—	—	—
Buses													
Intercity Bus (Motor Coaches)	BUS-40	12.0	8.5	40	6	6.3[a]	24	3.7	—	—	—	—	—
	BUS-45	12.0	8.5	45	6	8.5[a]	26.5	4.0	—	—	—	—	—
City Transit Bus	CITY-BUS	10.5	8.5	40	7	8	25	—	—	—	—	—	—
Conventional School Bus (65 pass.)	S-BUS 36	10.5	8.0	35.8	2.5	12	21.3	—	—	—	—	—	—
Large School Bus (84 pass.)	S-BUS 40	10.5	8.0	40	7	13	20	—	—	—	—	—	—
Articulated Bus	A-BUS	11.0	8.5	60	8.6	10	22.0	19.4	6.2[b]	13.2[b]	—	—	—
Trucks													
Intermediate Semitrailer	WB-40	13.5	8.0	45.5	3	2.5[a]	12.5	27.5	—	—	—	—	27.5
Intermediate Semitrailer	WB-50	13.5	8.5	55	3	2[a]	14.6	35.4	—	—	—	—	37.5
Interstate Semitrailer	WB-62*	13.5	8.5	68.5	4	2.5[a]	21.6	40.4	—	—	—	—	42.5
Interstate Semitrailer	WB-65** or WB-67	13.5	8.5	73.5	4	4.5–2.5[a]	21.6	43.4–45.4	—	—	—	—	45.5–47.5
"Double-Bottom"-Semitrailer/Trailer	WB-67D	13.5	8.5	73.3	2.33	3	11.0	23.0	3.0[c]	7.0[c]	23.0	—	23.0
Triple-Semitrailer/Trailers	WB-100T	13.5	8.5	104.8	2.33	3	11.0	22.5	3.0[d]	7.0[d]	23.0	23.0	23.0
Turnpike Double-Semitrailer/Trailer	WB-109D*	13.5	8.5	114	2.33	2.5[a]	14.3	39.9	2.5[e]	10.0[e]	44.5	—	42.5
Recreational Vehicles													
Motor Home	MH	12	8	30	4	6	20	—	—	—	—	—	—
Car and Camper Trailer	P/T	10	8	48.7	3	10	11	—	5	19	—	—	—
Car and Boat Trailer	P/B	—	8	42	3	8	11	—	5	15	—	—	—
Motor Home and Boat Trailer	MH/B	12	8	53	4	8	20	—	6	15	—	—	—
Farm Tractor[f]	TR	10	8-10	16[g]	—	—	10	9	3	6.5	—	—	—

* = Design vehicle with 48-ft trailer as adopted in 1982 Surface Transportation Assistance Act (STAA).
** = Design vehicle with 53-ft trailer as grandfathered in with 1982 Surface Transportation Assistance Act (STAA).
a = This is overhang from the back axle of the tandem axle assembly.
b = Combined dimension is 19.4 ft and articulating section is 4 ft wide.
c = Combined dimension is typically 10.0 ft.
d = Combined dimension is typically 10.0 ft.
e = Combined dimension is typically 12.5 ft.
f = Dimensions are for a 150–200 hp tractor excluding any wagon length.
g = To obtain the total length of tractor and one wagon, add 18.5 ft to tractor length. Wagon length is measured from front of drawbar to rear of wagon, and drawbar is 6.5 ft long.

- WB_1, WB_2, and WB_4 are the effective vehicle wheelbases, or distances between axle groups, starting at the front and working towards the back of each unit.
- S is the distance from the rear effective axle to the hitch point or point of articulation.
- T is the distance from the hitch point or point of articulation measured back to the center of the next axle or center of tandem axle assembly.

Source: Adapted from *A Policy on Geometric Design of Highways and Streets,* American Association of State Highway and Transportation Officials, Washington, D.C., 2004. Used by permission.

Kinematic Characteristics

The primary element among kinematic characteristics is the acceleration capability of the vehicle. Acceleration capability is important in several traffic operations, such as passing maneuvers and gap acceptance. Also, the dimensioning of highway features such as freeway ramps and passing lanes is often governed by acceleration rates. Acceleration is also important in determining the forces that cause motion. Therefore, a study of the kinematic characteristics of the vehicle primarily involves a study of how acceleration rates influence the elements of motion, such as velocity and distance. We therefore review in this section the mathematical relationships among acceleration, velocity, distance, and time.

Let us consider a vehicle moving along a straight line from point o to point m, a distance x in a reference plane T. The position vector of the vehicle after time t may be expressed as:

$$r_{om} = x\hat{i} \tag{1.3}$$

Where,

r_{om} = positon vector for m in T
$\hat{i}$ = a unit vector parallel to line om
x = distance along the straight line

The velocity and acceleration for m may be simply expressed as

$$u_m = \dot{r}_{om} = \dot{x}_{\hat{i}} \tag{1.4}$$
$$a_m = \ddot{r}_{om} = \ddot{x}_{\hat{i}} \tag{1.5}$$

Where,

u_m = velocity of the vehicle at point m
a_m = accleration of the vehicle at point m
$\dot{x} = \frac{dy}{dx}$
$\ddot{x} = \frac{d^2y}{dx^2}$

Two cases are of interest: (1) acceleration is assumed constant; (2) acceleration is a function of velocity.

Case 1: Acceleration is constant When the acceleration of the vehicle is assumed to be constant,

$$\ddot{x}_{\hat{i}} = a \tag{1.6}$$

$$\frac{d\dot{x}}{dt} = a \tag{1.7}$$

$$\dot{x} = at + C_1 \tag{1.8}$$

$$x = \frac{1}{2}at^2 + C_1 t + C_2 \tag{1.9}$$

The constants C_1 and C_2 are determined either by the initial conditions on velocity and position or by using known positions of the vehicle at two different times.

Case 2: Acceleration is the function of velocity The assumption of constant acceleration has some limitations, because the accelerating capability of a vehicle at any time t is related to the speed of the vehicle at that time (u_t). The lower the speed, the higher the acceleration rate that can be obtained. One model that is used commonly in this case is

$$\frac{du_t}{dt} = \alpha - \beta u_t \tag{1.10}$$

Where α and β are constants.

In this model, the maximum acceleration rate that can be achieved is theoretically α, which means that α has units of acceleration as its unit. The term βu_t also should have units of acceleration as its unit, which means that b has the inverse of time (for example, sec^-1) as its unit.

Integrating Equation 1.10 gives,

$$-\frac{1}{\beta}\ln(\alpha - \beta u_t) = t + C$$

If the velocity is u_o at $t = 0$,

$$C = -\frac{1}{\beta}\ln(\alpha - \beta u_o)$$

$$-\frac{1}{\beta}\ln(\alpha - \beta u_t) = t - \frac{1}{\beta}\ln(\alpha - \beta u_o)$$

$$\ln\frac{(\alpha - \beta u_o)}{(\alpha - \beta u_t)} = \beta t$$

$$\therefore t = \frac{1}{\beta}\ln\frac{(\alpha - \beta u_o)}{(\alpha - \beta u_t)} \tag{1.11}$$

$$\alpha - \beta u_t = (\alpha - \beta u_o)e^{-\beta t}$$
$$u_t = \frac{\alpha}{\beta}(1 - e^{-\beta t}) + u_o e^{-\beta t} \tag{1.12}$$

The distance $x(t)$ traveled at any time t may be determined by integrating equation 1.12

$$\begin{aligned}
x = \int_0^t u_t dt &= \int_0^t \left[\frac{\alpha}{\beta}(1 - e^{-\beta t}) + u_o e^{-\beta t}\right] dt \\
&= \left[\frac{\alpha}{\beta}\left(t + \frac{e^{-\beta t}}{\beta}\right) - \frac{u_o}{\beta}e^{\beta t}\right]_0^t \\
&= \left[\frac{\alpha}{\beta}t + \frac{\alpha}{\beta^2}e^{-\beta t} - \frac{u_o}{\beta}e^{-\beta t}\right] \\
&= \frac{\alpha}{\beta}t + \frac{\alpha}{\beta^2}e^{-\beta t} - \frac{u_o}{\beta}e^{-\beta t} - \frac{\alpha}{\beta^2} + \frac{u_0}{\beta} \\
&= \left(\frac{\alpha}{\beta}\right) t - \frac{\alpha}{\beta^2}(1 - e^{-\beta t}) + \frac{u_o}{\beta}(1 - e^{-\beta t})
\end{aligned}$$

Numerical Example:

The acceleration of a vehicle can be represented by the following equation.

$$\frac{du_t}{dt} = 33 - 0.04u$$

where u is the vehicle speed in ft/sec. If the vehicle is traveling at 45 mi/h, determine its velocity after 5 sec of acceleration and the distance traveled during that time.

Solution:

- Convert 45 mi/h to ft/sec:
- $45 \times 1.47 = 66.15$ ft/sec
- Use Eq. 3.11 to determine velocity u_t after time t.

$$\begin{aligned} u &= \frac{\alpha}{\beta}(1 - e^{-\beta t}) + u_o e^{-\beta t} \\ \alpha &= 3.3 \\ \beta &= 0.04 \\ u_t &= \frac{3.3}{0.04}(1 - e^{-(0.04\times 5)}) + 66.15e^{-(0.04\times 5)} \\ &= 82.5(1 - 0.82) + 66.15 \times 0.82 \\ &= 14.85 + 54.24 \\ &= 69.09 \text{ ft/sec} \end{aligned}$$

- Convert ft/sec to mi/h.

$$\begin{aligned} u_t &= 69.09/1.47 \\ &= 47.00 \text{ mi/h} \end{aligned}$$

- Use Eq. 3.12 to determine distance traveled.

$$\begin{aligned} x &= \left(\frac{\alpha}{\beta}\right)t - \frac{\alpha}{\beta^2}(1 - e^{-\beta t}) + \frac{u_o}{\beta}(1 - e^{-\beta t}) \\ &= \left(\frac{3.3}{0.04}\right)5 - \frac{3.3}{(0.04)^2}(1 - e^{-0.04\times 5}) + \frac{66.15}{0.04}(1 - e^{-0.04\times 5}) \\ &= 412.5 - 2062.5(1 - 0.82) + 1653.75(1 - 0.82) \\ &= 412.5 - 371.25 + 297.68 \\ &= 338.93 \text{ ft} \end{aligned}$$

Dynamic Characteristics

Several forces act on a vehicle while it is in motion: air resistance, grade resistance, rolling resistance, and curve resistance. The extents to which these forces affect the operation of the vehicle are discussed in this section.

Air Resistance A vehicle in motion has to overcome the resistance of the air in front of it as well as the force due to the frictional action of the air around it. The force required to overcome these is known as the air resistance and is related to the cross-sectional area of the vehicle in a direction perpendicular to the direction of motion and to the square of the speed of the vehicle. Claffey has shown that this force can be estimated from the formula

$$R_a = \frac{1}{2}\frac{(2.15\rho C_d A u^2)}{g} \tag{1.13}$$

Where,

R_a = air resistance force (lb)

ρ = density of air (0.0766 lb/ft^3) at sea level; less at higher elevations

C_d = aerodynamic drag coefficient

(current average value for passenger cars is 0.4; for trucks, this value ranges from 0.5 to 0.8, but a typical value is 0.5)

A = frontal cross-sectional area (ft^2)

u = vehicle speed (mi/h)

g = acceleration of gravity (32.2 ft /sec2)

$$R_a = \frac{0.0772 \times \rho \times C_D \times A \times u^2}{2} \tag{1.14}$$

Where,

R_a = air resistance force (N)

ρ = density of air (1.227 kg/m^3) at sea level; less at higher elevations

C_d = aerodynamic drag coefficient

(current average value for passenger cars is 0.4; for trucks, this value ranges from 0.5 to 0.8, but a typical value is 0.5)

A = frontal cross-sectional area (m^2)

u = vehicle speed (Km/h)

Grade Resistance When a vehicle moves up a grade, a component of the weight of the vehicle acts downward, along the plane of the highway. This creates a force acting in a direction opposite that of the motion. This force is the grade resistance. A vehicle traveling up a grade will therefore tend to lose speed unless an accelerating force is applied. The speed achieved at any point along the grade for a given rate of acceleration will depend on the grade. Figure below shows the relationships between speed achieved and distance traveled on different grades by a typical heavy truck of 200 lb/hp during maximum acceleration. Note: grade resistance = weight * grade, in decimal.

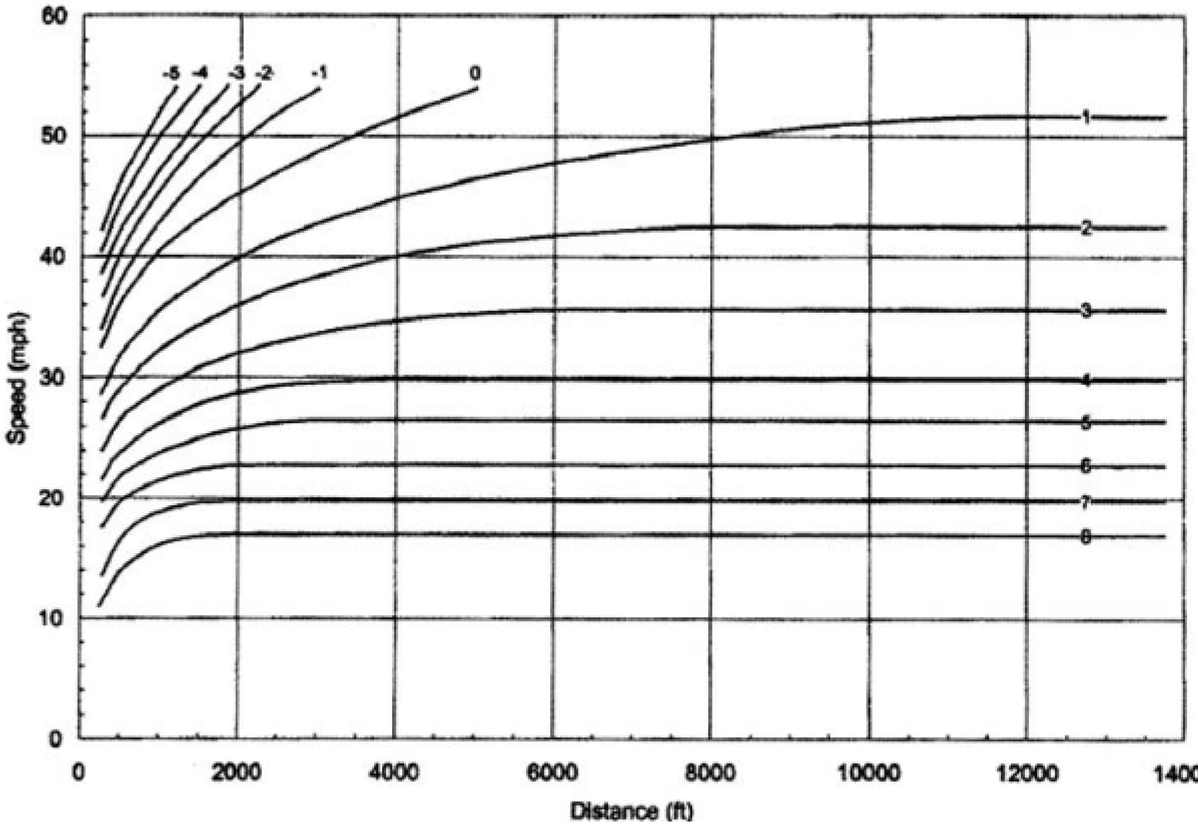


Fig: Speed-Distance Curves for Acceleration of a Typical Heavy Truck of 120 kg/kw [200 lb/hp] on Upgrades and Downgrades

Rolling Resistance There are forces within the vehicle itself that offer resistance to motion. These forces are due mainly to frictional effect on moving parts of the vehicle, but they also include the frictional slip between the pavement surface and the tires. The sum effect of these forces on motion is known as rolling resistance. The rolling resistance depends on the speed of the vehicle and the type of pavement. Rolling forces are relatively lower on smooth pavements than on rough pavements.

The rolling resistance force for passenger cars on a smooth pavement can be determined from the relation

$$R_r = (C_{rs} + 2.15 C_{rv} u^2) W \tag{1.15}$$

Where,

R_r = rolling resistance force (lb)
C_{rs} = constant (typically 0.012 for passenger cars)

C_{rv} = constant (typically $0.65 * 10^{-6}$ sec2/ft2 for passenger cars)
u = vehicle speed (mi/h)
W = gross vehicle weight (lb)

$$R_r = (C_{rs} + 0.0772 \times C_{rv} \times u^2)W \tag{1.16}$$

Where,
R_r = rolling resistance force (N)
C_{rs} = constant (typically 0.012 for passenger cars)
C_{rv} = constant (typically 6.99×10^{-6} sec^2/m^2 for passenger cars)
u = vehicle speed (Km/h)
W = gross vehicle weight (N)

For trucks, the rolling resistance can be obtained from

$$R_r = (C_a + 1.47C_b u)W \tag{1.17}$$

Where,
R_r = rolling resistance force (lb)
C_a = constant (typically 0.2445 for trucks)
C_b = constant (typically 0.00044 sec/ft for trucks)
u = vehicle speed (mi/h)
W = gross vehicle weight (lb)

$$R_r = (C_a + 0.278 \times C_b \times u) \times W \tag{1.18}$$

Where,
R_r = rolling resistance force (N)
C_a = constant (typically 0.2445 for trucks)
C_b = constant (typically 1.44×10^{-3} sec/m for trucks)
u = vehicle speed (Km/h)
W = gross vehicle weight (N)

The surface condition of the pavement has a significant effect on the rolling resistance. For example, at a speed of 50 mi/h on a badly broken and patched asphalt surface, the rolling resistance is 51 lb/ton of weight, whereas at the same speed on a loose sand surface, the rolling resistance is 76 lb/ton of weight.

Curve Resistance When a passenger car is maneuvered to take a curve, external forces act on the front wheels of the vehicle. These forces have components that have a retarding effect on the forward motion of the vehicle. The sum effect of these components constitutes the curve resistance. This resistance depends on the radius of the curve, the gross weight of the vehicle, and the velocity at which the vehicle is moving. It can be determined as

$$R_c = 0.5\frac{(2.15u^2W)}{gR} \tag{1.19}$$

Where,

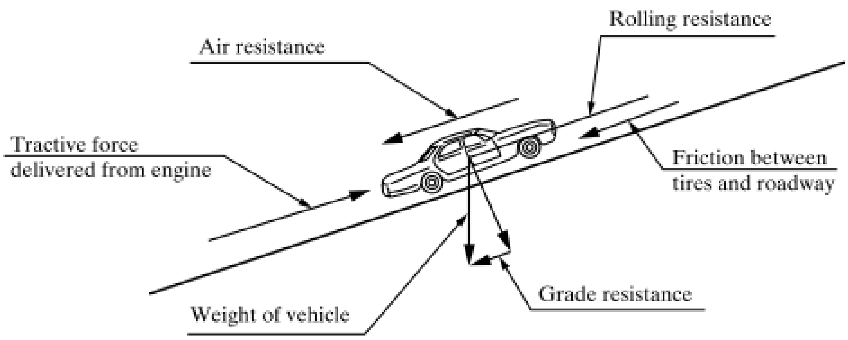
R_c = curve resistance (lb)
u = vehicle speed (mi/h)
W = gross vehicle weight (lb)
g = acceleration due to gravity (ft/sec^2)
R = radius of curvature (ft)

$$R_c = 0.5 \times \frac{0.0772 \times u^2 \times W}{g \times R} \tag{1.20}$$

Where,

R_c = curve resistance (N)
u = vehicle speed (Km/h)
W = gross vehicle weight (N)
g = acceleration due to gravity (m/sec^2)
R = radius of curvature (m)

Power Requirements Power is the rate at which work is done. It is usually expressed in horsepower (a U.S. unit of measure), where 1 horsepower is 550 lb-ft /sec. The performance capability of a vehicle is measured in terms of the horsepower the engine can produce to overcome air, grade, curve, and friction resistance forces and put the vehicle in motion. Figure below shows how these forces act on the moving vehicle.



The power delivered by the engine is

$$P = \frac{1.47Ru}{550} \tag{1.21}$$

Where,

P = horsepower delivered (hp)

R = sum of resistance to motion (lb)
u = speed of vehicle (mi/h)

$$P = \frac{0.278 \times F \times u}{76.04} \tag{1.22}$$

Where,
P = horsepower delivered (hp)
R = sum of resistance to motion (Kg)
u = speed of vehicle (Km/h)

Numerical Example:

Determine the horsepower produced by a passenger car traveling at a speed of 65 mi/h on a straight road of 5% grade with a smooth pavement. Assume the weight of the car is 4000 lb and the cross-sectional area of the car is 40 ft^2.

Solution: The force produced by the car should be at least equal to the sum of the acting resistance forces.

$$R = (\text{air resistance}) + (\text{rolling resistance}) + (\text{upgrade resistance})$$

Note: There is no curve resistance since the road is straight.

- Use Eq. 3.13 to determine air resistance.

$$\begin{aligned} R_a &= 0.5\left(\frac{2.15pC_DAu^2}{g}\right) \\ &= 0.5\,\frac{2.15 \times 0.0766 \times 0.4 \times 40 \times 65 \times 65}{32.2} \\ &= 172.9 \text{ lb} \end{aligned}$$

- Use Eq. 3.14 to determine rolling resistance.

$$\begin{aligned} R_r &= (C_{rs} + 2.15C_{rv}u^2)(4000) \\ &= (0.012 + 2.15 \times 0.65 \times 10^{-6} \times 65 \times 65)(4000) \\ &= (0.012 + 0.006)(4000) \\ &= 0.018 \times 4000 \\ &= 72 \text{ lb} \end{aligned}$$

$$\text{Grade resistance} = 4000 \times \frac{5}{100} = 200 \text{ lb}$$

- Determine total resistance.

$$R = R_a + R_r + \text{grade resistance} = 172.9 + 72 + 200 = 444.9 \text{ lb}$$

- Use Eq. 3.17 to determine horsepower produced.

$$\begin{aligned} P &= \frac{1.47Ru}{550} \\ &= \frac{1.47 \times 444.9 \times 65}{550} \\ &= 77.3 \text{ hp} \end{aligned}$$

1.4.3 Pedestrian Characteristics

Pedestrian characteristics relevant to traffic and highway engineering practice include those of the driver, discussed in the preceding sections. In addition, other pedestrian characteristics may influence the design and location of pedestrian control devices. Such control devices include special pedestrian signals, safety zones and islands at intersections, pedestrian underpasses, elevated walkways, and crosswalks. Apart from visual and hearing characteristics, walking characteristics play a major part in the design of some of these controls. For example, the design of an all-red phase, which permits pedestrians to cross an intersection with heavy traffic, requires knowledge of the walking speeds of pedestrians. Observations of pedestrian movements have indicated that walking speeds vary between 3.0 and 8.0 ft /sec. Significant differences have also been observed between male and female walking speeds. At intersections, the mean male walking speed has been determined to be 4.93 ft /sec, and for females, 4.63 ft /sec. A more conservative value of 4.0 ft /sec is normally used for design purposes. However, Rouphail and others have shown that the average walking speed depends on the population of elderly pedestrians. For example, the average walking speed is 4.0 ft /sec when the percentage of elderly pedestrians is 20 percent or lower, but reduces to 3.0 ft /sec when the percentage of elderly pedestrians is higher than 20 percent. This factor therefore should be taken into consideration for the design of intersection pedestrian signals at locations where a high number of older pedestrians is expected. Consideration also should be given to the characteristics of handicapped pedestrians, such as the blind. Studies have shown that accidents involving blind pedestrians can be reduced by installing special signals. The blind pedestrian can turn the signal to a red phase by using a special key, which also rings a bell, indicating to the pedestrian that it is safe to cross. Ramps are also now being provided at intersection curbs to facilitate the crossing of the intersection by the occupant of a wheelchair. Also, consideration should be given to the relatively lower average walking speed of the handicapped pedestrian, which can vary from a low of 1.97 ft /sec to 3.66 ft /sec.

1.4.4 Road Characteristics

The characteristics of the highway discussed in this section are related to stopping and passing because these have a more direct relationship to the characteristics of the driver and the vehicle discussed earlier.

Sight Distance

Sight distance is the length of the roadway a driver can see ahead at any particular time. The sight distance available at each point of the highway must be such that, when a driver is traveling at the highway's design speed, adequate time is given after an object is observed in the vehicle's path to make the necessary evasive maneuvers without colliding with the object. The two types of sight distance are (1) stopping sight distance and (2) passing sight distance.

1. Stopping Sight Distance The stopping sight distance (SSD), for design purposes, is usually taken as the minimum sight distance required for a driver to stop a vehicle after seeing an object in the vehicle's path without hitting that object. This distance is the sum of the distance traveled during perception-reaction time and the distance traveled during braking. The SSD for a vehicle traveling at u mi/h is

$$SSD = 1.47ut + \frac{u^2}{30\left(\frac{a}{g} \pm G\right)} \tag{1.23}$$

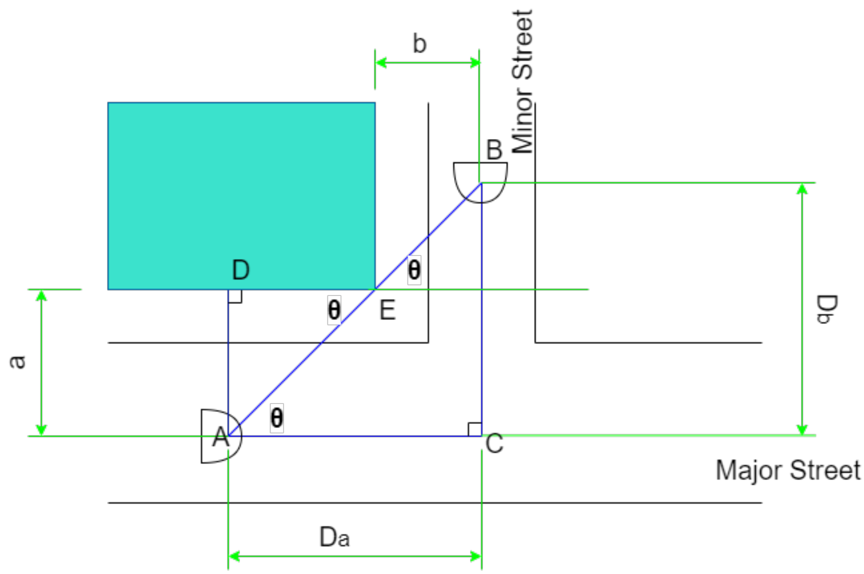
Where,

u = speed of the vehicle when brakes are appled
t = distance travelled between perception-reaction time
g = acceleration due to gravity
a = acceleration of the vehicle
$G = \tan\gamma$ (% Grade/100)

In general, Stopping Sight Distance (SSD) = lag distance + braking distance
lag distance = speed of vehicle (m/s) × total reaction time (s)
Braking distance = $\frac{v^2}{2gf}$ [For road without gradient]
Where,

v = speed of the vehicle (m/s)
g = acceleration due to gravity (9.8 m/s^2)
f = coefficient of friction

Calculation of speed limit at Major/ Minor Street intersection



Let a vehile 'A' is moving on a major street having a design speed of V_{maj} and vehicle 'B'is moving on a minor street having a design speed of V_{min}. V_{maj} and V_{min} should be computed such that safe operation can be achieved at the intersection.

From the fig, $\triangle$ CBA is the visibility traingle,
To avoid collision,
D_a = SSD for vehicle A + Clearance D_b = SSD for vehicle B + clearnce

We have,

$$D_a = V_{maj} \times t + \frac{V_{maj}^2}{2gf} + \ clearance \tag{1.24}$$

$$D_b = V_{min} \times t + \frac{V_{min}^2}{2gf} + \ clearance \tag{1.25}$$

From similar triangles, $\triangle$ CBA and $\triangle$ DAE,

$$\frac{D_b}{D_a} = \frac{a}{D_a - b} \tag{1.26}$$

$$D_b = \frac{a}{D_a - b} \times D_a \tag{1.27}$$

When design speed for major street is known and other parameters (t, f, and clearance) are known, the design speed for minor street can be computed using above equations and vice versa.

2. Passing Sight Distance The passing sight distance is the minimum sight distance required on a two-lane, twoway highway that will permit a driver to complete a passing maneuver without colliding with an opposing vehicle and without cutting off the passed vehicle. The passing sight distance will also allow the driver to successfully abort the passing maneuver (that is, return to the right lane behind the vehicle being passed) if he or she so desires. In determining minimum passing sight distances for design purposes, only single passes (that is, a single vehicle passing a single vehicle) are considered. Although it is possible for multiple passing maneuvers to occur (that is, more than one vehicle pass or are passed in one maneuver), it is not practical for minimum design criteria to be based on them.

Numerical Example:
Drivers with an average 20/40 vision travel at 55 mph in the curb lane of a freeway, where the exit ramps are designed for 25 mph. What should be the minimum distance of the sign with 6 inch letters placed ahead of the exit? (perception-reaction time = 2.5 sec, deceleration rate = 5 ft/s2 , drivers with 20/20 vision can read the sign at 60 ft per inch of letter height)
Solution:

initial velocity (u) = 55 mph; final velocity (v) = 25 mph ; perception reaction time (t) = 2.5 sec; deceleration rate (a) = 5 ft/s2

now,
Decelerating distance (d_1) = $\frac{u^2-v^2}{2\times a}$ = $\frac{55^2-25^2}{2\times 5} \times \frac{5280}{3600}^2$ = 516 ft

Also,
distance for a 20/20 vision to read the sign (D_{normal}) = 60 × 6 = 360 ft
visual acuity (X) = 20/40 = 0.5

$\therefore D_{sight} = D_{normal} \times X$ = 360 × 0.5 = 180 ft

then,
Then, distance required to react the stimulus (d_2) = u × t = 55 × (2.5/3600) = 202 ft

Now,
the minimum distance where the sign should placed ahead of the exit ramp(d) = $d_1 + d_2 - d_{sight}$= 516 + 202 – 180 = 538 ft

1.5 Review Problems

1. Explain Perception-Reaction process with a relevant example.

2. Define Visual Acuity and 20/20 vision. Derive the expression for visual angle.

3. Derive the mathematical expression for kinematic characteristics of a vehicle.

4. Determine the difference in air resistance between a passenger car and a single-unit truck if both vehicles are traveling at a speed of 96.5 km/h. Assume that the frontal cross-sectional area of the passenger car is 2.79 m^2 and that for the truck is 10.70 m^2. [1903.6 N]

5. Determine the rolling resistance on a passenger car that is traveling at 105 km/h if the weight of the car is 9000 N. [161.1 N]

6. A single-unit three-axle truck traveling on an interstate highway at a speed of 88.5 km/h approaches a horizontal curve with a radius of 274.25 m. Determine the curve resistance that acts on the truck as it traverses the curve if the weight on each axle is 22675 N. [7644.1 N]

7. A 13600 N passenger car is traveling at 88.5 km/h on a flat section of road with a horizontal curve of 300 m radius. If the cross-sectional area of the vehicle is 2.7 m^2, determine the horsepower required to overcome the resistance on the vehicle. [65.56 hp]

8. Two passenger cars are traveling at 88.5 km/h.The weight of car A is 9000 N, and that of car B is 18000 N.The cross-sectional area for car A is 3.15 m^2, and that of car B is 3.6 m^2. Determine the maximum grade on which passenger car A can travel without its total resistance exceeding that of car B traveling on a straight and level section of road.[2.36 %]

9. A passenger car being driven on a flat and straight section of a highway at a speed of 105 km/h reaches a curved section of the highway with a grade of 5% and a radius of 450 m. Determine:

   (a) The additional force that will be required to maintain the original speed of 105 km/h [1302.63 N]

   (b) The percentage increase in the total force to maintain the original speed of 105 km/h [59.3 %]

   Assume that the weight of the car is 907 kg, the cross-sectional area is 3.15 m^2, and the car is being driven at sea level.

10. A truck and a passenger car traveling on a section of highway at a speed of 80 km/h enter a short curved section of the road with a grade of 5% and a radius of 270 m. Determine the ratio of the additional force required by the truck to that required by the car for both vehicles to maintain their original speed of 80 km/h. Assume the weight of the car is 11250 N and that of the truck 54000 N. [5 : 1]

11. As a part of traffic management on minor road approach of an intersection, compute the speed limit for safe operation as per the Nepal Road Standard. The major road has been designed using 60 kmph. Assume constants a and b as 15m and 20m respectively. (As per the Nepal Road Standard, for a speed of 60 kmph, f = 0.38, t = 2.5 sec, and clearance = 5m) [36.54 kmph]

Chapter 2

Motor Vehicle and Transport Management Act

Please refer the "Motor Vehicles and Transport Management Act, 2049 (1993)" for this chapter.

Chapter 3

Transportation Economics

3.1 Introduction

All decisions related to planning, design, and improvement of transportation infrastructure have economic implications. Transport economics includes the issues such as transport location, movement of people and freight/goods, transport demand, transport planning and forecasting, direct and indirect cost of transport, pricing of transport services, investments in transport infrastructure and services, transport and social-economic development, and transport regulation. In this chapter the transport economics is considered from the micro economic perspective. We consider various aspects of the direct costs and pricing of transport infrastructure and services of different transportation modes.

Considered from this microeconomic perspective, it can be said that the transport sector consists of the demand and supply component. The transport demand is derived demand due to needs of people and freight/goods shipments to change the physical place. For example, many people live at one place but work and/or have a leisure on the others, which requires them to travel forward and backward. The location of companies providing raw materials is different than those of the users of these materials—manufacturers of the semifinal and final products, which requires transportation of these raw materials from the former to the latter. In addition, the manufacturers of the final products are often located far away from the retailers of these products, which again require transportation, this time of the final products. Thus, it can be said that the transport demand is derived demand. In many cases transportation demand is proportional to the volumes of peoples' activities and the quantities of final products they consume during a given period of time.

The transport demand is handled by the transport supply/capacity provided by transport companies. The transport companies generally provide transport infrastructure with the supportive facilities and equipment, and rolling stock/vehicles carrying out transport services. In order to make them operational, the corresponding material, labor, ie, employed staff/personnel, and energy/fuel, are consumed. In terms of time, the transport infrastructure has particularly the long life-cycle, which is, for example, about 20, 30, 40, or even 60 years. That of rolling stock/vehicles is shorter (20–25 years) mainly due to its/their physical and also technological obsolescence, after they need to be replaced. In this context, two categories of transport companies can be distinguished: that providing transport infrastructure called "the infrastructure providers," and that providing transport services called "transport operators." They both constitute the transport systems within particular transport modes.

According to the economic jargon, the above-mentioned components of transport supply/capacity represent the main inputs to transport processes. The outputs of the transport process are the transport services produced in the given quantities and at the specified quality. They are consumed by users— passengers and/or freight shippers/receivers at the same time as they are produced. This implies that they are the short-lived similarly as transport demand without possibility to be stored/warehoused and left to be consumed sometimes latter on.

The most important economic categories of a given transport system are its costs, revenues, and their relationship.

The costs are expenses for maintaining the transport infrastructure and carrying out transport services. These costs are passed to users of these services (passengers and freight/goods shippers and receivers) in the form of prices/charges. These charges bring revenues to transport companies. In general, these prices/charges, ie, revenues, are set up to cover the company's costs and provide some profits. Nevertheless, the difference between revenues and costs can generally be positive or negative, thus representing the profits or losses. Profits or losses are calculated for the given period of time (usually for a quarter of year, year, or few years).

The prices charged to users of transport services represent for them direct costs. In addition, users are imposed the indirect costs, which are usually considered to be the cost of time during trip/travel/ transportation. In such case, the sum of direct and indirect costs is called the users' generalized trip/travel/transportation costs for users.

We describe in this chapter the main elements of economics of transportation systems considered from the engineering perspective. The chapter analyzes direct costs and revenues of the providers of transport infrastructure and transport operators. Costs and revenues could be analyzed at the level of the individual component (infrastructure, rolling stock) and/or of the entire company.

3.2 Scope of Transportation Economics

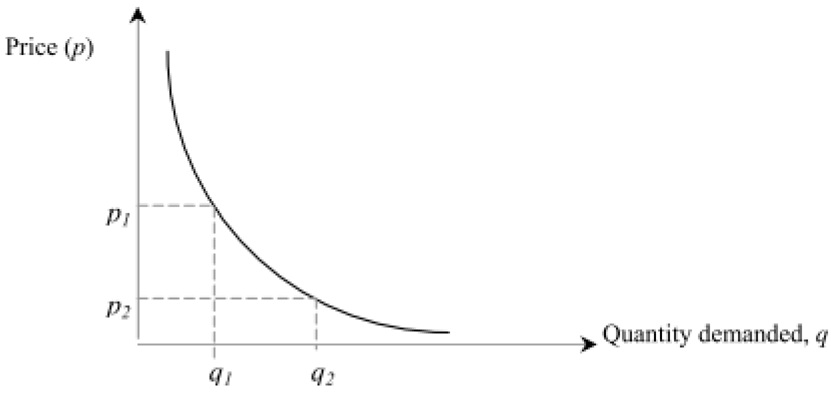
The study of economics is divided into macro-economics and micro-economics. Macro-economics is associated with the wealth of society on a regional scale, and deals with the behavior of aggregate concepts. On the other hand, micro-economics involves the behavior of relatively smaller entities such as firms and individuals. Transportation economics, while considered a branch of applied micro-economics, is associated with certain unique issues such as:

1. the demand for transportation is not direct, but is derived
2. the consumption of each transportation facility (i.e., each trip) is unique in time and space
3. technological differences among different modes and economies of scale
4. governmental interventionist policies and regulations in transportation

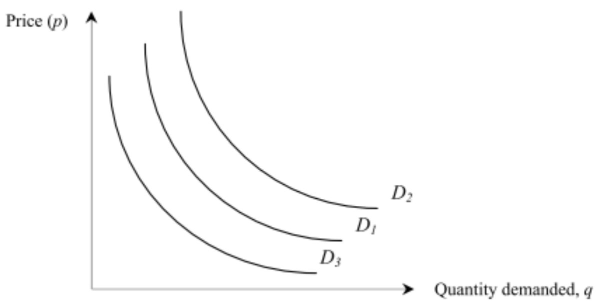
Transportation economics specifically addresses demand of transportation services, supply of transportation facilities, elasticity of demand and supply, price mechanisms, and transportation cost analysis

3.3 Transportation Demand

Like all other goods and services, the demand for any specific transportation facility demands on factors pertaining to the consumer such as income, and characteristics of the facility such as the cost associated with its use (in terms of time and price) relative to rival facilities. A typical example of such demand is that for auto travel: lower incomes of consumers, coupled with lower costs and travel times associated with transit are expected to lead to reduced demand for auto travel. Transportation demand analysis involves demand functions (which represents the willingness of consumers to purchase the transportation "product" at various alternative prices, i.e., the demand-price curve, and demand models (which estimate the probability that an individual (or fraction of a set of individuals) will choose a particular product over the other.



A hypothetical example of an aggregate transportation demand function is provided as shown in the above figure. This represents the amount of travel people are willing to make by transit at various transit fare (price) levels. Transportation demand functions, either in the form of a graph or an equation, are useful in transportation planning because they enable the determination of expected demand at any price. A specific demand curve represents the demand-price relationship given a set of conditions specific to the transportation product in question (referred to as alternative-specific attributes, such as travel time, comfort, convenience), and also specific to the users (income levels and other socioeconomic characteristics). Changes in such conditions often result in changes in the levels of transportation demand, even at fixed price of that product. For example, increased unemployment would likely lead to reduced demand for travel. Also, an increase in costs associated with auto use is likely to result in increased transit demand, even if transit fares remain the same. When such changes in conditions (other than price) occur, they are represented as a shift in the demand curve shown in the figure below (upward shift for increased demand, $D_1 \rightarrow D_2$; and downward shift for decreased demand, $D_1 \rightarrow D_3$).



3.3.1 Causes of Shift in Transportation Demand Curve

There is a possibility of a change in demand of a transportation facility even when its price remains same, and this is reflected as a shift in the demand curve for that transportation facility. Factors that influence such demand shifts are discussed below:

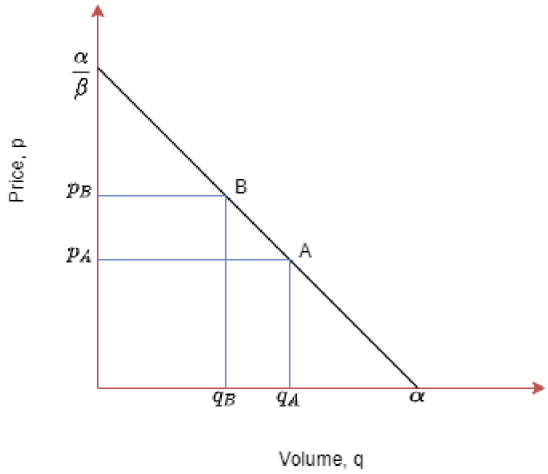
- ***Sudden change in customer preference***: (Season, life, style, e.t.c) For example, more people seem to ride transit in the winter season.
- ***Change in the level of attribute of interest (e.g., increase in price) of related goods***: For complementary products, a decrease in price of product increases the demand for other product, shifting the latter's demand curve to the right (e.g., parking spaces, automobile use). For rival products, an increase in price of a product increases the demand for its rival product, shifting the latter's demand curve to the right (e.g., transit, auto).
- ***Change in regional income***: An increase in income shifts the demand curve to the right. A normal good is one whose demand increases as a person's income increases.
- ***Change in number of potential customers***: An increase in market size shifts the demand curve towards right.
- ***Expectation of impending change in the level of the attribute of interest***: e.g., A news report predicting higher prices in the future can cause a shift in the demand curve at the current price as customers purchase increases in the anticipation of the price change.

3.3.2 Transportation Demand Functions

Transportation demand models are used to determine the volume of travel demanded, at various levels of service and have been described as "a representation of human behavior which can be used to predict how individuals or firms will change transportation choices in response to changes in future conditions". Within the context of transportation economics, a trip maker is defined as a consumer, in the economics meaning of the word, as the trip maker, by planning a trip, seeks to consume the service offered by transportation facilities. There are two types of demand functions:

1. *Disaggregate demand functions*: These predict the behavior of a single consumer in response to changes in future conditions.
2. *Aggregate demand functions*: These predict the behavior of a group of consumers such s a household, in response to changes in future conditions.

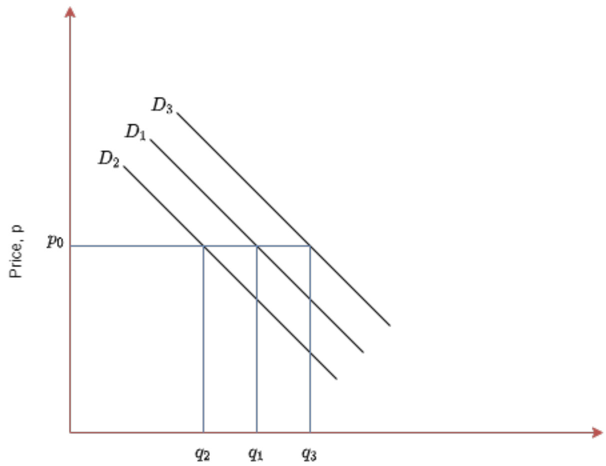
Linear Demand Function



A linear demand function for travel as shown in the above figure for a given pair of origin and destination points, at a specific time of day and for a particular purpose. Such a demand function is useful for predicting travel over a wide range of conditions. This demand function is useful for predicting travel over a wide range of conditions. This demand function assumes a particular level and distribution of income, population, and socioeconomic characteristics. Note that it is an aggregate demand curve, representing the volume of trips demanded at different prices by a group of travelers. Functionally,

$$q = \alpha - \beta p \tag{3.1}$$

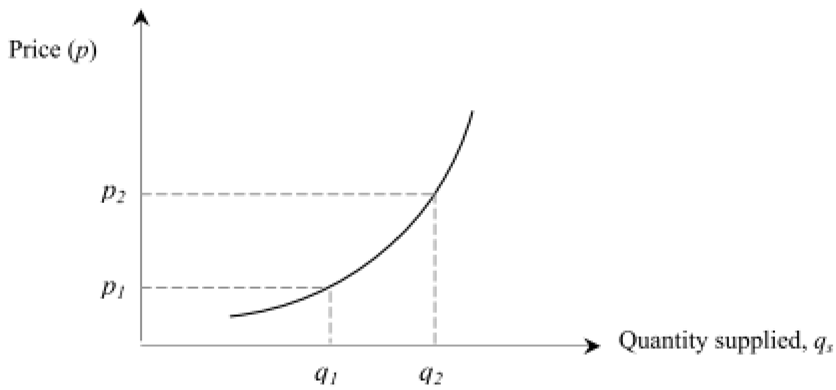
Where, q is the quantity of trips demanded, p is the price, and α and β are constant demand parameters.The demand function is drawn with a negative slope expressing a similar situation where a decrease in perceived price usually results in an increase in travel, although it is not always true.



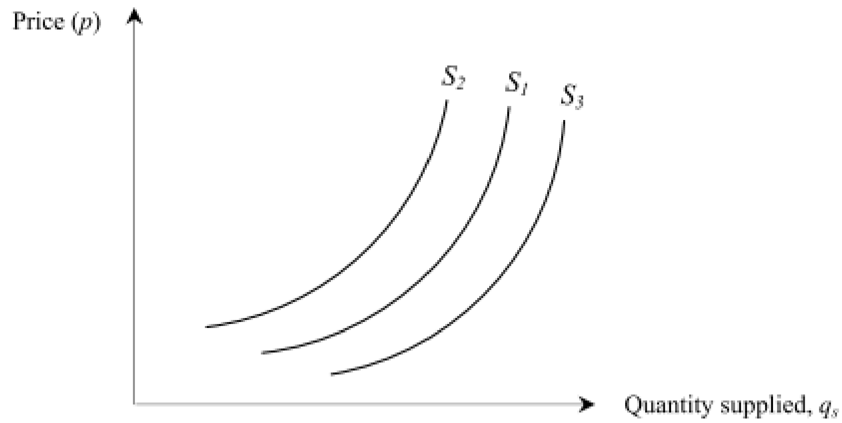
The above figure shows a series of shifted demand curves, representing changes in the quantity of travel due to variables other than the perceived price. Naturally, at a price p_0, one could expect different quantities q_1, q_2, and q_3 as the demand curve changes from D_1, D_2, and D_3. If the curve moves upward(from D_1 to D_3), it indicates an increase in trips.

It is important to distinguish short-run changes in the quantity of travel due to price changes represented by movement along a demand curve from long-run changes due to activity or behavioral variables, represented by shifts in demand functions.

3.4 Transportation Supply



The supply of a transportation product represents the quantity of that product a producer is willing to offer at a given price. However, transportation supply may also be associated with the quality of the product. At a given time, transportation facility supply depends on price: higher prices are an incentive for producers to make more profits who therefore increase supply levels. A hypothetical example of a transportation supply function is provided as shown in the figure above. This represents the amount of transportation products that suppliers are willing to make available at various prices. Transportation supply functions are useful in transportation planning because they enable the determination of expected supply at any future price. A specific supply curve represents the supply-price relationship given a set of conditions specific to the transportation product in question (referred to as alternative-specific attributes, such as travel time, comfort, convenience), and also specific to the producers (such as technology, policy and governmental intervention through policies and regulation. Changes in such conditions often result in changes in the levels of transportation supply, even at fixed price of that product. For example, improved increased unemployment would likely lead to reduced demand for travel. Also, an increase in costs associated with auto use is likely to result in increased transit demand, even if transit fares remain the same. When such changes in conditions (other than price) occur, they are represented as a shift in the supply curve shown in the figure below (upward shift for increased demand, $S_1 \rightarrow S_2$; and downward shift for decreased demand, $S_1 \rightarrow S_3$).



Increases in transportation supply may be traditionally thought of in terms of increasing the fleet size of a transit company or building new roads or increasing the number of lanes for existing roads. However, it is possible to increase supply without such physical capital-intensive investments. For instance, the use of intelligent transportation systems could lead to increased supply without any physical enlargements of the road network.

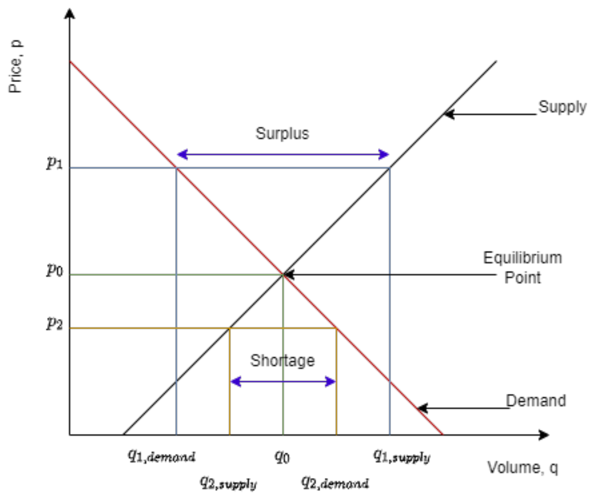
3.4.1 Causes of Shift in Transportation Supply Curve

The supply of transportation service may change even if the price remains constant, for reasons such as:

- ***Price of rival transportation services***: The supply of a service may decrease if there is a decrease in the price of competing transportation service. This may apply to toll roads, where profit is the primary intent, and to a lesser extent, non-toll roads.
- ***Number of transportation modes***: An increased number of modes, such as construction of a subway in a city that already has buses and light rail transit and facilities for autos, indicates an increase in supply, shifting the supply curve towards right.
- ***Price of relevant input***: If the cost of resources used to produce a transportation service increases, the transportation agency would be less capable of supplying the same quantity at a given price, and the supply curve will shift towards left.

- ***Technology***: Technological advances that increase facility capacity or efficiency causes the supply curve to shift towards right.

3.5 Equilibrium of Transportation Demand and Supply



We have seen that the demand function is a relation between quantity demanded of a good and its price. Similarly, the supply function(Service function) represents the quantity of goods a producer is willing to offer at a given price. If the demand and supply function for a transportation facility are known, then it is possible to deal with the concept of equilibrium. Equilibrium is said to be attained when factors that affect the quantity demanded and those that determine the quantity supplied result in being statically equal (or converging toward equilibrium).

Numerical example:
The travel time on a stretch of highway lane connecting two activity centers has been observed to follow the equation representing the service function:

$$t = 15 + 0.02v$$

Where t and v are measured in minutes and vehicles per hour respectively. The demand function for travel connecting the two activity centers is $v = 4000 - 120t$.

- Sketch these two equations and determine the equilibrium time and speed of travel.

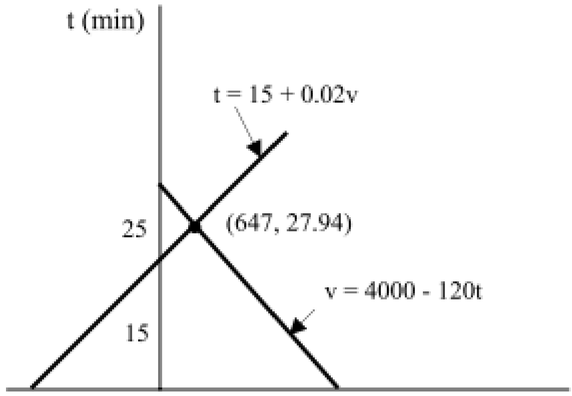
- If the length of the highway lane is 20 miles, what is the average speed of vehicles traversing this length?

Solution:

Demand function: $v = 4000 - 120t$
Supply function: $t = 15 + 0.02v$
Solving these two equations simultaneously yields:

$v = 647$ vehicles/hour
$t = 27.94$ minutes

$\therefore speed = \frac{20 \times 60}{27.94} = 42.95mph$



3.6 Sensitivity of Travel Demand

In the planning and evaluation of transportation systems and associated investments, it is often useful to have knowledge of the changes in transportation demand caused by changes in attributes of the transportation system or its environment. A particular instance is the change in demand for a given mode in response to changes in price of that mode. Given the functional form of the travel demand function, it is possible to derive a marginal effects model that estimates any one of the following:

- Change in demand in response to unit change in attribute
- Change in demand in response to unit percent change in attribute
- Percent change in demand in response to unit percent change in attribute

The law of demand states that a fall in the price of a good raises the quantity demanded. The price elasticity of demand measures how much the quantity demanded responds to a change in price. Demand for a good is said to be elastic if the quantity demanded responds substantially to changes in the price. Demand is said to be inelastic if the quantity demanded responds only slightly to changes in the price.

Transportation demand elasticity may be defined as the degree of responsiveness of transportation demand in response to a unit change in demand-related or attributes such as price or income. This is typically expressed as follows:

$$e_p = \frac{\textit{percentage change in quantity demanded}}{\textit{percentage change in price}}$$

$$e_p = \frac{\frac{\partial q}{q}}{\frac{\partial p}{p}} = \frac{\partial q}{\partial p} \times \frac{p}{q} \tag{3.2}$$

Arc elasticity can be calculated as:

$$\textit{Arc elasticity} = \frac{\frac{\partial q}{q}}{\frac{\partial p}{p}} = \frac{\partial q}{\partial p} \times \frac{p}{q} = \frac{\frac{Q_1 - Q_0}{\frac{Q_1 + Q_0}{2}}}{\frac{P_1 - P_0}{\frac{P_1 + P_0}{2}}} = \frac{Q_1 - Q_0 \times (P_1 + P_0)/2}{P_1 - P_0 \times (Q_1 + Q_0)/2} \tag{3.3}$$

Where, Q_0 and Q_1 represent quantity of travel demanded corresponding to prices P_0 and P_1 respectively.

3.6.1 Elasticity and Revenue Along a Linear Demand Curve

For a linear demand function, we can determine the elasticity of demand by taking derivative of equation 3.1:

$$\frac{dq}{dp} = -\beta$$

Also from equation 3.1,

$$p = \frac{\alpha - q}{\beta}$$

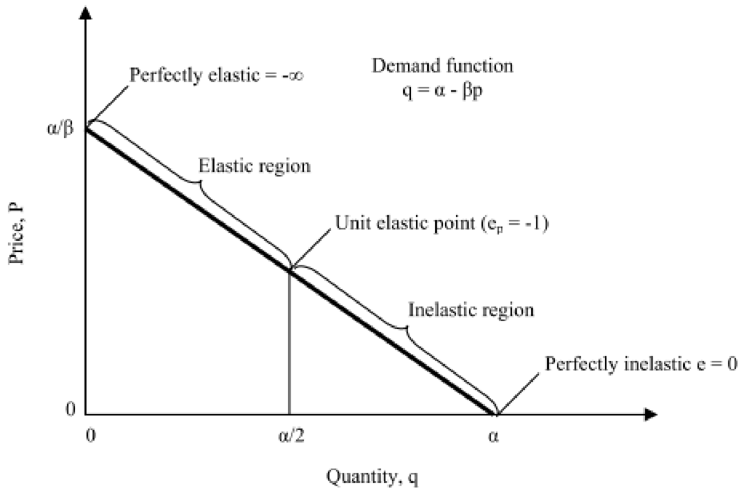
We have,

$$e_p = \frac{\partial q}{\partial p} \times \frac{p}{q}$$

Substituting the values of $\frac{dq}{dp}$ and p we get,

$$\therefore\ e_p = -\beta \times \frac{\alpha - q}{\beta\, q} = 1 - \frac{\alpha}{q} \tag{3.4}$$

When the elasticity is less than –1 (i.e., more negative than –1), the demand is described as being elastic, meaning that the resulting percentage change in quantity will be larger than the percentage change in price. In this case, demand is relatively sensitive to price change. However, when elasticity is between 0 and –1, the demand is described as being inelastic or relatively insensitive. These ranges are shown in the figure below:



A linear demand curve has several interesting properties. Notice, as one moves down the the demand curve, the price elasticity of demand becomes smaller (i.e., more inelastic). In fact, the elasticity at a given point equals the length of the demand line segment below the point divided by the length of the line segment above it. Another point to is that the slope of the line is constant, but the elasticity changes from ∞ at the top, where the demand line intersects the vertical axis, ti 0, where the demand line intersects the horizontal axis. Because elasticity changes along the demand curve, it is essential to specify what range of prices or quantity the elasticity was measured.

value	*Nature of Demand*	*Relation between Price and Revenue*	Impact of increase in price
$e > 1$	elastic	negatively related	reduces revenue
$e < 1$	inelastic	positively related	increases revenue
$e = 1$	unit elasticity	none	remains the same

3.7 Kraft Demand Function

We occasionally come across a demand function where the elasticity of demand for travel with respect to its price is essentially constant. The demand function for such a situation corresponds to the equation.

$$Q = \alpha(p)^{\beta} \tag{3.5}$$

Where α and β are constant parameters of the demand function. To prove that this function is has a constant elasticity, we differentiate this function with respect to price, we get:

$$\frac{dQ}{dp} = \alpha\,\beta\,p^{\beta-1}$$

Now substituting this value in standard elasticity equation 3.2,

$$e_p = \frac{dQ}{dp} \times \frac{p}{Q}$$

$$= \alpha\,\beta\,p^{\beta-1} \times \frac{p}{Q}$$

$$= \alpha\,\beta\,p^{\beta-1} \times \frac{p}{\alpha(p)^{\beta}} \;\; [from\, equation\, 3.5]$$

$$\therefore e_p = \beta \tag{3.6}$$

Thus, β, the exponent of price (a constant parameter), is the price elasticity.

Numerical Example:
The elasticity of transit demand with respect to price has been found to be equal to -2.75, which means that a 1% increase in transit fare will result in 2.75 decrease in the number of passengers using the system. A transit line on this system carries 12,500 passengers per day, charging 50 cents per ride. The management wants to raise the fare to 70 cents per ride. What advice would you offer to the management?
Solution:
Since the elasticity of the demand is a constant, the demand function is a Kraft Demand Model,

Q_1 = 12500
p_1 = 50
p_2 = 70
β = -2.75

We have,
$Q_1 = \alpha(p_1)^{\beta}$
$12500 = \alpha \times 50^{-2.75}$
$\therefore \alpha = 5.876 \times 10^{8}$

An increase in fare from 50 to 70 cents will attract a demand of,

$Q_2 = \alpha(p_2)^{\beta}$
$\therefore Q_2 = 5.876 \times 10^{8} \times 70^{-2.75} \;= 4955\, passengers$

Therefore, the increase in fare from 50 to 70 cents (40% increase) is likely to reduce the patronage on this line from 12500 passengers per day to 4955

(60 % decrease). In terms of revenue, the results are as follows:

Revenue before increase in fare (R_1) = 50 $\times$ 12500 = \$6250
Revenue after increase in fare (R_2) = 70 $\times$ 4955 = \$3468.50
$\therefore$ Loss in revenue = $R_1 - R_2$ = \$6250 - \$3468.50 = \$2781.5

Hence, the management is not advised to increase the fare.

3.8 Factors Affecting Elasticity

As stated earlier, elasticity is the change in demand in response to a unit change in levels of attributes of the transportation system or its environment. Such attributes include characteristics of the transportation system such as price and level of service associated with a given mode, price and level of service of competing modes, and characteristics of the socio-economic system such as income, level of employment, household size, car ownership, etc. Of these factors, of particular interest are price and income. The elasticity of demand in response to price and income are known as price elasticity and income elasticity, respectively.

3.8.1 Income Elasticity

The elasticity of demand with respect to income, or income elasticity, is the change in demand for a good in response to a unit change in income of the consumer of that good. Income elasticity has a special significance in travel demand modeling. Often, the transportation planner seeks to evaluate the impact of changing socioeconomic trends on the demand for or share of various modes of transportation. A major indicator of economic trends is income. In disaggregate demand modeling, it is sought to determine the sensitivity of changing income on the demand for a particular mode. Income elasticity is generally defined as the change in demand in response to a unit change in income. In transportation economics, a good service is considered normal if there is a direct relationship between the demand for that commodity and the income of the consumer. Besides increased demand, if the share of the demand for that good (in the consumer's total income) also increases, then the good is described as a superior good. If the demand of a good decreases with increasing income then the good is described as inferior. In developed countries, automobile travel is considered superior while mass transit is considered an inferior good.
Mathematically,

$$e_i = \frac{\%\ change\ in\ quantity\ of\ goods/service\ demanded}{\%\ change\ in\ income} \tag{3.7}$$

3.8.2 Price Elasticity

The elasticity of demand with respect to price, or price elasticity, is the change in demand for a good in response to a unit change in the price of the good. A study of price elasticity is important because it is often sought to assess the impacts of changing prices of a good or rival goods (due to past supply and demand conditions) on the demand of that good. The level of price elasticity depends on factors such as price of rival goods, income-share of the good, the scope of definition of the good, and whether the good is considered a luxury or a necessity. Some factors that affect price elasticity are as follows:

- *Price of rival goods*: A consumer who spends a substantial percentage of income on a particular good is more likely to seek a substitute good when the price of the good increases.

- *Scope of definition*: Goods that have narrow definitions are more likely to have more substitutes, and are therefore expected to have a more elastic demand.

- *Price an availability of rival goods*: The lower the price and greater availability of substitutes, the greater the elasticity of demand of the good with respect to price.

- *Luxury vs. substitute goods*: Goods that are considered necessities typically have inelastic price, while luxury goods have relatively elastic price.

3.9 Elasticity and Total Revenue

It is possible to tell what the revenue (price × output) of a firm is likely to be if the price of a unit changes. Some general relationship between elasticity and revenue are given below:

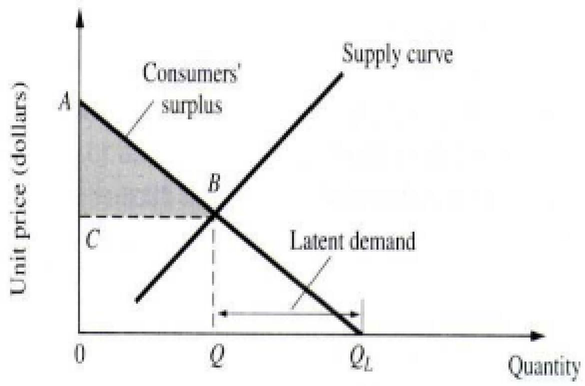
- When demand is inelastic ($e < 1$), price and total revenue move in the same direction: If the price increases, total revenue also increases.

- When demand is elastic ($e > 1$), price and total revenue move in opposite directions: If the price increases, total revenue decreases.

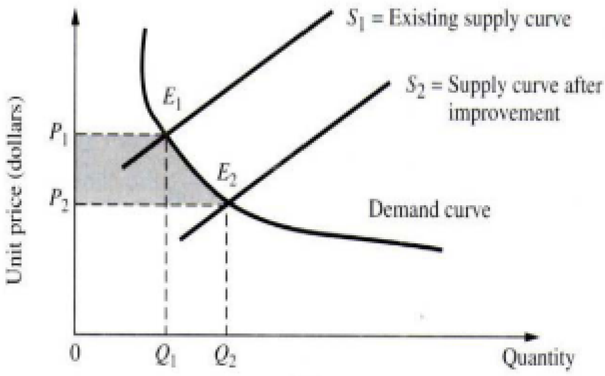
- If demand is unit elastic ($e = 1$), total revenue remains constant when the price changes.

3.10 Application of Price Elasticity: Consumer Surplus and Latent Demand

Consumer surplus is a measure of the monetary value made available to consumers by the existence of a facility. It is defined as the difference between what the consumers might be willing to pay for a service and what they actually pay. A patron of a bus service pays a fare of, say, 50 cents per trip but would be willing to pay up to as much as 75 cents per trip. In this case, her consumer surplus is 25 cents.



Consumer surplus may be determined from a survey of customers to determine how much they are willing to pay for a good or service. Recent studies have evaluated the willingness of road users to pay for improvements that would enhance safety on their highways. Such information can be combined with data on actual payments to determine the amount of consumer surplus per person or for a group of persons. Alternatively, consumer surplus can also be determined from the demand-supply curve, as the demand curve can be considered as an indicator of the utility of the good or service in terms of price, as illustrated in the figure above. The area ABC represents the total consumer surplus. Maximization of consumer surplus is the maximization of the economic utility of the consumer. The use of the consumer surplus concept is common in the area of the evaluation of transit systems.



Generally, the economic impacts of improvements in a transportation system can be evaluated in terms of consumer surplus, and can be represented as the area under a demand curve and a shift in the supply curve. The figure above indicates the case of an urban arterial street with a traffic supply S1, intersecting a demand curve at E1. An additional lane is added, thus shifting the supply curve to S2 and therefore intersecting the demand curve at E2. The change in consumer surplus can be quantified as the trapezoidal area P1, P2, E1, E2. In other words,

$$E = (P1 - P2) \times \frac{(Q1 + Q2)}{2}$$

The consumer surplus can be measured as the difference between the maximum amount which consumers are willing to pay for a specified quantity of a good or service rather than going without it. In general, the area AOQB in Figure 5-9 represents the total community benefit, BCOQ is equal to the market value, and ACB is equal to the consumer surplus or net community benefit. For a constant supply curve, a higher elasticity is associated with a smaller consumer surplus, and vice versa.

The foremost figure also illustrates an additional concept that is useful to transportation engineers: latent demand. It can be observed that travelers between Q and D do not make trips at the current time, but would do so if the price per trip were lower than the equilibrium price. The number of such potential travelers is referred to as latent demand. An application of this concept is the investigation of latent demand associated with introduction of incentives for non-peak travel, such as reducing transit fares during non-peak hours. From Figure 5-8, it can be seen that if transit travel were free (zero trip price), the travel demand (quantity of trips demanded) would be QL – Q.

Numerical Example:
A bus company with an existing fleet of one hundred 40-seater buses increases its fleet size by 20% and reduces its $1.00 fare to 90 cents per ride. Determine the impact of the price change on the company's profitability, change in consumer surplus, and the price elasticity of demand. It is assumed that the current load factor of 90% would increase to 95% after the price change. Assume that all buses in the fleet are being used during the peak hours.

Solution:
Existing situation:

Demand = 100 buses × 40 seats × 0.90 = 3600 persons per hour
Revenue = Demand × Price = 3600 × 1.00 = $3600 per hour

After the price drop and supply increase:

Demand = 120 buses × 40 seats × 0.95 = 4560 persons/hour
Revenue = 4560 × 0.90 = $4104 per hour

After the price change, the company gains $4104 - $3600 = $ 504 per hour.

Change in consumer surplus = (1.00 – 0.90) × (3600 + 4560)/2 = $408/hr

Also, Price elasticity of demand, $e_p = \frac{Q_1 - Q_0 \times (P_1 + P_0)/2}{P_1 - P_0 \times (Q_1 + Q_0)/2} = \frac{960 \times 0.95}{0.1 \times 4080} = -2.235$

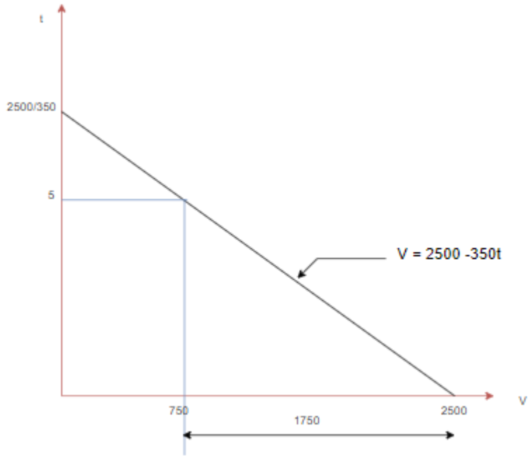
Numerical Example:
The Demand for a transit service between a city and its largest suburb during an off-peak hour V, is given by 2500 -350t where t is the travel time in minutes. At the current time, the transit trip takes an average of 5 minutes. Determine the time elasticity of demand and the latent demand at this travel.
Solution:
$V = 2500 - 350t$

Diff the above equation with respect to t we get,
$\frac{dV}{dt} = -350$

Time elasticity of demand,
$e_t = \frac{dV/V}{dT/t}$
$or, e_t = -350 \times \frac{5}{(2500 - 350 \times 5)}$
$\therefore e_t = -2.3$

Also, at t = 5 minutes,
$V = 2500 - 350 \times 5 = 750$

$\therefore$ *Latent demand* $= 2500 - 750 = 1750$



3.11 Cost Analysis in the Evaluation of Transportation Systems

A complete and balanced evaluation of alternative investment options is possible only by giving due consideration to both benefits and costs associated with each alternative. For this reason, it is essential to have knowledge of the costs of each aspect of the provision of a transportation good or service. This way, the future costs of such aspects can be determined using average cost values or better still, cost models that estimate cost as a function of investment and facility attributes.

In economic theory, three types of costs are encountered: fixed costs (which are independent of the volume of goods produced), variable costs (which depend on the volume of good produced), and total costs (the sum of fixed and variable costs). For each of these costs, it is possible to find the average cost (dividing the total production cost by the number of goods produced) and marginal cost (the incremental cost of producing an additional unit). It is therefore possible to determine average fixed costs, marginal fixed costs, average total costs, etc.

3.11.1 Economic Laws Related to Cost

- *Law of Diminishing Returns*: States that an increase in input of one unit of a factor of production generally causes an increase in output, but only up to a point, after which increasing inputs of that factor will result in progressively less increase in output.

- *Law of Increasing Returns to Scale*: States that in practice, the production of units is often likely to increase at a faster rate than the increase in the factors of production. This may be due to technological features, specialization.

3.11.2 Average Cost

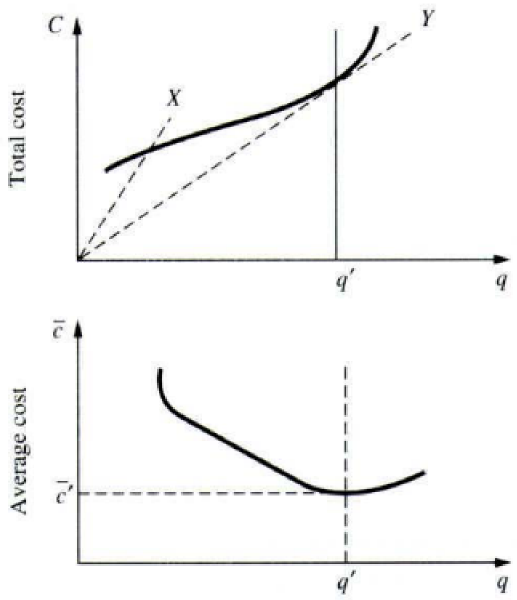
The Average Cost is the per unit cost of production obtained by dividing the total cost by the total output. By per unit cost of production, we mean that all the fixed and variable cost is taken into the consideration for calculating the average cost. Thus, it is also called as Per Unit Total Cost. Mathematically,

$$C = cq = \alpha + \beta(q) \tag{3.8}$$

where,
C = total cost of a product
c = unit cost
q = magnitude of output
α = fixed cost of production
$\beta(q)$ = variable cost of production
Therefore, Average cost:

$$\overline{c} = \frac{C}{q} = \frac{\alpha + \beta(q)}{q} = \frac{\alpha}{q} + \frac{\beta(q)}{q} \tag{3.9}$$

The relationships of the total and average cost functions are shown the figure below. It can be seen that as output q increases, the average cost of production decreases and then increase at higher levels of production. When the production level reaches q' , the average cost is a minimum $\overline{c'}$. The decrease in average cost with increasing output is referred to as economies of scale. In the figure, there is obviously there is economy of scale for production levels between 0 and q' . However, there is no economy of scale beyond q' because the average cost increases. This concept is useful to engineers in deciding whether additional capacity or growth would yield higher profits, and is important in the economic evaluation of transportation system improvements.

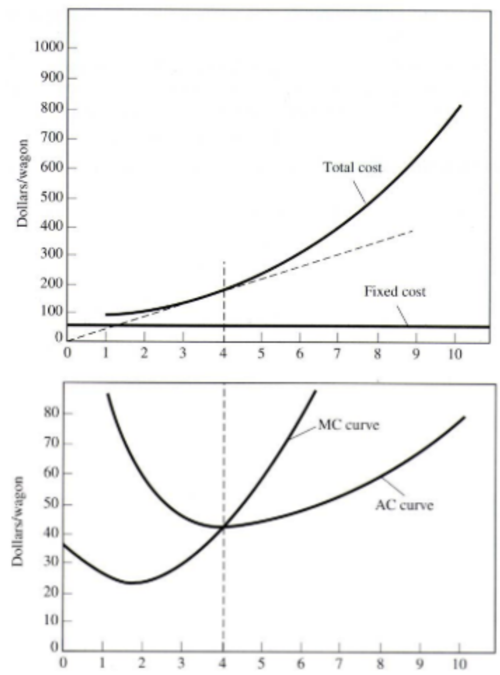


3.11.3 Marginal Costs

The marginal costs of a transportation good or service is the additional cost associated with the production of an additional unit of output. The following example illustrates the concepts of average and marginal costs.

Case Study

(1) Number of wagon/train	(2) Fixed cost, FC	(3) Variable cost, VC	(4) Total cost, TC	(5) Average cost, AC	(6) Marginal cost/unit, MC
1	55	30	85	85.0	
2	55	55	110	55.0	25
3	55	75	130	43.3	20
4	55	105	160	40.0	30
5	55	155	210	42.0	50
6	55	225	280	46.7	70
7	55	315	370	52.9	90
8	55	425	480	60.0	110
9	55	555	610	67.8	130
10	55	705	760	76.0	150



The table above presents the cost of running a train system with variable number of wagons. For each system size, the fixed and variable costs are provided in the first three columns. The total, average and marginal costs are then computed and presented in the next three columns.
The various costs are computed as follows:

$$Total\ cost = TC(x) = FC + VC(x)$$

$$Average\ Total\ Cost = AC(x) = \frac{TC(x)}{x} = \frac{FC}{x} + \frac{VC(x)}{x}$$

$$Marginal\ Cost = MC(x) = TX(x) - TC(x-1)$$

The second graph in the figure above presents the curves corresponding to average cost, marginal cost and total costs for the train system. It can be seen that the point of minimum cost ($40) occurs at the intersection of the average cost (AC) and marginal cost (MC) curves. It is also observed that the projection of this point to the first graph in the figure above corresponds to the point where the gradient of the tangent drawn from the origin has the minimum slope.

When the output is a continuous function, the differential form of the marginal cost is used, in which the marginal cost is the rate of change of total cost with respect to a change in output. In this form the equation is:

$$MC(x) = \frac{dTC(x)}{dx} = \frac{dVC(x)}{dx} \tag{3.10}$$

From the geometry of the AC and MC curves, it is also seen that the average costs is proportional to the slope of a line connecting the origin of the total cost curve with a point on that curve corresponding to he total output. In our example, the slope of such a line begins at infinity at zero output and then decreases to its lowest point, when x = 4. Beyond this point, the slope increases again. On the other hand, the marginal cost curve is the slope of the tangent drawn at any point on the total cost curve.

3.11.4 Costs and Production

In general, a private company will continue to produce and market a good or service as long as it is returning a profit. The net profit (P) is equal to the total revenue (R) minus the total costs (C).

$$P = R - C = pq - cq \tag{3.11}$$

Where p is the selling price of one unit of product q, and c is the production cost of one unit.

The necessary condition for profit maximization is:

$$\frac{dP}{dq} = \frac{dR}{dq} - \frac{dC}{dq} = 0$$

$$or, \ \frac{dP}{dq} = \frac{d(pq)}{dq} - \frac{d(cq)}{dq} = 0$$

$$\therefore \frac{d(pq)}{dq} = \frac{d(cq)}{dq}$$

Let,
Marginal Revenue (MR) $= \frac{dR}{dq} = \frac{d(pq)}{dq}$
Marginal Cost (MC) $= \frac{dC}{dq} = \frac{d(cq)}{dq}$

$\therefore MR = MC$

This equation implies that the firm seeking to maximize profits should produce at the point where marginal revenue equals marginal costs.

3.11.5 Cost Elasticity

The cost elasticity of a good or service is defined as the ratio of percentage change in cost C to a unit percentage change in supply q. The difference between cost elasticity and price elasticity is obvious.

$$e_c = \frac{\%change\ cost}{\%change\ in\ supply} = \frac{\Delta C/C}{\Delta q/q} = \frac{q}{C} \times \frac{\Delta C}{\Delta q} \tag{3.12}$$

In the limit when $\Delta q = 0$, $\frac{\Delta C}{\Delta q} = \frac{dC}{dq}$, then,

$$e_c = \frac{dC/dq}{C/q} = \frac{MC}{AC} \tag{3.13}$$

Numerical Example:
A transport company hauling goods by truck has cost function $C = 1.5q^{1.25}$, where C is the total costs of supply, q.

- Determine the average cost and the marginal cost of production
- Prove that the cost elasticity is 1.25
- Is there an economy of scale?

Solution:

$\overline{c} = \frac{C}{q} = \frac{1.5q^{1.25}}{q} = 15q^{0.25}$

$MC = \frac{dC}{dq} = 1.5 \times 1.25 \times q^{0.25} = 1.875q^{0.25}$

$e = \frac{MC}{AC} = \frac{18.75q^{0.25}}{15q^{0.25}} = 1.25$

$\therefore$ Economy of scale does not exist because the average cost increases with increases q.

3.12 Transportation Projects Evaluation

Construction of highways along with the other associated infrastructures, construction of a new freight center, acquisition of new vehicles for the transit e.t.c. represent transportation projects faced by transportation experts. In the initial stage of these projects, it is necessary to perform their evaluation, that is, as precise as possible, analyze and review economic, environmental, equity, as well as other project impacts. The planners and engineers must properly answer the following questions: (a) Are the transportation project's benefits greater than the projects' costs?; (b) What is the best project's alternative in the case when project has few mutually exclusive alternatives?; (c) How to allocate available funds among competitive transportation projects?; (d) When to start the considered project?

Frequently, financial resources are scarce, and appropriate engineering economic analysis can significantly help planners and decision-makers to allocate available resources properly. In the first step ofany project evaluation, it is necessary to analyze the project's socio-economic context, and to clearly define the project's objectives. In the next steps, the analysts should clearly recognize the type of costs and benefits, compare them and make recommendations to the decision makers.

Transportation projects usually extend over many years. On the other hand, the purchasing power of money decreases over time. The main cause of this phenomenon is the inflation that exists in every society. A discount rate regulates the value ofmoney for time. This rate is used to represent future monetary quantities in terms oftheir today's value. Compounding and discounting are techniques that enable us to compare money values at different points in time. Let us briefly explain compounding technique.

Let us assume that we want to invest \$100 these days, at an annual interest rate (r) of 5%. It will be \$100 + \$5 = \$105 in 1 year. After two years it will be 105\$ + \$0.05 * 105 = \$110.25. After 3 years we will have \$110.25 + \$0.05 * \$110.25 = \$115.7625. Discounting represent reverse operation of compounding. The compounding technique helps us to find the answer to the following question: what is the present value (PV) of a known future amount of money?

In our example, the PV of \$105 next year, when $r = 5\%$, is \$100. The PV equals

$$PV = V_t/(1+r)^t \quad t = 0, 1, 2, ..., n \tag{3.14}$$

Where n is the project duration (in years), r is the rate of discount, and V_t is the value in year t.

There are two interest rates that are used in transportation projects evaluation. The first one is the real interest rate that is exclusive of inflation, while the second one is the nominal interest rate that is inclusive of inflation. Project value is usually expressed as a NPV. This value represents a project's value or cost for its whole life cycle in today's dollars.

When evaluating transportation projects many governments and funding agencies in the world (OECD, World Bank, etc.) require a cost-benefit analysis (CBA) to be performed. In this way, the CBA represents common evaluation language between the governments, funding agencies and the transportation project supporters.

3.13 Cost-Benefit Analysis

The CBA is a method that calculate and compare project's costs and benefits to society over period of time. The CBA monetizes all project's inputs and outputs. In other words, the CBA converts the inputs and the outputs into a monetary values. The CBA helps decision-makers to rank and prioritize various project's alternatives including also alternative "no action" ("no action," or "do nothing" case assumes continued operation of the existing facility, exclusive of any major investments). The specific transportation project should start only when the CBA clearly shows to the decision makers that the total benefits to society outweigh the total costs. When performing CBA, the analysts enumerate all project's costs and benefits to society. In the next step, they assign monetary values to costs and benefits, and discount them to a NPV. All costs, as well as all benefits are added into a single number. The transportation project is evaluated by using total costs, and the total benefits values.

In the first step of the CBA, it is necessary to identify transportation project's alternatives to be evaluated. Alternatives may represent "do nothing" case, rehabilitation of existing facility, construction of a new facility, etc. Transportation projects have consequences over time. The analysts should also define the time period over which the life cycle costs and benefits of all of the alternatives will be calculated.

The project's economic performances are measured by the following indicators:
NPV: Net Present Value
IRR: Internal Rate of Return
B/C: The Benefit-Cost ratio

The majority of experts consider the NPV as the most important CBA indicator. The NPV is defined in the following way:

$$NPV = \sum_{t=0}^{\infty} \frac{B_t - C_t}{(1+r)^t} \tag{3.15}$$

Where,

B_t: Benefits in year t
C_t: Costs in year t
n: Project duration (in years)
r: interest rate
t: year index

Project's benefits and costs are forecast over the project duration.For example,benefits from road investment could be shorter traveling distance, shorter travel time, reduced number of traffic accidents, etc. The road improvement costs could be project design costs, labor costs equipment costs, material costs, etc.

The analysts use a NPV to express a project's worth for its complete life cycle in today's money value. We see that the NPV decreases in r (interest rate) increases. In case when $NPV > 0$, the project may be accepted. However, when $NPV < 0$, the considered project should be rejected.Finally, when $NPV = 0$, we conclude that the considered project adds no monetary value. The final decision about such transportation project should be based on some additional criteria.

The internal rate of return (IRR) is the indicator that also measures the project's performances. The IRR is the discount rate/interest rate at which the $NPV = 0$. We calculate the IRR by solving the following equation:

$$\sum_{t=0}^{\infty} \frac{B_t - C_t}{(1 + IRR)^t} = 0 \tag{3.16}$$

The average values of the observed $IRR's$ in a sample of investment projects sponsored by the European Union (EU) at the end of the 20th century are approximately equal to 15% in the cases of roads and highways, 10% in the cases of railways and underground, and 25% in the cases of ports and airports.

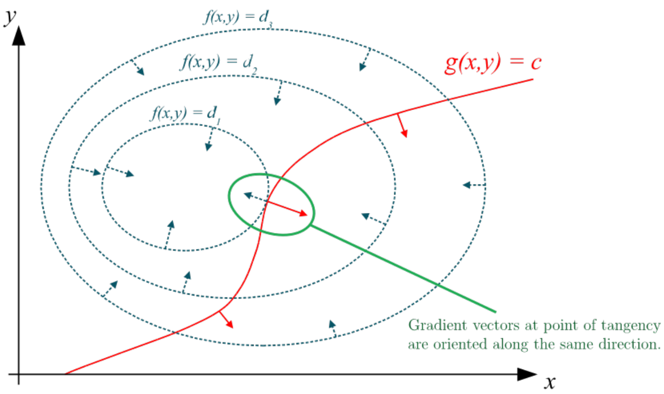
After calculation net benefits B and net costs C, the benefit/cost ratio $\left(\frac{B}{C}\right)$ should be also calculated. (Frequently, it is not easy to estimate future costs, and, especially project's benefits.) The benefit/cost ratio $\left(\frac{B}{C}\right)$ informs us about the improvement in traffic operations (expressed in dollars) per dollar invested.

Analysts and engineers usually perform a sensitivity analysis to conclude how sensitive final results are to changes in hypothesis about the costs, benefits, and discount rate.

The main weakness of the CBA is that all transportation project benefits are evaluated only in monetary terms. It is very complicated to value all the costs and benefits of transportation projects in monetary terms. In other words, in many situations, social and environmental aspects of the considered transportation projects are not treated adequately. For example, many traffic safety programs, actions and projects involve the prevention of loss of life. The logical and ethical question is how should we value a life saved? A pure economic approach would suggest to us that the value of life is equal to the PV of lifetime earnings. Obviously, there are numerous opponents to such an oversimplified and ethically questioned approach.

3.14 Constrained Optimization using Lagrange Multipliers

The general technique for optimizing a function $f = f(x, y)$ subject to a constraint $g(x, y) = c$ is to solve the system $\nabla f = \lambda \nabla g$ and $g(x, y) = c$, for x, y, $and\, \lambda$. Then, we evaluate the function at each point (x, y) that results from a solution to the system to find the optimum values of $f(x, y)$ subject to the constraint.



example: The cost function of an automobile company selling two products x and y is, $C = 6x + \frac{96}{x} + \frac{4y}{x} + \frac{3x}{y}$. The restriction is that the firm is required by contract to produce a minimum quantity of x and y totaling 10. Determine the optimal values of two products and λ. Also, compute the value of function C.

solution:

Let $f(x, y) = C = 6x + \frac{96}{x} + \frac{4y}{x} + \frac{3x}{y}$

$g(x, y) = x + y$

Objective: Minimize $f(x, y)$

subject to: $g(x, y) = x + y = 10$

Now,

$$\nabla f = \vec{i}\frac{\partial f(x,y)}{\partial x} + \vec{j}\frac{\partial f(x,y)}{\partial y}$$
$$= \left(6 - \frac{96}{x^2} - \frac{y}{x^2} + \frac{3}{y}\right)\vec{i} + \left(\frac{4}{x} - \frac{3x}{y^2}\right)\vec{j}$$

$$\nabla g = \vec{i}\frac{\partial(x+y)}{\partial x} + \vec{j}\frac{\partial(x+y)}{\partial y}$$
$$= \vec{i} + \vec{j}$$

We have,

$\nabla f = \lambda \nabla g$
$\left(6 - \frac{96}{x^2} - \frac{y}{x^2} + \frac{3}{y}\right)\vec{i} + \left(\frac{4}{x} - \frac{3x}{y^2}\right)\vec{j} = \lambda(\vec{i} + \vec{j})$

Equating along corresponding components we get,

$6 - \frac{96}{x^2} - \frac{y}{x^2} + \frac{3}{y} = \lambda$ ....(1)
$\frac{4}{x} - \frac{3x}{y^2} = \lambda$ ....(2)

From eqn (1) and (2) we get,

$6 - \frac{96}{x^2} - \frac{y}{x^2} + \frac{3}{y} = \frac{4}{x} - \frac{3x}{y^2}$
or, $\frac{4}{x} + \frac{96}{x^2} + \frac{4y}{x^2} - \frac{3x}{y^2} - \frac{3}{y} = 6$
or, $\frac{4x+96+4y}{x^2} - \frac{3x+3y}{y^2} = 6$
or, $\frac{4(x+y)+96}{x^2} - \frac{3(x+y)}{y^2} = 6$
or, $\frac{4\times10+96}{x^2} - \frac{3\times10}{y^2} = 6$ $[\because x + y = 10]$
or, $\frac{136}{x^2} - \frac{30}{y^2} = 6$
or, $136y^2 - 30x^2 - 6x^2y^2 = 0$
or, $68y^2 - 15x^2 - 3x^2y^2 = 0$
or, $68(1-x)^2 - 15x^2 - 3(1-x)^2y^2 = 0$
or, $-3x^4 + 60x^3 - 247x^2 - 1360x + 6800 = 0$................(3)

Solving eqn (3) we get,

either, $x = 4.419$
or, $x = -4.707$
or, $x = 10.144 + 2.464\, i$
or, $x = 10.144 - 2.464\, i$

We will neglect the complex roots and only consider real roots for our solution. Also, since the question states that the values of $x\ and\ y$ are quantity of automobiles, they cannot be negative and are integer values, i.e, $x, y \geq 0 \in Z$.

Thus,

$x = 4$ and $y = 10 - x = 10 - 4 = 6$

$C = 6 \times 4 + \frac{96}{4} + \frac{4\times6}{4} + \frac{3\times4}{6}$
$\therefore C = 56$

$\therefore \lambda = \frac{4}{x} - \frac{3x}{y^2} = \frac{4}{4} - \frac{3\times4}{6^2} = \frac{2}{3}$

$\therefore$, $x = 4$, $and\ y = 6$ is the optimal solution to minimize the given cost function subject to the specified constraints.

3.15 Numerical Examples:

Example #1: Selecting a transportation mode

An individual is planning to take a trip between the downtown area of two cities, A and B, which are 400 miles apart. There are three options available:

1. **Travel by air:** This trip will involve driving to the airport near city A, parking, waiting at the terminal, flying to airport B, walking to a taxi stand, and taking a taxi to the final destination.
2. **Travel by auto:** This trip will involve driving 400 miles through several congested areas, parking in the downtown area, and walking to the final destination.
3. **Travel by rail:** This trip will involve taking a cab to the railroad station in city A, a direct rail connection to the downtown area in city B, and a short walk to the final destination.

Since this is a business trip, the person making the trip is willing to pay up to $25 for each hour of travel time reduced by a competing mode. (For example, if one mode is two hours faster than another, the traveler is willing to pay $50 more to use the faster mode.) After examining all direct costs involved in making the trip by air, auto, or rail (including parking, fuel, fares, tips, and taxi charges) the traveler concludes that the trip by air will cost $250 with a total travel time of five hours, the trip by auto will cost $200 with a total travel time of eight hours and the trip by rail will cost $150 with a total travel time of 12 hours.

- Which mode is selected based on travel time and cost factors alone?
- What other factors might be considered by the traveler in making a final selection?

Solution: Since travel time is valued at $25/hr, the following costs would be incurred:

$$Air : 250 + 25 * 5 = \$375$$
$$Auto : 200 + 25 * 8 = \$400$$
$$Rail : 150 + 25 * 12 = \$450$$

In this instance, the air alternate reflects the lowest cost and is the selected mode. However, the traveler may have other reasons to select another alternative. These may include the following considerations.

Safety: While each of these modes is safe, the traveler may feel "safer" in one mode than another. For example, rail may be preferred because of concerns regarding air safety issues.

Reliability: If it is very important to attend the meeting, the traveler may select the mode that will provide the highest probability of an on-time arrival. If the drive involves travel through work zones and heavily congested areas, rail or air would be preferred. If potential air delays are likely due to congestion, flight cancellations, or inclement weather, another mode may be preferred.

Convenience : The number of departures and arrivals provided by each mode could be a factor. For example, if the railroad provides only two trains/day and the airline has six flights/day, the traveler may prefer to go by air.

Example #2: Computing the Toll to Maximize Revenue Using a Demand Curve

A toll bridge carries 5000 veh/day. The current toll is 150 cents. When the toll is increased by 25 cents, traffic volume decreases by 500 veh/day. Determine the amount of toll that should be charged such that revenue is maximized. How much additional revenue will be received? (Assume a linear demand function)

Solution:

Let V be the number of vehicle per day using the bridge and x be the increment in toll, then using the slope-intercept form of a straight line the demand function corresponding to the increment in toll can be defined by,

$V = mx + C$ (1)
Where, m is the slope and C is the y intercept.

When there is no increase in toll, that is, $x = 0$ then $V = 5000$,

Substituting these values in the above demand function we get,
$C + m \times 0 = 5000$
$\therefore C = 5000$

Also, 25 cents increase in toll decreases the traffic volume of the bridge by 500 vehicles per day(vpd), that is,

$m = \frac{change\ in\ vehicles\ per\ day}{change\ in\ toll}$
$or, m = -\frac{500}{25}$ ($\because$ increase in toll causes decrease in vpd)
$\therefore m = -20$

Substituting the values of m and C in eqn (1) we get,

$V = 5000 - 20x$

The toll after increment, $T = 150 + x$

Revenue is the product of toll and volume,

$$
\begin{aligned}
R &= V * T \\
&= (5000 - 20x)(150 + x) \\
&= 750,000 + 2000 - 20x^2
\end{aligned}
$$

For maximum value of R,

$\frac{dR}{dx} = 0$
$2,000 - 40x = 0$
$\therefore x = 50 cents$

Toll for maximum revenue (T_{max}) $= 150 + 50 = 200 cents$
The additional revenue, AR

$$
\begin{aligned}
&= V_{max} * T_{max} - V_{current} * T_{current} \\
&= (5000 - 20 * 50) * 200 - 5000 * 150 \\
&= 50,0000\, cents \\
&\therefore AR = \$500
\end{aligned}
$$

Example #3:
The demand function for transportation from the suburbs to downtown in a large city is as follows:

$$Q = T^{-0.3}\, C^{-0.2}\, A^{0.1}\, I^{-0.25}$$

Where,

Q = number of transit trips
T = travel time on transit(hours)
C = fare on transit (dollars)
A = cost of automobile trip (dollars)
I = average income (dollars)

1. There are currently 10,000 persons per hour riding the transit system, at a flat fare of $1 per ride. What would be the change in ridership with a 90 cent fare? What would be company gain/lose per hour?
2. By auto, the trip costs $ 3 (including parking). If the parking charge were raised by 30 cents, how would it affect the transit ridership?
3. The average income of auto riders is $15,000 per year. What raise in salary will rider require to cover their costs in view of the change in parking charge noted in the second question?

Solution:

Part 1:
This is essentially a Kraft model. The price elasticity of demand for transit trips,
$e_p = -0.2$
% Reduction in fare $(\frac{dC}{C}) = \frac{100-90}{100} = 10\%$

% Increase of Patronage $(\frac{dQ}{Q}) = e_p \times \frac{dC}{C} = 0.2 \times 0.1 = 0.02 = 2\%$
Now,
Initial Patronage $(Q_1) = 10000$
Final Patronage $(Q_2) = 10000 + 0.02 \times 10000 = 10200$

Also,
Revenue before increment of patronage $(R_1) = Q_1 \times 1 = 10000$
Revenue after increment of patronage $(R_2) = Q_2 \times 0.9 = 9180$

$\therefore$ The company will lose $820 per hour.

Part 2:
Similarly, the automobile price cross-elasticity of demand, $e_{cp} = 0.1$
% rise in auto costs $(\frac{dA}{A}) = \frac{3.3-3}{3} = 10\%$

% increase of patronage $(\frac{dQ}{Q}) = e_{cp} \times \frac{dA}{A} = 0.1 \times 0.1 = 1\%$
Now,
Final patronage $= 10000 + 0.01\ times\ 10000 = 10100$

$\therefore$ The transit patronage increases from 10000 riders to 10100 riders.

Part 3:
Similarly, the income elasticity of demand, e_i = -0.25
$\frac{dQ}{Q} = 1\% = 0.01$

We have,

$\frac{dI}{I} = \frac{\frac{dQ}{Q}}{e_i} = \frac{0.01}{0.25} = 0.04 = 4\%$

So, a 4% increase in income would cover a 30 cent increase in auto cost. If a the average income is % 15000, $600 raise in salary would change the minds of those auto drivers who were planning to ride the transit system.

3.16 Review Problems

1. Describe the causes of shift in transportation demand curve.
2. Describe the causes of shift in transportation supply curve.
3. Define the law of diminishing returns. Illustrate the concept with a relevant example.
4. Describe the equilibrium state in a transportation demand and supply curve. Illustrate the concept with a relevant example.
5. Discuss why the average cost curve is a U-shaped curve and also explain the reason for marginal cost curve to intersect at the minimum of the average cost curve. Illustrate the concept with a lucid diagram.
6. Derive the relation of point elasticity for a linear demand curve and illustrate it's behavior along the curve.
7. Show that the price elasticity of demand for travel is constant for the Kraft's Demand Model.
8. Derive the condition for profit maximization.
9. It is estimated that the demand for a newly constructed parking facility will be related to the price of usage as follows: $V = 1500 - 25P$, Where V is the number of vehicles using the parking lot per day and P is the average daily parking fee in dollars. For the first month of operation, parking at the facility is free.
   (a) How many vehicles would be expected to park at the facility during the first month? *[answer: 1500 vehicles/day]*
   (b) After the second month, when a $10 daily fee is charged, how many vehicles would be expected to use the facility, and what would be the loss in consumer surplus? *[answer: $13.750]*
10. A 20% increase in downtown parking costs resulted in a 5% reduction in downtown auto trips and a 20% increase in transit patronage for downtown routes. Determine the elasticity of auto and transit demand with respect to parking costs. *[answer: 1, -0.25]*

11. The aggregate demand function for a bus transit service servicing a newly developed suburban area is represented by the equation, $V = 300 - 40P^2$, where V is the number of trips made per month and P is the average price of the ticket for the trip. In a given month, the average price was \$0.75. What is the price elasticity of demand of the bus transit service for the same month? *[answer: -0.162]*

12. Consideration is being given to increasing the toll on a bridge now carrying 4500 veh/day. The current toll is \$1.25/veh. It has been found from past experience that the daily traffic volume will decrease by 400 veh/day for each 25 cents increase in toll. In order to maximize revenues, what would the new toll charge be per vehicle, and what would the traffic in veh/day be after the toll increase? How much additional revenue will be generated? *[answer: $T_{max} = 203.125\, cents$, V = 3250 vpd, and Additional Revenue = \$ 976.5625]*

13. Show that the price elasticity of demand for an exponential demand function, $V = \alpha e^{\beta p}$, is βp. (Where V = demand, p = price of the transportation service, and α *and* β are constants)

Chapter 4

Insurance

4.1 Introduction

There is no single definition of insurance. Insurance can be defined from the viewpoint of several disciplines, including law, economics, history, actuarial science, risk theory, and sociology. But each possible definition will not be examined at this point. Instead, we will examine the common elements that are typically present in any insurance plan. However, before proceeding, a working definition of insurance—one that captures the essential characteristics of a true insurance plan—must be established.

After careful study, the Commission on Insurance Terminology of the American Risk and Insurance Association has defined insurance as follows. ***Insurance is the pooling of fortuitous losses by transfer of such risks to insurers, who agree to indemnify insureds for such losses, to provide other pecuniary benefits on their occurrence, or to render services connected with the risk.*** Although this lengthy definition may not be acceptable to all risk managers and insurance scholars, it is useful for analyzing the common elements of a true insurance plan.

4.2 Basic Characteristics of Insurance

Based on the preceding definition, an insurance plan or arrangement typically includes the following characteristics:

4.2.1 Pooling of Losses

Pooling or the sharing of losses is the essence of insurance. *Pooling is the spreading of losses incurred by the few over the entire group, so that in the process, average loss is substituted for actual loss.* In addition, pooling involves the grouping of a large number of exposure units so that the law of large numbers can operate to provide a substantially accurate prediction of future losses. Ideally, there should be a large number of similar, but not necessarily identical, exposure units that are subject to the same perils. Thus, pooling implies (1) the sharing of losses by the entire group and (2) the prediction of future losses with some accuracy based on the law of large numbers.

The primary purpose of pooling, or the sharing of losses, is to reduce the variation in possible outcomes as measured by the standard deviation or some other measure of dispersion, which reduces risk. For example, assume that two freight business owners each own an identical storage building valued at \$50,000. Assume there is a 10% chance in any year that each building will be destroyed by a peril, and that a loss to either building is an independent event. The expected annual loss for each owner is \$5,000 as shown below:
Expected Loss $= 0.90 \times \$0 + 0.10 \times \$50000 = \$5000$

A common measure of risk is the standard deviation, which is the square root of the variance. The standard deviation (SD) for the expected value of the loss is \$15,000, as shown below:

$$\text{SD} = \sqrt{0.9 \times (0 - 5000)^2 + 0.1 \times (50000 - 5000)^2} = \$15000$$

Suppose instead of bearing the risk of loss individually, the two owners decide to pool (combine) their loss exposures, and each agrees to pay an equal share of any loss that might occur. Under this scenario, there are four possible outcomes:

Possible Outcomes	Probability
Neither Building is destroyed	$0.9 \times 0.9 = 0.81$
First building is destroyed and second building is intact	$0.1 \times 0.9 = 0.09$
First building is intact and second building is destroyed	$0.9 \times 0.1 = 0.09$
Both buildings are destroyed	$0.1 \times 0.1 = 0.01$

If neither building is destroyed, the loss for each owner is \$0. If one building is destroyed, each owner pays \$25,000. If both buildings are destroyed, each owner must pay \$50,000. The expected loss for each owner remains \$5,000 as shown below:

Expected loss $= 0.81 \ltimes \$0 + 0.09 \ltimes \$25000 + 0.09 \ltimes \$25000 + 0.01 \times \$50000 = \$5000$

Note that while the expected loss remains the same, the probability of the extreme values, \$0 and \$50,000, have declined. The reduced probability of the extreme values is reflected in a lower standard deviation as shown below:

$$\text{SD} = \sqrt{0.81(0-5000)+0.09(25000-5000)+0.09(25000-5000)+0.01(50000-5000)}$$
$$\text{SD} = \$10607$$

Thus, as additional individuals are added to the pooling arrangement, the standard deviation continues to decline while the expected value of the loss remains unchanged. For example, with a pool of 100 insureds, the standard deviation is \$1,500; with a pool of 1,000 insureds, the standard deviation is \$474; and with a pool of 10,000, the standard deviation is \$150.

In addition, by pooling or combining the loss experience of a large number of exposure units, an insurer may be able to predict future losses with greater accuracy. From the viewpoint of the insurer, if future losses can be predicted, objective risk is reduced. Thus, another characteristic often found in many lines of insurance is risk reduction based on the law of large numbers.

The law of large numbers states that the greater the number of exposures, the more closely will the actual results approach the probable results that are expected from an infinite number of exposures. For example, if you flip a balanced coin into the air, the a priori probability of getting "heads" is 0.5. If you flip the coin only 10 times, you may get heads eight times. Although the observed probability of getting heads is 0.8, the true probability is still 0.5. If the coin were flipped 1 million times, however, the actual number of heads would be approximately 500,000. Thus, as the number of random tosses increases, the actual results approach the expected results.

4.2.2 Payment of Fortuitous Losses

A second characteristic of private insurance is the payment of fortuitous losses. Most insurance policies exclude intentional losses. *A fortuitous loss is one that is unforeseen and unexpected by the insured and occurs as a result of chance.* In other words, the loss must be accidental. The law of large numbers is based on the assumption that losses are accidental and occur randomly. For example, a person may slip on an icy sidewalk and break a leg. The loss would be fortuitous.

4.2.3 Risk Transfer

Risk transfer is another essential element of insurance. With the exception of self-insurance, a true insurance plan always involves risk transfer. *Risk* *transfer means that a pure risk is transferred from the insured to the insurer, who typically is in a stronger financial position to pay the loss than the insured.* From the viewpoint of the individual, pure risks that are typically transferred to insurers include the risk of premature death, excessive longevity, poor health, disability, destruction and theft of property, and personal liability lawsuits.

4.2.4 Indemnification

A final characteristic of insurance is indemnification for losses. *Indemnification means that the insured is restored to his or her approximate financial position prior to the occurrence of the loss.* Thus, if your home burns in a fire, a homeowners policy will indemnify you or restore you to your previous position. If you are sued because of the negligent operation of an automobile, your auto liability insurance policy will pay those sums that you are legally obligated to pay. Similarly, if you become seriously disabled, a disability-income insurance policy will restore at least part of the lost wages.

4.3 Characteristics of an Ideally Insurable Risk

Private insurers generally insure only pure risks. However, some pure risks are not privately insurable. From the viewpoint of a private insurer, an insurable risk ideally should have certain characteristics. There are ideally six characteristics of an insurable risk:

4.3.1 Large Number of Exposure Units

The first requirement of an insurable risk is a large number of exposure units. Ideally, there should be a large group of roughly similar, but not necessarily identical, exposure units that are subject to the same peril or group of perils. For example, a large number of wood frame dwellings in a city can be grouped together for purposes of providing property insurance on the dwellings.

The purpose of this first requirement is to enable the insurer to predict losses based on the law of large numbers. Loss data can be compiled over time, and losses for the group as a whole can be predicted with some accuracy. The loss costs can then be spread over all insureds in the underwriting class.

4.3.2 Accidental and Unintentional Loss

A second requirement is that the loss should be accidental and unintentional; ideally, the loss should be unforeseen and unexpected by the insured and outside of the insured's control. Thus, if an individual deliberately causes a loss, he or she should not be indemnified for the loss.

There are several reasons for this requirement. First, the loss should be accidental because the law of large numbers is based on the random occurrence of events. A deliberately caused loss is not a random event because the insured knows when the loss will occur. Thus, prediction of future experience may be highly inaccurate if a large number of intentional or nonrandom losses occur. Second, moral hazard is increased if the insured deliberately intends to cause a loss. Finally, it is poor public policy to allow insureds to collect for intentional losses.

4.3.3 Determinable and Measurable Loss

A third requirement is that the loss should be both determinable and measurable. This means the loss should be definite as to cause, time, place, and amount. Life insurance, in most cases, meets this requirement easily. The cause and time of death can usually be readily determined, and if the person is insured, the face amount of the life insurance policy is the amount paid.

Some losses, however, are difficult to determine and measure. For example, under a disability-income policy, the insurer promises to pay a monthly benefit to the disabled person if the definition of disability stated in the policy is satisfied. Some dishonest claimants may deliberately fake sickness or injury to collect from the insurer. Even if the claim is legitimate, the insurer must still determine whether the insured satisfies the definition of disability stated in the policy. Sickness and disability are highly subjective, and the same event can affect two persons quite differently. For example, two accountants who are insured under separate disability-income contracts may be injured in an auto accident, and both may be classified as totally disabled. One accountant, however, may be more determined to return to work. If that accountant undergoes rehabilitation and returns to work, the disability-income benefits will terminate. Meanwhile, the other accountant would still continue to receive disability-income benefits according to the terms of the policy. In short, it is often difficult to determine when a person is actually disabled. However, all losses ideally should be both determinable and measurable.

The basic purpose of this requirement is to enable an insurer to determine if the loss is covered under the policy, and if it is covered, how much should be paid. For example, assume that Shannon has an expensive fur coat that is insured under a homeowners policy. It makes a great deal of difference to the insurer if a thief breaks into her home and steals the coat, or the coat is missing because her husband stored it in a drycleaning establishment but forgot to tell her. The loss is covered in the first example but not in the second.

4.3.4 No Catastrophic Loss

The fourth requirement is that ideally the loss should not be catastrophic. This means that a large proportion of exposure units should not incur losses at the same time. As we stated earlier, pooling is the essence of insurance. If most or all of the exposure units in a certain class simultaneously incur a loss, then the pooling technique breaks down and becomes unworkable. Premiums must be increased to prohibitive levels, and the insurance technique is no longer a viable arrangement by which losses of the few are spread over the entire group.

Insurers ideally wish to avoid all catastrophic losses. In reality, however, that is impossible, because catastrophic losses periodically result from floods, hurricanes, tornadoes, earthquakes, forest fires, and other natural disasters. Catastrophic losses can also result from acts of terrorism.

Several approaches are available for meeting the problem of a catastrophic loss. First, reinsurance can be used by which insurance companies are indemnified by reinsurers for catastrophic losses. Reinsurance is an arrangement by which the primary insurer that initially writes the insurance transfers to another insurer (called the reinsurer) part or all of the potential losses associated with such insurance. The reinsurer is then responsible for the payment of its share of the loss.

Second, insurers can avoid the concentration of risk by dispersing their coverage over a large geographical area. The concentration of loss exposures in a geographical area exposed to frequent floods, earthquakes, hurricanes, or other natural disasters can result in periodic catastrophic losses. If the loss exposures are geographically dispersed, the possibility of a catastrophic loss is reduced.

Finally, financial instruments are now available for dealing with catastrophic losses. These instruments include catastrophe bonds, which are designed to help fund catastrophic losses.

4.3.5 Calculable Chance of Loss

Another requirement is that the chance of loss should be calculable. The insurer must be able to calculate both the average frequency and the average severity of future losses with some accuracy. This requirement is necessary so that a proper premium can be charged that is sufficient to pay all claims and expenses and yields a profit during the policy period.

Certain losses, however, are difficult to insure because the chance of loss cannot be accurately estimated, and the potential for a catastrophic loss is present. For example, floods, wars, and cyclical unemployment occur on an irregular basis, and prediction of the average frequency and severity of losses is difficult. Thus, without government assistance, these losses are difficult for private carriers to insure.

4.3.6 Economically Feasible Premium

A final requirement is that the premium should be economically feasible. The insured must be able to afford the premium. In addition, for the insurance to be an attractive purchase, the premiums paid must be substantially less than the face value, or amount, of the policy.

To have an economically feasible premium, the chance of loss must be relatively low. One view is that if the chance of loss exceeds 40 percent, the cost of the policy will exceed the amount that the insurer must pay under the contract. For example, an insurer could issue a $1,000 life insurance policy on a man who is age 99, but the pure premium would be close to that amount, and an additional amount for expenses would also have to be added. The total premium would exceed the face amount of insurance.

Based on the preceding requirements, most personal risks, property risks, and liability risks can be privately insured because the ideal characteristics of an insurable risk generally can be met. In contrast, most market risks, financial risks, production risks, and political risks are difficult to insure by private insurers. These risks are speculative, and calculation of a correct premium may be difficult because the chance of loss cannot be accurately estimated. For instance, insurance that protects a retailer against loss because of a change in consumer tastes, such as a style change, generally is not available. Accurate loss data are not available. Thus, it would be difficult to calculate an accurate premium. The premium charged may or may not be adequate to pay all losses and expenses. Since private insurers are in business to make a profit, certain risks are difficult to insure because of the possibility of substantial losses.

4.4 Types of Insurance

Insurance can be classified as either private or government insurance. Private insurance includes life and health insurance as well as property and liability insurance. Government insurance includes social insurance programs and other government insurance plans.

4.4.1 Private Insurance

Life Insurance

Life insurance pays death benefits to designated beneficiaries when the insured dies. The benefits pay for funeral expenses, uninsured medical bills, estate taxes, and other expenses. The death proceeds can also provide periodic income payments to the deceased's beneficiary. Life insurers also sell annuities, individual retirement account (IRA) plans, 401(k) plans, and individual and group retirement plans. Some life insurers also sell (1) individual and group health insurance plans that cover medical expenses because of sickness or injury, (2) disability income plans that replace income lost during a period of disability, and (3) long-term care policies that cover care in nursing facilities.

Health Insurance

Although many life insurers we described also sell some type of individual or group health insurance plan, the health insurance industry overall is highly specialized and controlled by a relatively small number of insurers. These companies include Blue Cross Blue Shield Association, AETNA, United Health Group, and Well Point. Medical expense plans pay for hospital and surgical expenses, physician fees, prescription drugs, and a wide variety of additional medical costs.

Property and Liability Insurance

Property insurance indemnifies property owners against the loss or damage of real or personal property caused by various perils, such as fire, lightning, windstorm, or tornado. Liability insurance covers the insured's legal liability arising out of property damage or bodily injury to others; legal defense costs are also paid.

Property and liability insurance is also called property and casualty insurance. In practice, nonlife insurers typically use the term property and casualty insurance (rather than property and liability insurance) to describe the various coverages and operating results. Casualty insurance is a broad field of insurance that covers whatever is not covered by fire, marine, and life insurance; casualty lines include auto, liability, burglary and theft, workers compensation, and health insurance.

4.4.2 Government Insurance

Numerous government insurance programs are in operation at the present time. Social insurance is the one discussed in this chapter, whereas other government insurance also exist but differ as per the national and state policy of a country.

Social Insurance

Social insurance programs are government insurance programs with certain characteristics that distinguish them from other government insurance plans. These programs are financed entirely or in large part by mandatory contributions from employers, employees, or both, and not primarily by the general revenues of government. The contributions are usually earmarked for special trust funds; the benefits, in turn, are paid from these funds. In addition, the right to receive benefits is ordinarily derived from or linked to the recipient's past contributions or coverage under the program; the benefits and contributions generally vary among the beneficiaries according to their prior earnings, but the benefits are heavily weighted in favor of low-income groups. Moreover, most social insurance programs are compulsory. Covered workers and employers are required by law to pay contributions and participate in the programs. Finally, eligibility requirements and benefit rights are usually prescribed exactly by statute, leaving little room for administrative discretion in the award of benefits. %

4.5 Readings from Insurance Regulation, 2049 (1993)

The following are the readings extracted from *"Insurance Regulation, 2049"*.

4.5.1 Definitions

Unless the subject or context otherwise requires, in this Regulation :

- "Act" means the Insurance Act, 2049
- Certificate" means an Insurer Registration Certificate provided to an Insurer for operating the Insurance Business by registering him as an Insurer pursuant to Rule 8
- "Premium" means the Insurance Premium to be collected by the Insurer from the insured i consideration of the Insurance Business
- "Advisory Committee" means the Insurance Tariff Advisory Committee constituted pursuant to section 41 of the Act

4.5.2 Categories of Insurance Business

Subject to the provisions made in the Act and this Regulation, the Insurance Business to be operated by an Insurer shall be divided into the following categories :

- Life Insurance Business
- Non-Life Insurance Business
- Re-Insurance Business

Notwithstanding anything contained in sub-rule (1), Nepal Government may prescribe other categories of Insurance Business as required on the advise of the Board.

Life Insurance Business

1. The Insurer may operate the following Insurance Business under the Life Insurance Business:
   (a) Whole Life Insurance
   (b) Endowment Life Insurance
   (c) Term Life Insurance
2. Notwithstanding anything contained in sub-rule (1), the Board may prescribe other categories of the Life Insurance Business as required.
3. The conditions and privileges of the Life Insurance Policy to be executed pursuant to this Rule shall be as specified by the Board.

Non-Life Insurance Business

1. The Insurer may operate the following Insurance Business under the Non-Life Insurance Business :
   (a) Fire Insurance
   (b) Motor Insurance
   (c) Marine Insurance
   (d) Engineering and Contractor's Risk Insurance
   (e) Aviation Insurance
   (f) Miscellaneous Insurance
2. Notwithstanding anything contained in sub-rule (1), the Board may prescribe other categories of Non-Life Insurance Business as required.
3. The conditions and privileges of the Non-Life Insurance Policy to be executed pursuant to this rule shall be as specified by the Board.

Re-Insurance Business

1. The Insurer may re-insure the risks which are in excess from the risks assumed by it.

2. The Categories of Re-insurance Business to be made pursuant to sub-rule (1) and other arrangement shall be as specified by the Board

4.5.3 Provisions Relating to a Surveyor

Application to be submitted for surveyor's license

Any person or any corporate body having a qualification, as mentioned in regulation, and desirous to work as a Surveyor pursuant to sub-section (1) of Section # 30A of the Act, shall submit an application to the office of the Board in the format of Schedule -12.

Surveyor's License to be provided

1. After receiving an application for the Surveyor's license pursuant to rule-26, the Board shall make an inquiry whether the applicant is qualified or not pursuant to Rule 28, and if it deems appropriate to provide the Surveyor's license to him it shall register his name as a Surveyor in the registration-book pursuant to Schedule - 13.

2. After making the registration of the name of the applicant on the registration-book pursuant to sub-rule (1), the Board shall provide a Surveyor's license to the applicant in the format of Schedule - 14 by collecting from the applicant a fee of twelve thousand Rupees for the class A license, that of nine thousand Rupees for the class B license, that of seven thousand rupees for the class C license and that of five thousand Rupees for the class D license.

Qualifications of a Surveyor

1. An applicant desirous of making an application for the Surveyor's License pursuant to Rule 26 shall have possessed any one of the following qualifications :

   (a) Having gained at least ten years of work experiences on the Insurance Business, holding an officer level post at the office of any Insurer

   (b) Having possessed at least a Bachelor Degree in Engineering subject

   (c) Having possessed at least a Bachelor Degree in Insurance subject from a Chartered Insurance Institute of international standard or from an organization recognized by such institute, or

(d) Having passed the Chartered Accountancy Examination

2. An applicant having possessed the qualification as referred to in clause (a) or (b) or (c) or (d) of sub-rule (1) shall, prior to obtaining the Surveyor's License, have obtained a certificate indicating his participation in, and completion of, the surveyor training conducted by the Board

Classification of the Surveyor

The Surveyors who have been working as Surveyors after having obtained the Surveyor's license and completed the following period shall be classified as follows and provided with the Surveyor's License, after the commencement of this Regulation:

- A Surveyor who has regularly worked as a Surveyor for a period more than fifteen years, Class "A"

- A Surveyor who has regularly worked as a Surveyor for a period from ten to fifteen years, Class "B"

- A Surveyor who has regularly worked as a Surveyor for a period From five to ten years, Class "C"

- A Surveyor who has regularly worked as a Surveyor for a period of five years, Class "D"

Provisions relating to the renewal of Surveyor's License

1. The Surveyor shall submit an application to the office of the Board in the format of Schedule-15 along with the renewal fee of →twelve thousand Rupees for "A" class license, nine thousand Rupees for "B" class license, seven thousand Rupees for "C" class license and five thousand Rupees for "D" class license within the time-limit pursuant to subsection (1) of Section 31 of the Act for the renewal of the License. On receipt of such application, the Board shall renew the Surveyor's license

2. If a Surveyor has submitted an application to the office of the Board stating the reasons for inability to submit an application for the renewal of his license within the time-limit pursuant to sub-rule (1), and if the reasons are found to be appropriate, the Board may renew the license by receiving an additional fee of five hundred Rupees for first two months from the date of expiry of the renewal time-limit and after that fifty Rupees per day for up to four months

Limitation of Survey

1. The limitation relating to survey which a surveyor of each category pursuant to Rule 28A can make shall be as prescribed by the Board

2. The Code of Conduct of the Surveyor shall be as prescribed by the Board

In the case of cancellation of Surveyor's License

If the Surveyor's license is cancelled pursuant to Section 33 of the Act, no other Surveyor's license shall be provided to him to work as a Surveyor up to a period of five years from the date of such cancellation.

4.5.4 Provisions relating to payment against Insurance Claim

Process against the payment of Life Insurance Claim

1. The Insurer shall issue a discharge voucher in the name of the Insured who has already paid the last installment of the Life Insurance Premium requesting him to come to collect payment against the claim along with the Insurance Policy and other documents required for making payment against such Life Insurance claim within fifteen days from the date of payment of such installment

2. In case an Insured submits the Insurance Policy and other documents including the discharge voucher to the Insurer for the payment of claim against the Life Insurance claim pursuant to sub-rule (1), the Insurer shall conduct an inquiry as required and make a payment against the Life Insurance claim within seven days from the date of expiry of the period of the Life Insurance Policy

3. In case any person who has taken up an Insurance Policy dies before the expiry of the period of the Insurance Policy, the person designated by him, if any, and in case no person has been designated, the nearest heir from among the persons mentioned in sub-section (1) of Section 38 of the Act shall submit an application for the payment against the claim to the Insurer to receive the amount of the Life Insurance stating the details as follows :

   (a) The details relating to the claim
   (b) A Certificate of death of the insured
   (c) In case the insured has died in an accident and if such risk is covered by the Life Insurance, the postmortem report of the government physician relating to the cause of death, and if there is no such report, a report of the police

(d) A certificate of relationship with the insured
(e) The documents regarding the certification of the age in case the age has not been certified
(f) Other details specified by the Board

4. After the receipt of the application pursuant to sub-rule (3), the Insurer shall make an inquiry into the details including the documents submitted regarding to the claim of Life Insurance, and shall examine other matters also if necessary, and shall determine the liability within fifteen days from the date of receipt of such documents by it and shall issue the discharge voucher in the name of the applicant requesting him to come to collect the payment against the claim. The Insurer shall make the payment against the Insurance claim within fifteen days from the date of receipt of the discharge voucher from the applicant

5. If it is found, while making an inquiry into the details pursuant to sub-rule (4) that the Insurance claim need not to be paid by determining the liability, the Insurer shall provide a written information to the applicant clearly stating the reasons thereof.

Process of payment against Non-Life Insurance Claim

1. If any claim has to be made under the Insurance Policy by an Insured who has taken up a Non-Life Insurance Policy, the Insured shall submit an application to the Insurer stating all the details relating to it

2. On receipt of an application of the Insured for the payment against the Insurance claim of the Non-Life Insurance Pursuant to sub-rule (1), the Insurer shall immediately designate a Surveyor to make necessary inquiry, if necessary

3. The Surveyor deputed pursuant to sub-rule (2) shall make necessary inquiry and shall determine the liability of the Insurer within fifteen days and shall submit a report to the Insurer including the comprehensive details and also inform the Insured relating to it mentioning the amount to be received by the Insured subject to the terms and conditions and facilities of the Insurance Policy

4. The Insurer shall determine the liability and shall provide the payment against the claim of the Non-Life Insurance to the Insured generally within thirty-five days from the submission of the report by the surveyor pursuant to sub-rule (3).

Chapter 5

Passenger Transport Operation

5.1 Public Transportation

Public transportation is a generic term used to describe the family of transit services available to urban and rural residents. Thus, it is not a single mode but a variety of traditional and innovative services, which should complement each other to provide system-wide mobility.

5.1.1 Transit Modes

The modes included within the realm of public transportation are:

- **Mass transit**, characterized by fixed routes, published schedules, designated networks, and specified stops. Mass-transit vehicles include buses, light rail (trolleys) or rapid transit that either share space in mixed traffic or operate on grade-separated rights of way.

- **Paratransit** is characterized by flexible and personalized service intended to replace conventional fixed-route, fixed-schedule mass-transit lines. Paratransit is available to the public on demand, by subscription, or on a shared-ride basis. Examples include taxi, car rental, dial-a-ride, and specialized services for elderly, medical, and other designated users.

- **Ridesharing** (as the name implies) is characterized by two or more persons traveling together by prearrangement, such as carpool, vanpool, or shared-ride taxi.

Buses are the most common transit mode. They operate on streets and have an extensive network of lines. In some cities they have been upgraded by provision of exclusive bus lanes and provision of bus preferential signals.

Light Rail Transit (LRT) represents the most common mode of semirapid transit. Its articulated electric vehicles operated in short trains on largely separated tracks provide more attractive and permanent services than buses at a much lower investment cost than metro systems require. LRT is presently being developed in many cities around the world that want to make transit services more efficient and largely independent of traffic congestion.

Metro systems have by far the highest performance - capacity, speed, reliability - of all transit modes. They require very high investment, but in the long run they are essential for efficient functioning and quality of life in large cities.

Transit modes are defined by their right-of-way (ROW) category, technology and types of operations. Three ROW categories are:

- ***ROW Category C :*** It represents public streets with general and mixed traffic. Street transit modes include mostly buses, but also trolleybuses and tramways/streetcars.
- ***ROW Category B***: It represents transit ways that are partially separated from other traffic. Typically they are street medians with rail tracks, which are longitudinally separated, but cross street intersections at grade. Bus lanes physically separated from other traffic also represent ROW category B. This ROW requires a separate strip of land and certain investment for construction.
- ***ROW Category A***: It is fully separated, physically protected ROW on which only transit vehicles operate. This category includes tunnels, aerial (elevated) structures or fully protected at-grade tracks or roadways. Thus, vertical position of the ROW is not as important as its separation from other traffic, because total independence of TUs allows many physical and operational features that are not possible on ROW categories B and C. Therefore, the modes with ROW category A are guided (rail, exceptionally rubber-tired) systems with trains, electric traction and signal control which offer very high capacity, speed, reliability and safety.

5.1.2 Transit Capacity and Level of Service

A basic attribute of any transit mode is its carrying capacity, defined as the number of vehicles or persons that pass a given point in a specified time (usually an hour). The numerical value of carrying capacity (usually referred simply as capacity), is dependent on two variables: (1) the number of vehicles that pass a point at a given time and (2) the number of passengers within each vehicle. For example, if for a given lane along a section of highway there are 60 buses that pass by in an hour (or one per minute), and each bus carries 50 seated passengers, then the carrying capacity of this highway lane is 60 buses/ln/hr or $50 \times 60 = 3000$ passengers/ln/hr.

Carrying capacity is influenced by (1) the "spacing" in seconds between each vehicle (called the headway) and (2) the "comfort factor" experienced by passengers (called the level of service). Thus, carrying capacity can be increased in two ways: (1) reduce the headway or (2) increase the number of passengers per vehicle. In the bus capacity example, the headway was 60 seconds and the level of service was that all passengers had a seat. Time spacing between buses could possibly be reduced, but there are limits to lowering headway values dictated by safe distance requirements between vehicles and/or the time spent at transit stops and terminals (called dwell time). Similarly, passenger loading could be increased by allowing standees, but this would decrease the comfort level for passengers. Were the bus equipped with computer tables and a refreshment area (thus offering a higher level of service), fewer passengers could be accommodated resulting in a lower carrying capacity but a higher level of service.

Accordingly, when reporting transit capacity, it is important to specify the units as either vehicles or passengers/hour and the corresponding level of service in terms of passengers/vehicle. Public transit is often compared with the automobile when issues of carrying capacity are involved, as it is commonly believed that transit capacity is superior to auto capacity. The capacity of a single lane of passenger vehicles is approximately 2,000 vehicles/hour which represents a headway of 1.8 seconds. Since most cars have at least five seats, the person capacity of a highway lane could be as great as $5 \times 2000 = 10000$ per/hour. Capacities of this magnitude never have been achieved, since most cars carry only one person with an average car occupancy of about 1.5. Why is this so? Have you ever driven in a car carrying five people? Not a pleasant experience, and a reason why car pooling is not very popular. Given the opportunity, most people choose to drive alone or with just one other person.

Travelers usually consider many more factors than simply the in-vehicle level of service, and they don't really consider how they can contribute to increasing "carrying capacity." In fact, if drivers were to optimize the carrying capacity of a highway, they would all drive at 35 miles per hour! Other major considerations in selecting the travel mode include: reliability, punctuality, cost, travel time, and safety. Transit systems that receive "high marks" for the out-of-vehicle level-of-service factors are typically the ones that use exclusive lanes or tracks with no interference from other vehicles or pedestrians and have adequate capacity at station stops and terminals. Thus, rapid transit services (whether bus or fixed guideway), are the superior mode but are more costly to build and maintain and require high volumes of demand to be feasible.

5.1.3 The Role and Future of Public Transportation

Public transportation is an important element of the total transportation services provided within large and small metropolitan areas. A major advantage of public transportation is that it can provide high-capacity, energy-efficient movement in densely traveled corridors. It also serves medium- and low-density areas by offering an option for auto owners who do not wish to drive and an essential service to those without access to automobiles, such as school children, senior citizens, single-auto families, and others who may be economically or physically disadvantaged.

For most of this century, public transportation was provided by the private sector. However, increases in auto ownership, shifts in living patterns to low-density suburbs, and the relocation of industry and commerce away from the central city, along with changes in lifestyle (which have been occurring since the end of World War II) have resulted in a steady decline in transit ridership. Since the early 1960s, most transit services have been provided by the public sector. Income from fares no longer represent the principal source of revenue, and over a 25- to 30-year period, the proportion of funds for transit provided by federal, state, and local governments has increased steadily. While it generally is believed that highways and motor transport will play a dominant role in providing personal transportation in the beginning decades of the twenty-first century, there are many unforeseen changes that could alter the balance between public and private transportation. Some could contribute to a decline in transit ridership while others might cause transit to become stronger, and for the remainder, there would be little or no effect. The potential changes that could influence transit usage are categorized here from the book Urban Mass Transportation Planning.

Factors Bad for Transit

- Growth of suburbs
- Industry and employment moving from the central city
- Increased suburb-to-suburb commuting
- Growth in private vehicle ownership
- Increased diversity in vehicle types such as SUVs, pickup trucks, and RVs

- High labor costs

Factors Good for Transit

- Emphasis by the government on air quality
- Higher prices of gasoline
- Depletion of energy resources
- Trends toward higher-density living
- Legislation to encourage “livable cities” and “smart growth”
- Increased number of people who cannot or choose not to drive

5.2 Bus Transit System

Buses represent the most widely used transit technology. Virtually every city in the world that has transit service operates buses. Large cities with rail transit also operate extensive bus networks, usually on lines with lower passenger volumes or as feeders to rail lines.

Bus service is easy to introduce or modify: basic service requires only purchase of vehicles, garage and maintenance facilities, and organization of service. Stops along the lines can be simple. Therefore, buses represent the most economical transit mode for lightly traveled lines. This flexibility of bus routes is an advantage for any necessary changes, but it is a disadvantage for major bus lines: they lack permanence, efficiency in carrying heavy passenger volumes, and image of permanent, physically fixed routes desired by passengers.

Compared to paratransit modes, bus transit is very labor-efficient: one driver operates a vehicle with capacity of 50-150 spaces. Compared to rail transit, buses are labor-intensive and have no economy of scale: on heavily traveled lines, for every additional 40-120 passengers, one bus and one driver must be added to the service.

5.2.1 Bus Vehicles

There is a range of bus vehicles by their size/capacity and body type. Some of the major types are discussed below:

Minibus is a 6-8 meters long vehicle, which has a capacity of 15-40 seats and standing spaces. It is used for lightly traveled lines, short shuttle lines, services in residential neighborhoods, etc.

Regular Bus is 10-12 m long, 2.50 m wide. It has 30-50 seats and 60-20 standing spaces (minimum number of seats corresponds to the maximum number of standing spaces)

Articulated Bus is a vehicle with the main body on two axles and an articulated section with the third axle. With their greater capacity, articulated buses are suited for heavily traveled lines. In few cities, with very heavy ridership, double-articulated buses with three body sections and four axles are used.

Double-decker Buses have two decks, the upper being for seated passengers only. Like articulated buses, double-deckers have a greater capacity than regular buses, but take less street space. They involve passengers climbing stairs, which is inconvenient. Riding on the upper deck, however, offers nice views for passengers. They are used extensively in the cities of the United Kingdom and many British Commonwealth countries, as well as in Berlin and a few other cities.

Low Floor Buses , perfected during the 1990s, have become standard in several industrialized countries. These buses have floors 35-40 cm above ground, so that entry from a curb is nearly flat, or a plate is provided for wheelchairs. Low-floor buses offer considerably greater comfort for passengers and speed up their boarding-alighting. Mechanical equipment on these buses is stored mostly on the roof, while the motor is in a compartment in the rear, where the floor is ramped up.

In selecting buses for a specific service, expected passenger volume is critical for vehicle design. Maneuverability and riding comfort are also considered. Thus, for lightly traveled bus lines in suburban areas with many narrow residential streets, or on hilly terrain, a minibus may be best suited because it is least expensive per vehicle-km, its small capacity is adequate and it can negotiate such alignments better than large buses. On the other hand, heavy passenger loads make regular or high-capacity buses more economical and superior in offering the required capacity. Average trip lengths influence the number and width of doors, as well as seating arrangement.

Relatively short trips and intensive exchange of passengers at stops requires two double channel doors on regular, 3 - 4 double channel doors on articulated buses, and single rows of seats on each side. For lines with moderate passenger loads and longer trips, 2+1 or 2+2 seating may be used. In the latter case, standing should be expected only in exceptional cases. In all cases access for passengers in wheelchairs is legally required to be provided by lifts, “kneeling bus” which can be lowered in the front, or by low-floor bus design.

5.2.2 Bus Travelways

The vast majority of buses operate on regular streets, ROW category C. Being in mixed traffic, their speed and reliability of service depend on traffic conditions. Their average speed is lower than the average speed of cars because they stop to pick up and drop off passengers. Buses are therefore not very competitive with car travel in the same corridor with respect to speed and reliability. Their advantages are much lower cost and convenience of not having to drive and park.

To make buses more efficient and attractive to passengers, bus preferential measures can be introduced. These include the following:

- ***Preferential signals***: buses in a separate approach lane at intersections get the green signal before other lanes, so that they can proceed through the intersection ahead of other traffic.

- ***Alternating stop locations*** at near- and far-side of intersections (before or after cross street) so that buses clearing one intersection on green signal use the green at the following intersection before they make the next stop. Also, spacing between bus stops should typically be about 250-400 m.

- ***Exclusive bus lanes***, which may be curb lanes or lanes in the median - ROW category B. This is the most significant improvement measure because it makes buses independent of traffic conditions on the same street.

- ***Buses on high-occupancy vehicle (HOV) lanes or roadways*** are used when bus lines with frequent service follow freeway alignment for a rather long distance. HOV facilities usually have traffic control that prevents congestion, but they do not provide the image of an exclusive, independent transit facility.

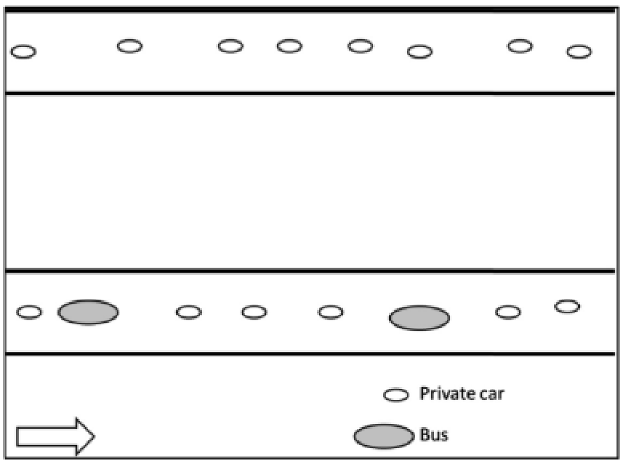
- ***Busway*** - special roadways reserved for buses only (ROW category B or A). Since busways require very high investment costs, they are used for some sections of lines. If ROW category A is required for a large section of line, it is usually better to introduce a rail system, so that the investment in high quality ROW is better used for electrically powered trains, rather than single bus vehicles

5.3 Number of Passengers Versus Number of Served Vehicles

Traffic engineers try by various techniques and measures to enable as many vehicles as possible to pass, during a specific period of time, through a traffic intersection or through a road section. In other words, traffic experts try to maximize the number ofvehicles that are served during a certain period of time. On the other hand, when it comes to public transit, we are trying to maximize the number of transported passengers. The number of passengers that can be served during the observed time interval represents transit line capacity. Similarly, the number of vehicles that can be served during the observed time interval represents the vehicle line capacity.

5.3.1 A Hypothetical Case Study

Let us assume that the freeway lane can serve 2200 vehicles per hour (the freeway lane capacity equals 2200 vehicles per hour). The upper part of the figure shows the situation when only private cars used freeway lane. The lower part of the figure refers to the situation when some buses also participate in freeway lane traffic. We assume that average number of passengers in bus and private car (average occupancies) are, respectively, equal to 50 and 1.4. We also assume that instead on any two private cars we can allow one bus to participate in freeway lane traffic.

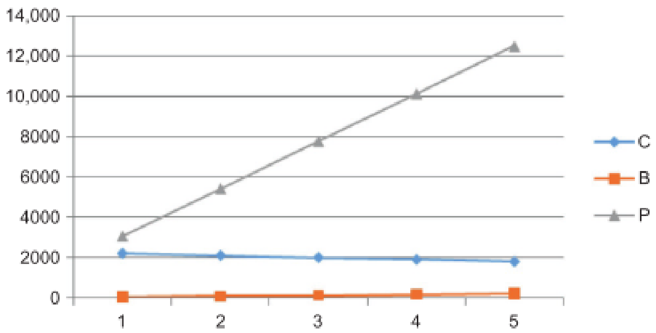


We denote by C, B, and P, respectively, the total number of cars, total number of buses, and total number of transported passengers. Ifwe allow, for example, 50 buses to enter the freeway lane, we need to reduce the total number of private cars for $50 \times 2 = 100$ private cars.In this case, the total number of transported passengers P by private cars and buses equals:

$P = C \times 1.4 + B \times 50 = 2100 \times 1.4 + 50 \times 50 = 5440$ Passengers

The table and the graph below show increase in the total number of transported passengers with the increase of the number of buses engaged.

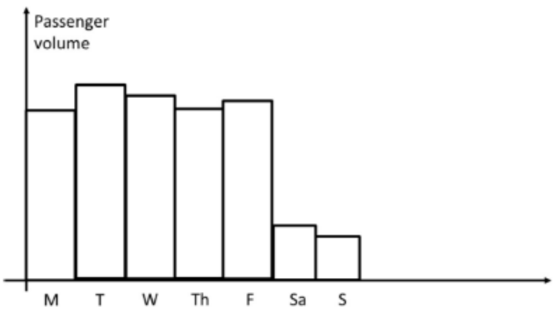
Number of Cars per Hour (*C*)	Number of Buses per Hour (*B*)	Number of Passengers per Hour (*P*)
2200	0	3080
2100	50	5440
2000	100	7800
1900	150	10,160
1800	200	12,520



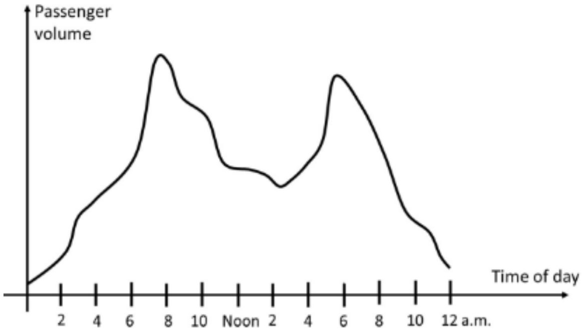
The graph above illustrates the effects that public transit can achieve. With a very small percentage share of public transport vehicles in the total number of vehicles, it is possible to increase dramatically the total number of passengers carried. To achieve such effects in reality, it is necessary constantly to increase the public transit attractiveness.

5.4 Passenger Flows in Public Transportation

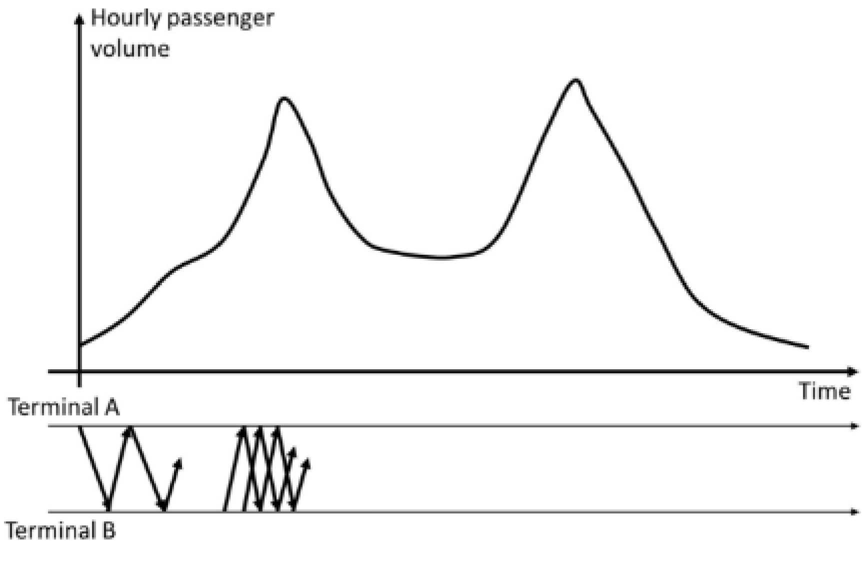
Passenger flows vary considerably in public transportation. During the working day, in most cities in the world, the daily numbers of passengers on most transit lines are relatively uniform. This is due to the fact that many passengers, each working day, use the same transportation mode and the same route when going to work. During a weekend the number of passengers in public transit is significantly smaller. The figure below shows daily variations in passenger volume.



The fig below shows hourly variations in passenger volume. The variations shown in figure are typical for many cities in the world. There are morning and evening peaks when people go, and when people come back from work.The differences in hourly passenger volumes could be very high by hours of a day.



Hourly variations in passenger volume have, as a consequence, different number of vehicle departures from the terminals during certain time intervals as shown in the figure below. In this way, hourly variations in passenger volume require the engagement of different number of vehicles during certain time periods. The number of engaged vehicles is much higher during the rush-hours. Outside the peak periods the transit operator has a surplus of vehicles and drivers. The transit operator, therefore, meets with a range of organizational problems that have to be solved ("empty" vehicle trips to garage and from garage, drivers working hours divided in two shifts, vehicle maintenance planning, etc.).

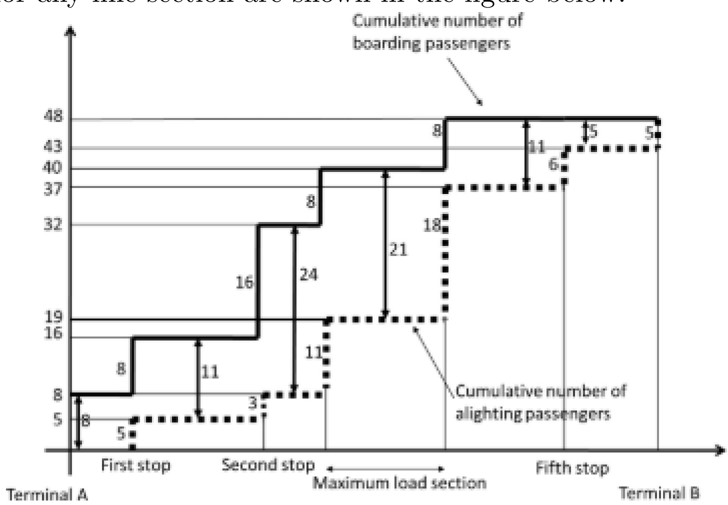


5.5 Passenger Flows Along a Transit Line

Headway in public transportation operations represents the time interval between vehicles past a specific point as shown in the fig below. Headways are expressed in minutes. It is essential to study passenger flow along the transit line, in order to determine the appropriate transit line headway.

Let us assume that we have data on the number of boarding passengers and number of alighting passengers on individual line stops and terminals (end stations on a transit line) as shown in the table below. The transit line has 5 stations. The terminals are denoted respectively by A and B.

The last column of the table shows the numbers of passengers in the vehicle, after departing from bus stops. Thus, for example, after leaving the station 2, there are 21 passengers in the vehicle that travels along line section between stop 3 and stop 4. The number of boarding passengers, number of alighting passengers and the number of passengers in the vehicle for any line section are shown in the figure below.

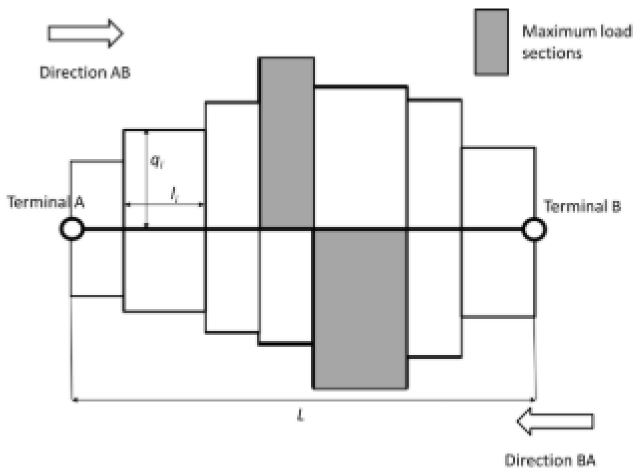


Bus Stop	Number of boarding Passengers	Number of Alighting Passengers	Number of Passengers in the Vehicle
Terminal A	8	0	8
1	8	5	11
2	16	3	24
3	8	11	21
4	8	18	11
5	0	6	5
Terminal B	0	5	0

The numbers of passengers on certain line sections respectively equal 8, 11, 24, 21, 11, and 5. The maximum passenger volume equals $\max 8, 11, 24, 21, 11, 5 = 24$ and corresponds to the section between stop 2 and stop 3. We call section between stop 2 and stop 3 the maximum load section (MLS).

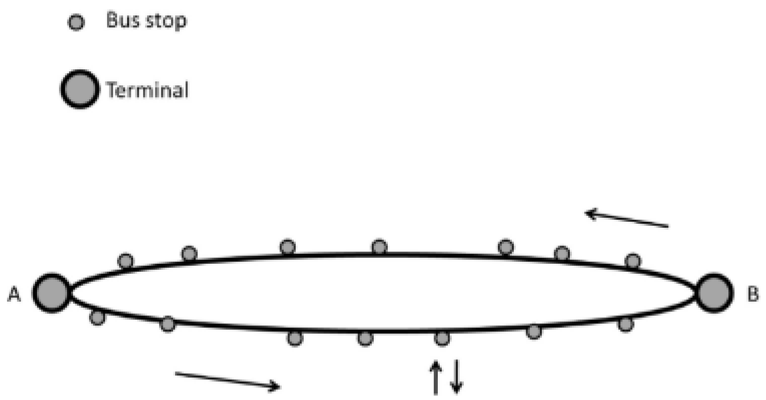
The data shown in the table and figure above are related to one vehicle trip. In real-life applications, 1 h is the basic time unit that is used when describing cumulative number of alighting and boarding passengers, as well as the maximum passenger volume on MLS. In other words, the maximum passenger volume is expressed in passengers per hour. The passenger volume profile is calculated for both directions of the transit line.

In the figure shown below, the line length is denoted by L. This length represents the one-way distance between the line terminals along the line alignment. The line length is measured in miles or kilometers. We denote by qi passenger volume on the i^th section. The length of the i^th transit section is denoted by li. The MLS in one direction is usually different from the MLS in another direction. The maximum passenger volumes in both directions should be taken into account when determining the number of vehicles to be engaged on the transit line.



5.6 Service Frequency and Headways

One of the most important problems encountered by public transit operators is how to match transportation supply and passenger demand on individual transit line. The matching problem is far more complex over the entire route network than on individual routes. Service frequencies and vehicle departure times on transit lines in the network reflect the manner in which transportation supply and passenger demand are matched. Service frequencies and vehicle departure times depend on passenger volume profile and on the number and type of vehicles in the fleet. The number of passengers that decide in the end to use public transit on a particular transit line depends, to the highest degree, on service frequency and vehicle departure times. For example, if service frequency is low or if vehicle departure times during the day are not convenient, a number of potential passengers will instead choose other modes of transportation.



Let us note the bus line shown in the figure above. Vehicles move from Terminal A to Terminal B. On the way to the Terminal B, vehicles stop at pre-defined bus-stops, where passengers enter and exit the vehicle. On arrival at the Terminal B, the driver rests for a while, and then the vehicle travels to Terminal A. On the way to Terminal A, car stops at bus-stops where passengers enter and exit the vehicle.

We denote by T turnaround time. This time is the time that elapses from the moment when vehicle leaves Terminal A to the moment when vehicle returns to Terminal A. Let us assume that we have on our disposal N vehicles that we can assign to the bus line. The service frequency represents the number of vehicles per time unit past a specific point in the same direction. The frequency equals:

$$f = \frac{N}{T} \tag{5.1}$$

The frequency is expressed in the number of vehicle per hour. Headway h in public transportation operations represents the time interval between vehicles past a specific point. Headways are expressed in minutes. Since the frequency represents the number of vehicles per time unit past a specific point in the same direction, we conclude that the frequency is the inverse of the headway, ie:

$$h = \frac{1}{f} \tag{5.2}$$

When calculating and rounding headways, it is desirable to obtain a so called clock headway. Clock headways have a feature that enables the generation of timetable that is repeated every hour, starting on the hour. Thus, for example, in the case when the headway is equal to 15 min, it is possible to have the vehicle departures from the terminal in 8:00, 8:15, 8:30, 8:45, 9:00, 9:15, 9:30, 9:45, 10:00, etc.

5.6.1 The Maximum Service Frequency

The maximum service frequency is defined by the maximum number of transit vehicles passing through a given point of a line/route i in one direction during a given period of time (usually 1 h) under prevailing operating conditions. This can be estimated as follows:

$$f_{max/i}(\tau) = \frac{\tau}{h_{min/i}} \tag{5.3}$$

Where,

τ is the given period of time

$h_{min/i}$ is the minimum headway, ie, the time interval between the successive transit vehicles passing through a given point of a line/route i in the same direction (min).

The minimum headway $h_{min/i}$ in Eq. 5.3 can be determined according to different criteria, but in many cases the prevailing factors are characteristics of the system such as technology and way of operations along the line and at the stations. These factors influence the minimum headway for a given line line/ route and for the stops/stations along it. Most frequently the stop/station headway is as much greater than that of the line(s)/route(s). In addition, the headway $h_{s/i}$ at the stop/station on a given line/route i, should be greater than the vehicle stop time $t_{s/i}$ at that stop/station. Consequently, the "ultimate" capacity of this stop/station will be:

$$C_{ss/i}(\tau) = \frac{\tau}{max[h_{s/i}, t_{s/i}]} \tag{5.4}$$

In addition, the following must be satisfied for all stops/stations on the line/route i:

$$f_{max/i}(\tau) \leq C_{ss/i} \tag{5.5}$$

In other words, the maximum number of transit vehicles passing through a given stops/stations cannot exceed the "ultimate" capacity of any of these stops/stations.

5.6.2 Passenger Waiting Time

Passengers' walk to a stop and passenger waiting time are the basic attributes of the public transit level of service. Walk to stop in the range of 400–800 m is considered acceptable for public transit users.

In order to estimate the average passenger waiting time, let us first consider the situation when the bus arrives at the bus stop regularly, according to the published timetable. We also assume that all passengers at the bus stop can enter the vehicle, and that the passengers appear at the bus stop in random moments of time. It has been shown that, in this case, the average waiting time per passenger at the station w is equal to the one half of the vehicle headway h, ie:

$$w = \frac{h}{2} \tag{5.6}$$

The average waiting time per passenger could be longer in the case of irregular bus arrivals. In the case of irregular bus arrivals, vehicle headway is not any more deterministic quantity. In this case, the vehicle headway is a random variable. If the planned headway equals, 10 min, in the case of irregular arrivals headway values could be, for example, 8, 9, 12, 15,.... minutes. It has been shown, that in the case of irregular vehicle arrivals at the stop, the average passenger waiting time equals:

$$E(W) = \frac{E(H)}{2} + \frac{var(H)}{2 \times E(H)} \tag{5.7}$$

Where, $E(H)$ is the expected value of the random variable H; and $var(H)$ is the variance of the random variable H (variance represents the square of standard deviation).

5.6.3 Headway Determination by "Square Root Formula"

Transit operator cost and passenger cost depend on chosen headway. Passenger cost, in the case when vehicles arrive regularly, has linear increase with headway. The greater the headway, the greater the passenger waiting time and passenger cost. On the other hand, greater headway means for transit operator smaller number of departures and lower costs.

Z is the total cost per hour;

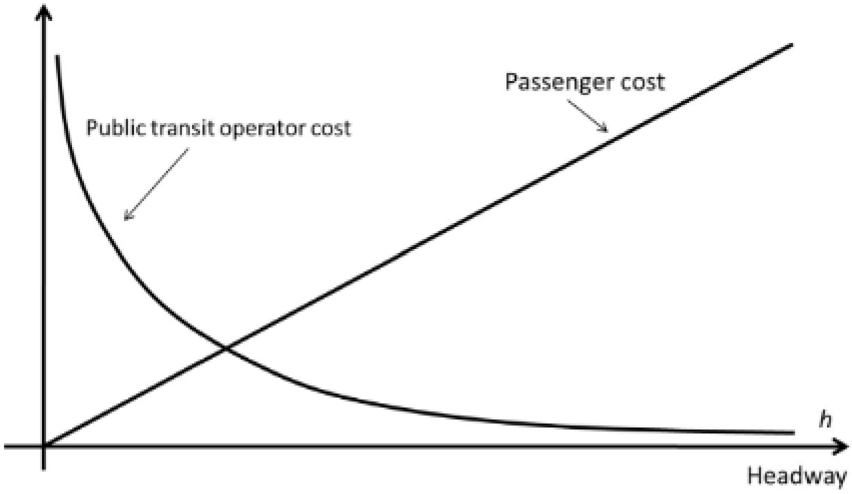
c is the transit operator cost per bus hour;

ν is the value of passenger waiting time per hour;

r is the total number of passengers on line per hour (ridership per hour);

N is the number of vehicles assigned to the bus line; and

h is the headway.



The transit operator cost per hour is equal to $N \times c$. We assume regular vehicle arrivals. In this case, the average waiting time per passenger at the station w is equal to the one half of the vehicle headway h, ie:

$$w = \frac{h}{2} \tag{5.8}$$

The waiting cost of all passengers is equal to the $\nu \times r \times \frac{h}{2}$. The total cost is:

$$Z = N \times c + \nu \times r \times \frac{h}{2} \tag{5.9}$$

Since, $N = \frac{T}{h}$, we can write,

$$Z = \frac{T}{h} \times c + \nu \times r \times \frac{h}{2} \tag{5.10}$$

The optimal headway is found by setting the derivative of Z with respect to h equal to zero:

$$\frac{dZ}{dh} = -c \times \frac{T}{h^2} + \frac{\nu \times r}{2} = 0 \tag{5.11}$$

Therefore the Optimum Headway,

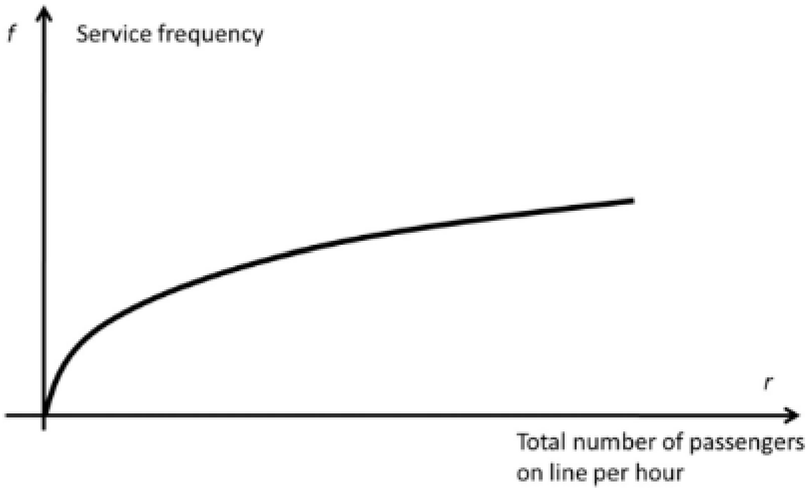
$$h = \sqrt{\frac{2 \times c \times T}{\nu \times r}} \tag{5.12}$$

Eq. 5.12 represents the "square root formula" for optimizing headway and service frequency. Minimal headway values in real-life are usually between 2 and 3 min. Maximal headways values are between 15 and 30 min. Outside of peak periods, on some transit lines, maximum headway values reach 60 min.

The service frequency f is the inverse of the headway, ie:

$$f = \frac{1}{h} = \sqrt{\frac{\nu \times r}{2 \times c \times T}} \tag{5.13}$$

Service frequency dependence of the total number of passengers on line per hour is shown in the figure below.



Numerical Example:
Let us assume the following line parameters: $c = 120\$$ per hour, $\nu = 10\$$ per passenger hour, $r = 1200$ passengers per hour, and $T = 1.5$ hr. Calculate the optimal headway.
Solution:

$$h = \sqrt{\frac{2 \times c \times T}{\nu \times r}} = \sqrt{\frac{2 \times 120 \times 1.5}{10 \times 1200}} = 0.173hr$$

$$\therefore h \approx 10min$$

5.6.4 Headway Determination by Maximum Load Method

When determining headways, transit operators try to provide enough space (especially during peak hours) to meet passenger demand. Majority of transit operators also define maximum headways on transit lines. For example, operator could define that maximum headway on a specific route, is equal to 30 min. Maximum headways guarantee a minimum service frequency offered to the passengers. The prescribed maximum headway is usually called *policy headway* and denoted by h_p.

The Maximum load method (that can have few variations) is based on counting passengers on the transit stop that is at the beginning of the MLS. Depending on the time of a day, the location of the MLS could change. For example, the transit stop, that is at the beginning of the MLS, could be Stop #5 between 10:00 am and 11:00 am, while Stop #7 could be at the beginning of the MLS between 5:00 pm and 6:00 pm Frequently, passenger counting is performed at the bus stop that has the highest daily passenger volume. The location of this transit stop is usually well known to the transit operator. The counting interval (whole day, between 7:00 am and 10:00 am, between 4:00 pm and 7:00 pm, etc.) is different for different transit operators and different cities.

We denote by P_{max} the average value of the maximum daily passenger volume. For example, transit operator monitored during seven days period the daily number of passengers that departed from the station # 6. The following 7 values were recorded: 1262, 1348, 1439, 1285, 1290, 1391, and 1287. The P_{max}, in this case is equal to:

$$P_{max} = \frac{1262 + 1348 + 1439 + 1285 + 1290 + 1391 + 1287}{7} = 1329$$

The service frequency, f, that should be offered in order to satisfy maximum passenger volume and desired vehicle occupancy is equal to:

$$f = \frac{P_{max}}{\alpha \times C_{car}} \tag{5.14}$$

Where, C_{car} is maximum number of passengers per car; and $0 \leq \alpha \leq 1$ is load factor.

The Load factor α is related to the concept of desired vehicle occupancy. The product $\alpha \times C_{car}$ defines the desired vehicle occupancy during the observed time period.

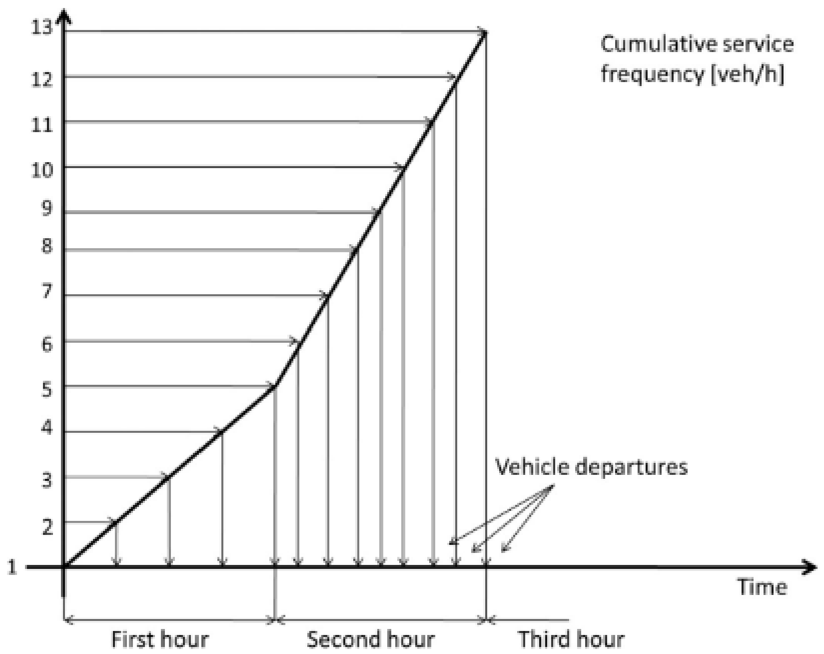
The corresponding headway is equal to:

$$h = \frac{1}{f} = \frac{\alpha \times C_{car}}{P_{max}} \tag{5.15}$$

5.7 Time Table

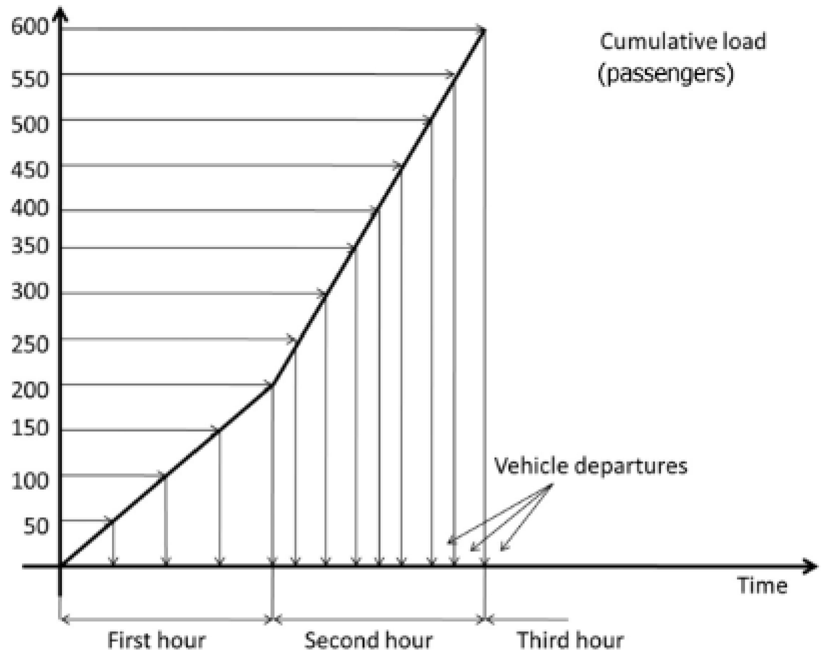
The transit line timetable is generated at one point (usually terminal). In the next step, by using information about average travel times between transit stops, timetable is generated for all transit stops. The transit operator's timetable contains information about vehicle departure times at all transit stops.

The basic input data for the timetable construction represent service frequency values in certain intervals of time during the day. Timetable can easily be generated in the following way. Let the abscissa represents time. Based on the known values of service frequency in certain intervals of time, we draw a cumulative frequency.



The cumulative frequency shown in the figure above is related to the case when service frequency within first hour is equal to 4, and service frequency within second hour is equal to 8. Let us first vehicle departure happen at the beginning of the first hour. We can go horizontally for every vehicle departure, until intersecting the cumulative curve. From the point of intersection we go vertically. The intersection of this vertical line and abscissa represents vehicle departure time. In this way, we generate timetable at specific point. In this way, we generate constant headways during specific time intervals.

Because of demand variability, constant headways frequently do not generate equal number of passengers at every vehicle departure. In order to achieve even-load at every vehicle departure, we draw cumulative loads on the transit stop that is at the beginning of the MLS.



Unequal number of passengers at specific vehicle departures can cause overcrowding in some vehicles, long boarding time at some transit stops, "bunching" of vehicles, and a decrease in the level of service offered to the passengers. On the other hand, equal load is related to unequal headways which could be inconvenient for passengers.

Trips between nodes in public transit networks may be made with or with no making transfers. Transfers generally cause inconvenience to passengers. Given that inadequately coordinated transfers can increase waiting times considerably, it is particularly important (when constructing timetables) to synchronize schedules cautiously in cases of larger headways. Unsuccessfully coordinated transfers can also reduce the number of passengers using public transit as a result of switching to competitor modes. When designing synchronized schedules, it is essential to try to minimize the total waiting times of all passengers at transfer nodes in a transit network.

5.8 Transit Line Capacity

The transit line is the basic element of public transit system. The transit route lengths in one direction are usually between 40 and 90 min, while stop spacing in urban areas are in the range from 120 to 400 m. It is desirable that transit route intersects few other transit routes. In this way, transfer points are generated that enable passengers to create various itineraries when making a trip. The waiting time at transfer points, which is less than 8 min, is considered as an acceptable waiting time. The transit line operating hours by weekdays are usually between 5:00 am and midnight. The spacing between transit lines, in the majority of cities, is in the range of 700–1000 m.

We denote by C_{car} the maximum number of passengers per car. Bus capacity represents the sum of number of seated passengers and legal standees. Public transit agencies and operators usually assume six passengers per square meter, as legal standees. In some countries, this figure could be higher.

The line/route (offered) "ultimate" capacity $C_{car}(\tau)$ is expressed by the maximum number of spaces, which can be transported in one direction during a given period of time τ (usually 1 hr) under prevailing operating conditions. In other words, the "ultimate" capacity of a public transportation lines represents the product of the maximum frequency of the service $f_{max}(\tau)$ and the maximum number of passengers in the transit vehicle.

$$C_{car}(\tau) = f_{max/i}(\tau) \times N_{car} \times C_{car}$$

The practical capacity C_{line} of a public transit line represents the product of the offered service frequency f and the maximum number of passengers in the transit vehicle.

$$C_{line} = f \times N_{car} \times C_{car} \tag{5.16}$$

Where, N_{car} is the number of cars/vehicles, and C_{car} s the maximum number of passengers per car.

The number of vehicles N_{car} per service frequency f can be variable. In the case of bus operations, $N_{car} = 1$, and the capacity of a public transportation line C_{line} equals:

$$C_{line} = f \times C_{car}(spaces/hr) \tag{5.17}$$

The turnaround time equals:

$$T = \frac{2 \times L}{u} \tag{5.18}$$

Where, L is the distance between Terminals A and B (for simplicity we assume that distance from A to B is equal to the distance from B to A); and u is the average vehicle speed.

Then,

$$C_{line} = \frac{N}{T} \times C_{car} \tag{5.19}$$

$$C_{line} = \frac{N}{\frac{2 \times L}{u}} \times C_{car} \tag{5.20}$$

$$\therefore C_{line} = \frac{N \times u \times C_{car}}{2 \times L} \tag{5.21}$$

As we can see, the line capacity C_{line} depends on the number of engaged vehicles N, the average speed u, vehicle capacity C_{car}, and length of the line L. By changing some of these quantities, it is possible to change line capacity.

Numerical Example:
The public transit line length equals 10 km in one direction. The average bus speed on a city heavy traffic equals 20 km/h. The total of 12 buses is assigned to the line. The capacity of every vehicle equals 50. Calculate the line capacity, frequency, and headway.
Solution:

The Line Capacity,

$C_{line} = \frac{N \times u \times C_{car}}{2 \times L}$

$C_{line} = \frac{12 \times 20 \times 50}{2 \times 10}$

$\therefore C_{line} = 600(spaces/hr)$

Then,

$T = \frac{2 \times L}{u} = \frac{2 \times 10}{20} = 1hr$

We have,

$f = \frac{N}{T} = \frac{12}{1}$

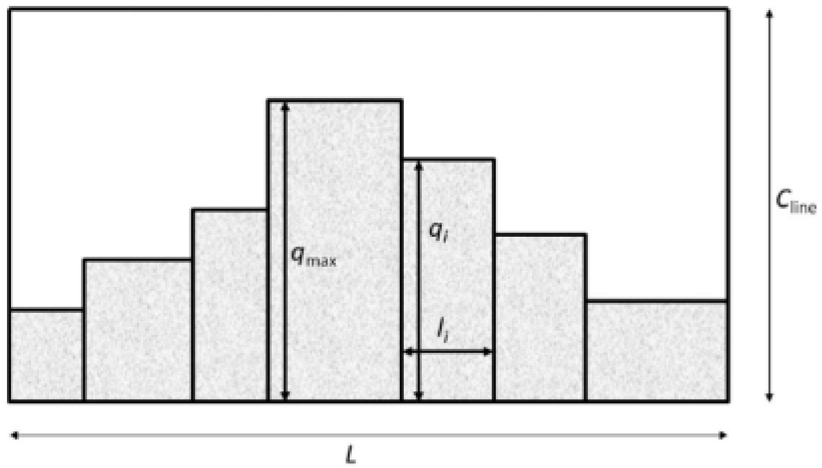
$\therefore f = 12vehicles/hr$

Also, $h = \frac{1}{f} = \frac{1}{12} = \frac{1}{\frac{12}{60}}$

$\therefore h = 5minutes$

5.8.1 Transit Line Capacity Utilization

The transit line capacity represents the number of spaces offered to passengers that pass a specific point in one direction during 1 h. Transit operator offers specific number of spaces to the passengers during specific period of time. It could happen that the offered capacity is insufficient in certain situations. On the other hand, is it possible that the offered capacity is underutilized. Therefore, there is a need to measure the utilization of the offered capacity. The figure below shows transit line capacity and passenger volume profile on the transit line between terminal A and terminal B.



The transportation work w_i, made by the transit operator when carrying q_i passengers along the section that has length equal to l_i, equals:

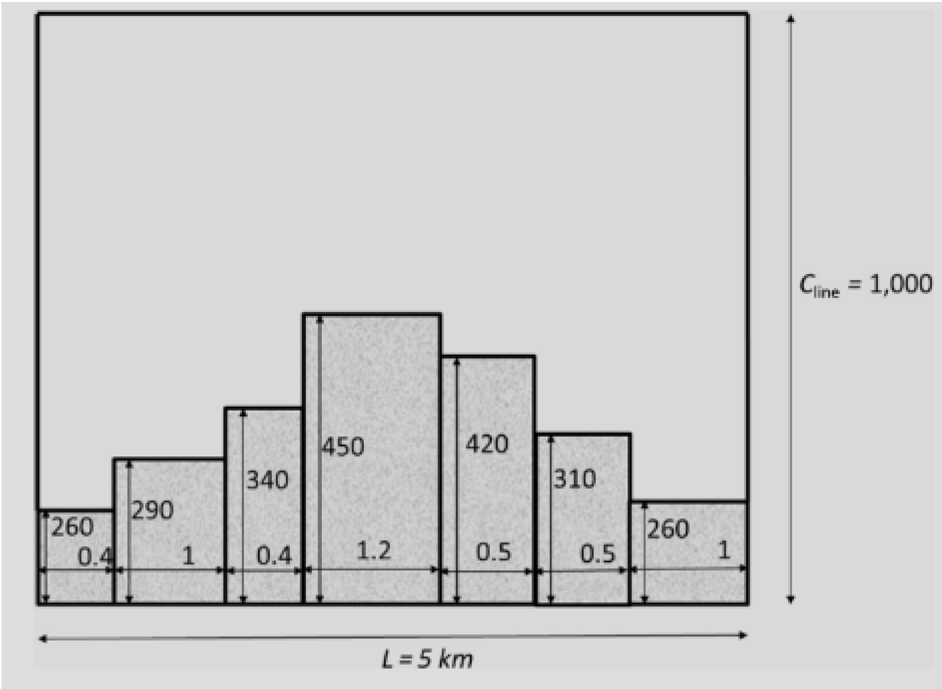
$$w_i = q_i \times l_i \tag{5.22}$$

Transit operator offers to the passengers the capacity that is equal to C_{line}. The transportation work that is possible to make is equal to the area of a rectangle with sides L and C_{line}. The realized transportation work is equal to the sum of the areas of shaded rectangles. The average transit line capacity utilization α is given by:

$$\alpha = \frac{\sum_{i=1}^{n} q_i \times l_i}{C_{line} \times L} \tag{5.23}$$

Where n is the number of line sections.

Numerical Example: The public transit line length equals 5 km in one direction. The average bus speed on a city heavy traffic equals 20 km/h. The total of 10 buses is assigned to the line. The capacity of every vehicle equals 50. The passenger volume profile of the line is given in the figure below.



Calculate the turnaround time, service frequency, headway, line capacity, and the average transit line capacity utilization α.

Solution:

Turnaround time,

$T = \frac{2 \times L}{u} = \frac{2 \times 5}{20} = 0.5hr$

$\therefore T = 30minutes$

Then, service frequency:

$\therefore f = \frac{N}{T} = \frac{10}{0.5} = 20vehicles/hr$

Also, headway:

$\therefore h = \frac{1}{f} = \frac{1}{20} = \frac{1}{\frac{20}{60}} = 3minutes$

Now, Line Capacity,

$\therefore C_{line} = \frac{N \times u \times C_{car}}{2 \times L} = \frac{10 \times 20 \times 50}{2 \times 5} = 1000spaces/hr$

Also, the average transit line capacity utilization:

$\alpha = \frac{\sum_{i=1}^{n} q_i \times l_i}{C_{line} \times L}$

$\alpha = \frac{260\times0.4+290\times1+340\times0.4+450\times1.2+420\times0.5+310\times0.5+260\times1}{1000\times5}$

$\therefore \alpha = 0.339$

5.9 Route Development

A transit system typically consists of many routes on each of which different number of transit units (for example, buses, trams, street, cars, and the like) ply. A route is a path that a transit unit follows during its journey from the origin terminus to the destination terminus and back. Along the route there can be many stops at which the transit units halt to let passengers alight and board. The subject of this section is to discuss what constitutes a good set of routes and how we can determine such a set.

5.9.1 Properties of a Good Route Set

A transit system's aim or for that matter any public transportation system's aim is to provide transport to a large portion of the population efficiently. In order to achieve this goal, the following properties of a route set are desirable.

Ridership: The percentage of the potential transit users who are served by the designed route set should be as high as possible. Given a route set and a statement of the origins and destinations of the potential transit users, we can determine the above quantity easily.

Riding Time: The time spent by any passenger on the transit unit should not be more than the time the passenger would have to spend going from his/her origin to his/her destination directly.

Transfer: Often in a transit network, we may have to change routes at some intermediate stops to go from the origin to the destination. This changing of routes is referred to as a transfer. The route set should be designed such that passengers are not required to make too many transfers while travelling from their origins to their destinations.

5.10 Stop Location and Stopping Policy

In this section, some issues related to then number of stops (and their locations) to be provided on a route and the stopping policy to be followed by the transit units plying on a route are discussed. The first subsection is on stopping policy and the next on stop locations.

5.10.1 Stopping Policy

Stopping policy relates to the policy followed by the transit system as regards to stopping on the route. There are different stopping policies of which any one may be used by a transit system. These are: (i) All stop, (ii) On-call stop, and (iii) Demand stop.

In the all stop case, the transit stops at all the stops at all the stops on the route irrespective of whether there is any demand for that stop. For a given transit unit, the demand for a stop exists if someone wants to board the transit unit from the stop or someone on the transit unit wants to alight at that stop. In the on-call stop case, the transit unit stops at a stop if and only if there is a demand for that stop. In the demand stop case, the transit unit stops anywhere along the route where it needs to stop to pick up or drop off passengers. In this case, although the route is fixed, the stop locations are not fixed or alternatively all points on the route are viable stop locations (although the transit unit does not stop at most of them).

In order to decide which is the best policy for a transit system (or a route in the transit system) a simple analysis, based on the expected number of places at which a transit unit a given route may have to stop, is done. The analysis procedure is shown below.

Let us assume:

1. The transit units are operating on a route with 'n' number of stops
2. A transit unit has to stop at a stop only is someone on the transit unit has to stop or someone is waiting at the stop
3. 'X' denotes the random variable that counts the number of people demanding to use the stop
4. The passenger demand for boarding/alighting at any stop has a mean rate of occurrence (λ) over any interval of time in that route
5. 'h' is the time headway at which transit unit operates on the route

From the definition of λ,Expected value of X or average number of passengers demanding to use a stop:

$$E(X) = \lambda h \tag{5.24}$$

The probability distribution of X is determined as follows:

1. The headway time is subdivided into n sub-intervals such that, $\Delta t = \frac{h}{n}$, the sub-interval are very small and the probability that passengers demand to stop more than one at that interval is negligible.

2. Probability that passengers demand to stop at every sub-interval is constant, say 'p'
3. Passenger demanding to stop at a sub-interval is independent of passengers demanding to stop at other sub-intervals

Thus, we can model the distribution of X approximately as a binomial random variable, where each sub-interval generates passenger demand to stop or not. Then,

$E(X) = \lambda h = np$
$\therefore p = \frac{\lambda h}{n}$

From approximate binomial distribution,

$$P(X = x) = \binom{n}{x} p^x \times (1-p)^{n-x}$$

Since the sub-intervals, $\Delta t \to 0$. Thus, $n \to \infty$,

$\lim_{n\to\infty} P(X = x) = \lim_{n\to\infty} \binom{n}{x}(\frac{\lambda h}{n})^x \times (1 - \frac{\lambda h}{n})^{n-x}$

$= \lim_{n\to\infty} \frac{n!}{(n-x)!x!} \times (\frac{\lambda h}{n})^x \times (1 - \frac{\lambda h}{n})^{n-x}$

$= \lim_{n\to\infty} \frac{n\times(n-1)...\times(n-x+1)\times(n-x)!}{(n-x)!x!} \times (\frac{\lambda h}{n})^x \times (1 - \frac{\lambda h}{n})^{n-x}$

$= \frac{(\lambda h)^x}{x!} \lim_{n\to\infty} \frac{n\times(n-1)...\times(n-x+1)}{n^x} \times (1 - \frac{\lambda h}{n})^n \times (1 - \frac{\lambda h}{n})^{-x}$

$= \frac{(\lambda h)^x}{x!} \times e^{-\lambda h} \lim_{n\to\infty} \frac{n}{n} \times \frac{(n-1)}{n} \times ... \times \frac{(n-x+1)}{n} \times (1 - \frac{\lambda h}{n})^{-x}$ $[\because \lim_{n\to\infty}(1 + \frac{x}{n})^n = e^x]$

$= \frac{(\lambda h)^x}{x!} \times e^{-\lambda h} \lim_{n\to\infty}(1 - \frac{1}{n}) \times (1 - \frac{2}{n}) \times ... \times (1 - \frac{(x+1)}{n}) \times (1 - \frac{\lambda h}{n})^{-x}$

$= \frac{(\lambda h)^x}{x!} \times e^{-\lambda h} \times 1$

$$Prob.\ (x\, passengers\, demand\, to\, use\, a\, stop) = P(x) = \frac{(\lambda h)^x e^{-\lambda h}}{x!} \quad (5.25)$$

Hence,

$$Prob.\ (transit\ unit\ stops\ at\ a\ STOP) = 1 - P(0) = 1 - e^{-\lambda h} \quad (5.26)$$

Therefore, on the entire route of n stops, the probability the transit unit stops s times, $\Pi(s)$, is given by:

$$\Pi(s) = \binom{n}{s} [1 - e^{-\lambda h}]^s [e^{-\lambda h}]^{n-s} \quad (5.27)$$

Hence, , $E(s)$, expected value of s or average number of times a transit unit stops is given by:

$$E(s) = n \times \Pi(s) = n[1 - e^{-\lambda h}] \tag{5.28}$$

Now, let v = the passenger demand per unit time for entire route

Then,
Total demand for boarding or alighting per unit time = $2v$
Total passenger demand per unit time for each stop = $\frac{2v}{n}$

Since λ is the rate at which passenger demand for boarding/ alighting arises at any stop,

$$\lambda = \frac{2v}{n}$$

Therefore,

$$E(s) = n \times \Pi(s) = n\left[1 - e^{\frac{-2vh}{n}}\right] \tag{5.29}$$

Case I: If demand is very large, that is, $v \to \infty$

$\lim_{v\to\infty} E(s) = n[1 - e^{-\infty}] = n\left[1 - \frac{1}{e^{\infty}}\right] = n\left[1 - \frac{1}{\infty}\right] = n[1 - 0]$

$\therefore \lim_{v\to\infty} E(s) = n$

This implies when demand is very large, then on average, a transit unit has to stop at all stops. Hence, in this case, it is better to use all stop policy.

Case II: If demand is very small compared to the total number of stops, that is, $v << n$

$E(s) = n\left[1 - e^{\frac{-2vh}{n}}\right]$

$= n\left[1 - \left\{1 + \left(\frac{-2vh}{n}\right) + \frac{(\frac{-2vh}{n})^2}{2} + \frac{(\frac{-2vh}{n})^3}{6} + ...\right\}\right]$ $[\because e^x = 1+x+\frac{x^2}{2}+\frac{x^3}{6}+...]$

Neglecting higher order terms of (vh/n), we get:

$E(s) = n\left(1 - 1 + \frac{2vh}{n}\right)$

$\therefore E(s) = 2vh$

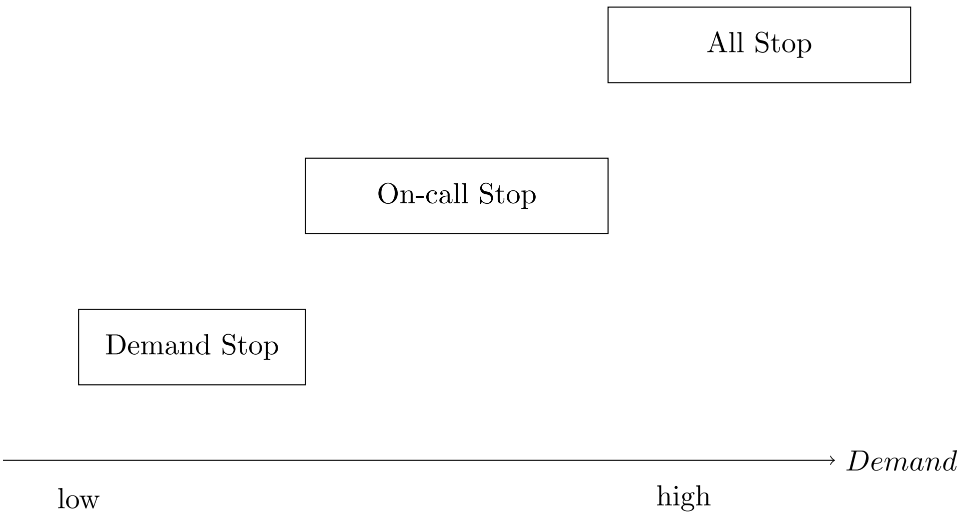
This implies that average number of places where transit unit has to stop is twice the number of passengers who use the transit unit for that period 'h', that is, no special benefit is obtained by a providing a fixed number of stops and stop locations. Hence, in case of low demand, it is better to follow a demand stop stopping policy with no predefined stop locations on the route.

Case III: For intermediate values of demand,

$$E(s) = n\left[1 - e^{\frac{-2vh}{n}}\right]$$
$$= n\left[1 - \left\{1 + \left(\frac{-2vh}{n}\right) + \frac{(\frac{-2vh}{n})^2}{2} + \frac{(\frac{-2vh}{n})^3}{6} + ...\right\}\right]$$
$$\therefore E(s) < n$$

Hence, on-call stop makes sense for this scenario.

The above analysis, in effect, brings out two aspects of stopping policy: (i) it relates demands to stopping policy and (ii) it shows that for the entire range of demand for a route, the expression $n\left[1 - e^{\frac{-2vh}{n}}\right]$ describe the average number of times a transit unit will have to stop.



5.10.2 Stop Location

Decisions on the number of stops to be on a route and where to put those stops effect the efficiency of operation on the route the route as well as the ridership on the route. The number of stops affect the efficiency of operation because every stop adds to the travel time on the route and hence to the round trip time. The number and location of stops also affect the ridership as too few and poorly located stops increase (i) the access time of passengers from their origins to a stop and (ii) the egress time of passengers from a stop to their destination; and too many stops increase the riding time of passengers as the operating speed falls with the increasing number of stops.

Hence in deciding the number of stops and stop locations, we have to strike a compromise between the operating efficiency and passenger convenience. Now, one of the simple models for deciding the number of stops is described.

A good way of determining the number of stops on a route (or the stop spacing) would be to determine the number of stops (or stop spacing) which minimizes the sum of the passenger travel time cost C_p and the operator cost C_o. The passenger travel time should include the access/egress time plus the in-vehicle riding time of the passengers. The operator cost is related to the number of transit units the operator has to operate in order to maintain the required frequency of service (or the required headway between the successive transit units). Mathematically, the total cost per unit time C_T is given by:

$$C_T = C_p + C_o = pT_uk_p + Nk_o \tag{5.30}$$

Where,

p is the number of passengers who use the route per unit time.

T_u is the average passenger travel time.

k_p is the cost of unit interval of time to a passenger.

N is the fleet of transit units running in order to maintain the time headway of h.

k_o is the cost of operating a single transit unit per unit time.

However, before attempting to minimize the total cost we need to determine T_u and N as these are functions of the number of stops n.

Determination of T_u

Let l be the average distance a passenger (or user) travels on a route and let L be the total one way length of the route. Let v be the operating speed of the transit unit on the route and let T_v be the total travel time of the transit unit over the length L.

Every time a transit unit stops at a stop location, it needs to decelerate from v to zero speed, remain stopped till all boarding and alighting is over, land then accelerate to v again. In this process, the transit unit loses time due to acceleration and deceleration and information can always be derived once the schedule of arrivals and departures at all stops is known.

5.10.3 Properties of a Good Schedule

A good schedule is one which distributes the available number of transit units in such a way that the following properties are satisfied.

Initial waiting time: The time between the arrival of a passenger at its origin stop for a particular route and the arrival of the next transit unit on that route is referred to as the initial waiting time. This should be as small as possible.

Transfer time: The time spent by passengers at an intermediate stop in order to transfer to another route (from the route on which the passenger came to the intermediate stop) is referred to as transfer time. This time should be small. Further, the transfer time of any passenger should not be greater than some allowable maximum value.

Policy headway: The time headway between the transit units of given route should not be greater that pre-specified maximum value, referred to as policy headway.

5.11 Optimal Paths in Transportation Networks

When traveling through the network we are faced with the problem of finding the paths that are "optimal." In other words, paths that we are looking for must possess optimal properties. The problems of finding optimal paths in transportation networks are known as shortest path problems. Depending of the context of the problem considered, the "shortest path" could be the shortest path, the fastest path, the most reliable path, the path with the greatest capacity, etc. network links are characterized by length. "Link length" could represent distance, travel time, travel cost, link reliability, etc. Link lengths are mostly treated as deterministic quantities.

Link lengths are treated in some problems as random variables. Most frequently, these link lengths represent travel times. There are random variations in travel times caused by weather conditions, randomness in traffic flows, traffic accidents, and other factors. In these cases, the shortest paths in probabilistic network should be determined. In some cases, when searching for the optimal path, we simultaneously try to optimize two or more objectives. For example, when searching for the best path, we could try simultaneously to take care of travel time as well as travel costs. In such cases, we are dealing with multicriteria shortest path problems.

5.11.1 Dijkstra's Shortest Path Algorithm

Dijkstra's algorithm is an algorithm we can use to find shortest distances or minimum costs depending on what is represented in a graph. You're basically working backwards from the end to the beginning, finding the shortest leg each time. The steps to this algorithm are as follows:

1. Start at the ending vertex by marking it with a distance of 0, because it's 0 units from the end. Call this vertex your current vertex, and put a circle around it indicating as such.

2. Identify all of the vertices that are connected to the current vertex with an edge. Calculate their distance to the end by adding the weight of the edge to the mark on the current vertex. Mark each of the vertices with their corresponding distance, but only change a vertex's mark if it's less than a previous mark. Each time you mark the starting vertex with a mark, keep track of the path that resulted in that mark.

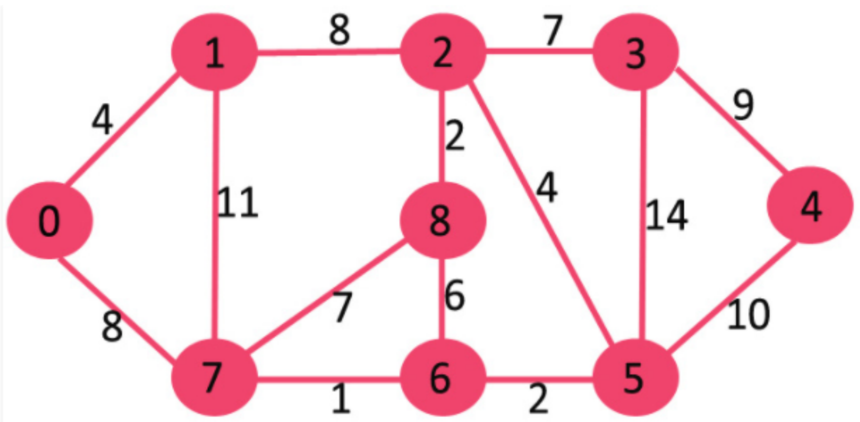
3. Label the current vertex as visited by putting an X over it. Once a vertex is visited, we won't look at it again.

4. Of the vertices you just marked, find the one with the smallest mark, and make it your current vertex. Now, you can start again from step 2.

5. Once you've labeled the beginning vertex as visited - stop. The distance of the shortest path is the mark of the starting vertex, and the shortest path is the path that resulted in that mark.

Numerical Example:

From the figure below, find the shortest path from node "0" to all other nodes, and also find the shortest route for node "4".



Solution:

Assign cost value of ∞ for all nodes which are not directly connected to node "0", and compute the distance for the nodes "1" and "7".

from/To	1	2	3	4	5	6	7	8
0	4*	∞	∞	∞	∞	∞	8	∞

Since $4 < 8$, we will assign node "7" as our next vertex, then:

from/To	1	2	3	4	5	6	7	8
0	4*	∞	∞	∞	∞	∞	8	∞
1	-	12	∞	∞	∞	∞	8* ($\because 8 < 15$)	∞

In the above step, we assigned the value 8 for node "7", because $8 < 15$(computed value from node "1"). Also, in this step, since $8 < 12$, we will assign node "7" as our next vertex, then:

from/To	1	2	3	4	5	6	7	8
0	4*	∞	∞	∞	∞	∞	8	∞
1	-	12	∞	∞	∞	∞	8*	∞
7	-	12	∞	∞	∞	9*	-	15

In the above step, node "7" is not directly connected to node "2" and the node is not solved yet, so the value remains as it is. Since, $9 < 12 < 15$, we assign node "6" as our next vertex, then:

from/To	1	2	3	4	5	6	7	8
0	4*	∞	∞	∞	∞	∞	8	∞
1	-	12	∞	∞	∞	∞	8*	∞
7	-	12	∞	∞	∞	9*	-	15
6	-	12	∞	∞	11*	-	-	$15_{(6,7)}$

In the above step, the computed value for node "8" from node "6" was 15, since the value didn't change we set the value as it is, that is, $15_{(6,7)}$ (This means that node 8 can be traversed using 6 an 7 in same distance). Also, $11 < 12 < 15$, so we assign node "5" as our next vertex, then:

from/To	1	2	3	4	5	6	7	8
0	4*	∞	∞	∞	∞	∞	8	∞
1	-	12	∞	∞	∞	∞	8*	∞
7	-	12	∞	∞	∞	9*	-	15
6	-	12	∞	∞	11*	-	-	$15_{(6,7)}$
5	-	12* ($\because 12 < 15$)	25	21	-	-	-	$15_{(6,7)}$

In the above step, the computed value for node "2" from node "5" was 15, and since $12 < 15$, we retain the earlier value, that is, 12. Also, $12 < 15 < 21 < 25$, we assign node "2" as our next vertex, then:

from/To	1	2	3	4	5	6	7	8
0	4*	∞	∞	∞	∞	∞	8	∞
1	-	12	∞	∞	∞	∞	8*	∞
7	-	12	∞	∞	∞	9*	-	15
6	-	12	∞	∞	11*	-	-	$15_{(6,7)}$
5	-	12*	25	21	-	-	-	$15_{(6,7)}$
2	-	-	19 ($\because 19 < 25$)	21	-	-	-	14* ($\because 14 < 15$)

in the above step, since $14 < 19 < 21$, we assign node "8" as our next vertex, then:

from/To	1	2	3	4	5	6	7	8
0	4*	∞	∞	∞	∞	∞	8	∞
1	-	12	∞	∞	∞	∞	8*	∞
7	-	12	∞	∞	∞	9*	-	15
6	-	12	∞	∞	11*	-	-	$15_{(6,7)}$
5	-	12*	25	21	-	-	-	$15_{(6,7)}$
2	-	-	19	21	-	-	-	14*
8	-	-	19*	21	-	-	-	-

In the above step, node "8" had no direct connection with either of the unsolved nodes, "3" and "4", so we resolve to assign node "3" as our next vertex as $19 < 21$, then:

from/To	1	2	3	4	5	6	7	8
0	4*	∞	∞	∞	∞	∞	8	∞
1	-	12	∞	∞	∞	∞	8*	∞
7	-	12	∞	∞	∞	9*	-	15
6	-	12	∞	∞	11*	-	-	$15_{(6,7)}$
5	-	12*	25	21	-	-	-	$15_{(6,7)}$
2	-	-	19	21	-	-	-	14*
8	-	-	19*	21	-	-	-	-
3	-	-	-	21* ($\because 21 < 28$)	-	-	-	-

Hence, the shortest distance from node "0" to all other nodes are given below:

node	shortest distance from node "0"/source
1	4
2	12
3	19
4	21
5	11
6	9
7	8
8	14

Hence the shortest route to reach node "4" from node "0" is 0-7-6-5-4.

5.12 Review Problems

1. Define the term public transportation. Briefly describe the different types of transit modes in public transportation.

2. List out the factors which are good and bad for transit.

3. Derive the expression for headway determination by "Square Root Formula".

4. Derive the expression for Transit Line Capacity, and use the expression to determine the average transit line utilization(α).

5. Let us assume that transit line parameters are respectively equal: $c =$ 120$ per hour, $\nu = 12$$ per passenger hour, $r = 1400$ passengers per hour, and $T = 120$ min $= 2$hr. Calculate the optimal headway by using the "square root formula."

6. Derive an expression for the average number of times a transit unit stops, $E(s)$, and use the relation to illustrate the relationship between stopping policy and demand.

7. Describe the properties of a good route set.

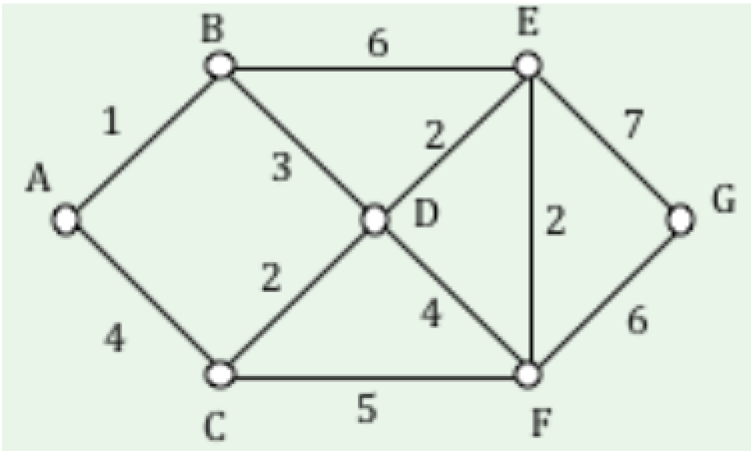
8. Discuss the properties of a good schedule.

9. The transit operator monitored during 10 days period the daily number of passengers that departed from the station #5. The following 10 values were recorded: 1160, 1245, 1440, 1280, 1180, 1380, 1220, 1358, 1178 and 1382. The maximum number of passengers per car is equal to 100. The transit operator would like to achieve the load factor that is equal to 0.75. Calculate the average value of the maximum daily passenger volumes. Determine the service frequency f that should be

offered in order to satisfy maximum passenger volume and the desired vehicle occupancy.

10. The public transit line length equals 12 km in one direction. The average bus speed on a city heavy traffic equals 25 km/h. The total of 14 buses is assigned to the line. The capacity of every vehicle equals 60. Calculate the line capacity, frequency, and headway.

11. Calculate the shortest path from node "A" to all other nodes, and also determine the shortest route from node "A" to node "G".



12. The number of boarding passengers, number of alighting passengers and the number of passengers in the vehicle are given in the following table:

Bus stop	# of boarding passengers	# of alighting passengers
Terminal A	9	0
1	10	6
2	14	4
3	9	11
4	7	17
5	0	7
Terminal B	0	4

Graphically illustrate the number of boarding passengers, number of alighting passengers and the number of passengers in the vehicle for all line sections. Also, graphically show the passenger load profile and identify the Maximum Load Section (MLS).

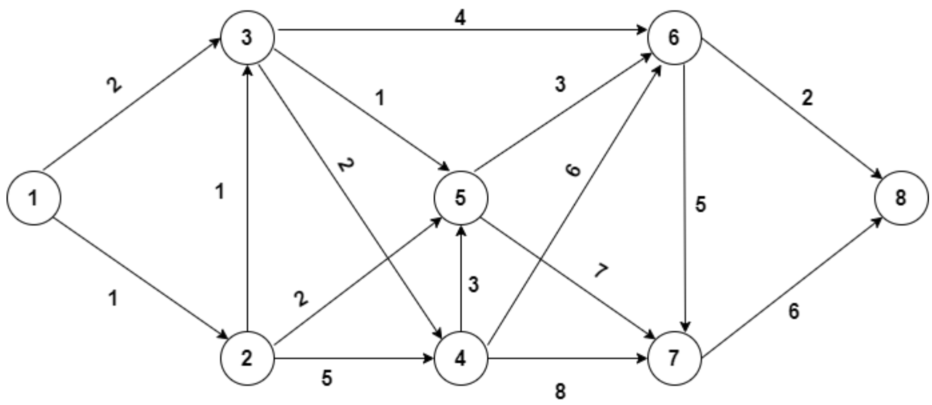
13. A public transit operator wants to provide service frequency between 6:00 and 7:00 that is equal to 4. Between 7:00 am and 8:00 am the

operator wants to offer service frequency equals to 6. The operator provides service frequency that is equal to 8 in the time interval between 8:00 am and 9:00 am. Generate vehicles timetable that will provide constant vehicle headways within time interval from 6:00 am to 9:00 am. Show the solution graphically.

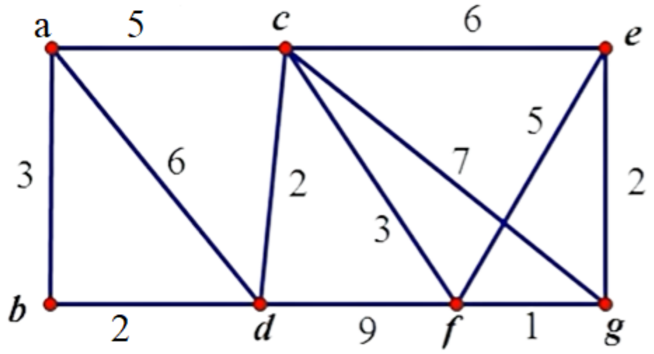
14. The network in the figure below gives the distances in miles between pairs of cities 1, 2, ..., 8. Find the shortest distance from city 1 to all other cities along with the shortest route between the following cities:

    (a) cities 1 and 8

    (b) cities 1 and 6



15. A bus transit service provider facilitates its service in a city. The city has seven terminals available for bus transit currently. The network of terminals is represented by a weighted graph given below, where the weighted values represent the distance between two terminals. Using Dijkstra's shortest path algorithm, determine the shortest distance from source terminal 'a' to all other terminals. Also, determine the shortest route/routes from source terminal 'a' to terminal



'g'.

Chapter 6

Freight Transport Operation

Some important terminologies:

1. ***Transport*** is the actual movement of goods from one location to another using a means or a vehicle of transport (e.g. trains, trucks, boats) and a transport infrastructure (e.g. roads, railways, canals).

2. ***Distribution*** often denotes all activities relevant to physical movement of goods, including transportation, but also transhipment and warehousing.

3. ***Logistics*** is generally used as an overarching term that includes all activities related to the movement and coordination of goods from their source of origin to the final point of delivery, and includes production and distribution. Here, movement does not just correspond to physical movement of goods but also the flow of information.

6.1 Main Actors of Freight Transport and Distribution

While freight transport and distribution include many elements, including various organizations and people, the three fundamental actors who actively take part in the domain are described here.

6.1.1 Shippers

The demand for freight transportation is generated by shippers. Each shipper will have their own logistics strategy, which includes whether to operate their own fleet or to use an external party to which it will outsource their logistics and distribution activities, as well as choosing the mode(s) of transport. The process through which the shippers will define their logistics activities generally takes a three-level decision structure:

1. The long-term decisions in the first level involve defining strategies in line with their customer network and production activities.
2. The second-level medium-term decisions include levels of inventories at production, warehousing, and distribution facilities, frequency and amount of shipping and flexibility of service.
3. At the short-term level, shippers decide on the attributes of the services required for its shipments, such as maximum rates, transport time, reliability and safety.

In making these decisions, they will consider the availability and the characteristics of the services offered on the market by carriers and intermediaries, such as freight brokers and third-party logistics providers.

6.1.2 Carriers

Carriers are people, businesses or organisations that operate and offer transportation services for shippers. They may either provide a customised service, where a vehicle or a fleet will be dedicated exclusively to a particular customer, or operate on the basis of consolidation, where each vehicle contains several pieces of freight for different customers with possibly different origins and destinations. In the latter case, carriers generally operate their services according to a published timetable, which prescribes routes, schedules and rates they offer.

6.1.3 Intermediaries

In some cases, the shipper operates their own fleets of vehicles and does not require an external carrier to ship goods on their behalf. In this case, the management of the relevant transportation and distribution is done inhouse. If a shipper does not own a fleet, then it may choose to work directly with one or several carriers.

Shippers may alternatively use a freight forwarder, an intermediary person or organization that acts as a third party and manages the shipments on behalf of the shipper by contracting one or several carriers. They also help to identify a suitable mode or a combination of modes for the shipper. Freight forwarders work closely with shippers and carriers, as well as other entities in the transportation network, such as ports or terminals, particularly if they additionally undertake ancillary services such as customs clearance.

6.2 Modes of Transportation

There exist different means of transporting freight over the network, each of which is referred to as a mode of transport. Transportation modes can be differentiated with respect to the type and specification of the vehicle used, the underlying technology, the relevant infrastructure and the nature of the associated operations. The three main modes of transport are air (e.g., cargo planes), land (including road, rail and off-road) and water (e.g., ships in oceans, barges in rivers). Other modes of transport, such as pipelines (e.g., to transport gas) and cable transport (e.g., elevators and cable-cars), also exist.

The term mode can also be used to denote different types of vehicles within a given domain of transport. For example, trucks, vans and bicycles can be seen as three separate modes operating within road transportation due to their distinct features, such as different capacities, capabilities and restrictions. In this example, trucks have the largest capacity but often have restrictions in traveling in urban areas, whereas bicycles are much smaller in capacity but do not suffer from the same type of restrictions as trucks or vans. Brief descriptions of the three main modes of transportation are explained here.

6.2.1 Road

Road has been, and continues to be, the most widely used mode of freight transport, both nationally and globally. One of the main reasons behind its popularity is the ability of road transport to offer a very quick service and often be available on demand.

A wide variety of vehicles are used for road freight transportation, which can be differentiated on the basis of size, capacity, weight and the type of energy used. Vehicle classification on road transportation is generally based on the Gross Vehicle Weight Rating (GVWR), which refers to the maximum allowable total weight of a vehicle including its empty mass, fuel and any load carried. The empty mass of the vehicle, but with fuel and fluids such as engine oil, is named as the curb weight. Vehicle classifications vary from one country to another. In the United States, eight classes exist, with vehicles in the lightest class having a GVWR up to around 3 tonnes, and those in the heaviest class with a GVWR higher than 15 tonnes. In the United Kingdom, more classes exist, with those of at most 3.5 tonnes gross weight described as light goods vehicles (LGVs) and those between 3.5 and 44 tonnes gross weight named as lorries or heavy goods vehicles (HGVs).

Most vehicles used in road freight transportation run on gasoline or diesel engines. Vehicles using alternative sources of fuel or energy have also been developed, such as those running on batteries, biofuels (such as bioalcohol or ethanol), biodiesel, compressed natural gas, hydrogen, and liquefied petroleum gas (LPG), for use in freight distribution. Within urban areas, human-powered vehicles, such as bicycles and tricycles, can also be used for goods deliveries. To overcome the sole dependency on human power, some of the freight bicycles have power assist motors to aid the cyclist.

The road network is composed of motorways (or highways), urban roads, rural roads, lanes or graded roads and includes bridges and tunnels. Traffic on the road network is controlled by means of traffic signals, signs or markings on the pavement. Various legal requirements are imposed on freight vehicles traveling on the road network, which include limitations on vehicle weight, dimensions, mandatory equipment, licenses and insurances. As for truck drivers, there also exist regulations on driving and working hours, which restrict the duration of driving time and require break and rest periods in long-haul journeys. These regulations aim at reducing driver fatigue, which is known to have adverse affects on road and driver safety. The regulations usually differentiate between on-duty time, which is the time spent working, including driving, waiting, loading and unloading and doing paperwork, and off-duty time, where the driver has no obligation to work. In the United States, for example, these regulations are known as Hours of Service, which limit the maximum consecutive driving time between two rest periods to 11 hours, at which point the driver must be off-duty for at least 10 consecutive hours. Furthermore, a truck driver cannot drive if 8 hours or more have elapsed since the end of the last off-duty period of at least 30 minutes. Similar regulations prevail in other countries, albeit with differences.

6.2.2 Rail

Rail freight transportation is known for its ability to offer cost-effective long-haul transportation services, primarily, but not exclusively, for bulk cargo. There are two major components of a rail system, namely the rail network infrastructure and freight trains.

The rail network is a large and complex structure composed of nodes and tracks (or track segments) as links between the nodes. The former include yards or terminals where classification or marshalling operations are performed, stations where cargo is picked up from or delivered to and junctions that are signal-controlled points in the rail network to allow trains to switch from one route to another.

Freight trains are composed of one or more locomotives, and several rail wagons (or cars). Locomotives move the train along the tracks by either pulling it from the front or push from the rear and range from the earlier types powered by steam to contemporary ones using electricity, magnetic force or diesel engine. Rail wagons carry the freight and come in a variety of forms, including specialized wagons for carrying particular types of cargo (e.g., autoracks for carrying automobiles or refrigerator cars for temperature-sensitive goods). A train is characterized by its route, origin, destination, intermediate stops, the physical path it travels on and the schedule information that includes departure and arrival times at each station where it stops. Each wagon also has an itinerary that specifies an origin and a destination station and need not correspond to the origin and destination of the train on which it is carried. Wagons may travel on several trains during their journey, usually in groups called blocks. Each block is assigned an origin and a destination, although individual cars in a single block may have different origins and destinations. A block is treated as a single unit for handling purposes. Once formed at its origin yard, a block will not be classified again until it arrives at its destination yard.

Classification or marshalling refers to a set of operations carried out at yards or terminals, where incoming trains are disassembled by decoupling the rail cars and new trains are formed using individual cars or blocks. Bekta¸s et al. (2009) provide a detailed description of the operations at a classification yard, according to which a train arriving at a yard first enters a receiving area, where the engines are taken away for inspection and maintenance, blocks are separated and cars are inspected. The classification operation begins from this point on and can be performed in two ways, depending on the type of the rail yard. In flat yards, a switching engine is used to push a group of cars out of the receiving tracks onto one of the classification tracks. In hump yards, classification is performed by using an artificially built hill, called the hump, where an engine pushes a group of cars out of the receiving area and up the ramp until it reaches the top of the hill. Due to the pull of the gravitational force, the cars roll down the incline on the other side of the hump, usually one car at a time, and are directed onto one of the classification tracks. Following this operation, each classification track becomes occupied by a group of cars that form the block. Each block then waits until the departure time of its outbound train. When the train is due to leave, they are pulled out of the classification tracks onto the departure tracks and are attached to the train. Following one last inspection of the whole train, the train and blocks leave the yard. The figure below shows the general structure of a classification hump yard.

A schematic representation of a rail classification (hump) yard showing the receiving area (RA), classification tracks (CT) and departure tracks (DT).

Classification is not the only operation that is performed at a rail yard. Other types of operations include inspection, crew change, refueling the trains and dropping and picking up blocks of cars. Among these, however, classification is known to be the major time-consuming operation.

6.2.3 Air

Air transportation is mainly used for goods that are time sensitive or have high value-to-weight ratios, including perishables due to the need for speed, fashion goods, emergency supplies and spare parts. Speed is a major advantage of air transportation but comes with the significant drawback of having a very high cost. This is one of the reasons as to why air transport has always had a marginal share in the amount of freight shipped worldwide, in comparison to other modes of transport.

Air freight transportation uses a system of air hubs, where outgoing cargo to be shipped is received and consolidated for shipping to a different hub in the network or incoming cargo from another hub is separated and shipped to its final destination. The network tends to have a hub-and-spoke structure to make the best use of economies of scale and help to achieve the lowest possible cost of transport. Specialized pallets or containers named unit load devices are used to stow cargo in aircraft, in a safe and efficient way. Air cargo can be carried in passenger planes along with passenger baggage. Alternatively, it can be transported in cargo aircraft designed specifically for freight, or by helicopters for access to areas that are difficult to reach.

6.2.4 Sea

Sea or maritime transport is the dominant mode with which a large majority of international trade is moved. It is suitable for transporting bulk goods, large packages or containers and similar types of cargo that are not time sensitive.

Vessels are the main vehicles of maritime transport, ranging from small boats to larger ships. Some of the common ship types include break-bulk vessels that are used to transport any type of loose cargo excluding liquid or loose bulk commodities (e.g. boxes, barrels, casks and bagged cargo), Roll-on/Roll-off (RoRo) vessels that carry wheeled cargo (e.g. cars, trucks, trailers), oil tankers that carry crude oil in liquid form, dry bulk vessels that carry bulk goods (e.g. coal, ore, grain) and reefer ships that carry perishable goods that require refrigeration (e.g. fish, meat, fresh produce).

The maritime transportation network consists of ports, which are hubs where ships depart from, stop at, or return to, used for loading, unloading and handling cargo. Ports come in different types based on their location and the nature of the cargo they handle. Seaports are the most common. Some ports only work with containerized cargo and include specialized equipment such as a gantry crane used to transfer containers between a vessel and other vehicles of land-based transport, such as trucks and rail wagons, and a stacker that is used to lift and swap containers, as well as perform loading and unloading operations on trucks and wagons.

In maritime transportation, there are larger variations in travel as compared with other modes of transport, caused by, for example, weather conditions or congestion at ports. In addition, travel and loading/unloading times are longer, particularly in the case of intercontinental trade.

There are three main modes of operation in maritime transportation:

1. ***Liner shipping*** is where vessels operate on regular routes with fixed schedules. The routes and schedules are published in a timetable, similar to bus services for passenger transportation. The routes are fixed, each of which is called a service route. There may be several vessels operating on a given service route, and this depends on the desired frequency over a given period of time (e.g., weekly).

2. ***Industrial shipping*** is where the cargo owner operates their own vessel fleet.

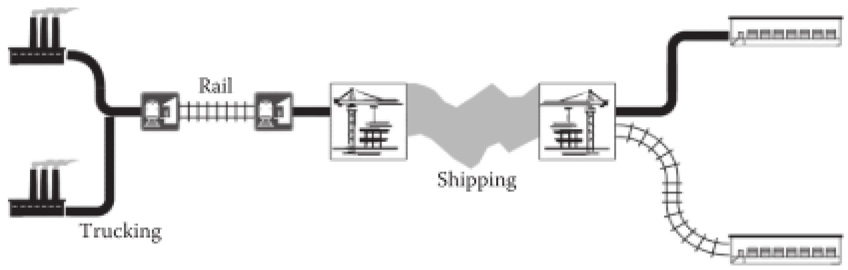
3. ***Tramp shipping*** is where a carrier operates a fleet of vessels to transport mandatory cargo between specified ports and within a certain time frame for contracted shippers but can also optionally carry spot cargo for other shippers, in order to increase their revenues. Tramp operators will need to decide on whether to accept or reject any spot cargos, which will depend on capacity and timing considerations in the light of the already scheduled routes for mandatory cargo.

6.2.5 Intermodal Transportation

A transportation network is said to be of unimodal type if it only operates a single mode of transport or of multimodal type if it operates at least two different modes of transport. Uni-modal transport cannot take advantage of the use of several modes of transportation, and it is either impractical or impossible to assume a single mode of transport for performing door-to-door delivery. If freight originates from an inland warehouse, for example, and is destined to another inland warehouse on another continent, then part of the journey will have to involve air or water transport.

However, even if there are no physical or geographical limitations, carrying small amounts of cargo on a relatively large vehicle from origin to destination is inefficient. One way to improve efficiency is consolidation. The aim of intermodal transportation is exactly this, that is, to consolidate loads for efficient long-haul transportation (e.g., by rail or large ocean vessels), while taking advantage of the efficiency of local pick-up and delivery operations by truck. For any shipment to be intermodal, freight must be transported from origin to destination by a sequence of at least two modes of transport, where transfers from one mode to another take place at intermodal terminals (such as sea ports or rail yards). An illustrative example is provided here.

Example:
A multimodal transportation network is illustrated in the figure below, on which three modes of transportation, namely road (trucks), rail and water (shipping), operate. Cargo originates from the two facilities seen on the far left of the figure and is destined to the warehouses appearing on the far right of the figure. The figure also shows how the cargo is shipped using an intermodal chain of modes. In particular, loaded trucks leave the origin facilities to a rail yard, where they are consolidated into a train and sent to another rail yard. Trucks are again used to transport the containers from this rail yard to a sea container terminal. This last operation may not be necessary if the sea container terminal has an interface to the rail network, in which case freight is transferred directly from the former onto the latter. Containers are then transported to a port on another continent by ocean shipping, from where they leave by either trucking or rail (or both) to their destinations.



One of the advantages of intermodal freight transport is the possibility of modal shift, which is defined as partially or fully transferring freight from one mode to the other in the network when it is being shipped to its final destination.

Intermodal Terminals

In intermodal transportation, transfers from one mode to another, and all associated operations related to, for example, loading and unloading, temporary storage or intermediate buffer, and even pre-delivery inspection or enhancement work on the goods are carried out at intermodal terminals. Here are some examples of intermodal terminals:

- ***Rail yards*** where transfers between the two modes of road (e.g. trucks and lorries) and rail take place.
- ***Sea ports*** that act as the interface between sea transport (vessels, ferries) and land transport (road or rail). A barge terminal is a smaller port used in inland river transportation and is linked to larger sea terminals.
- ***Inland distribution*** centers that are used only for road transportation, and act as transfer points between different modes of road transport, for example lorries and vans. Warehouses and cross-docking centres are examples to inland distribution centers.

Containerized Transport

Cargo flow in an intermodal transportation network can be carried by a variety of means, for example crates, pallets or boxes for fresh produce, specialized vessels or tanks for liquid bulk cargo. Irregularity of general cargo, in shape or size, has led to the development of standardized units of shipment, which came to be known as containers. A container is a steel or Aluminium box, with a standard size of 20 ft in length, known as the Twenty Foot Equivalent Unit (TEU), is widely used around the globe, but 40 and 45 ft containers are more commonly used in North America. Containers offer a number of advantages over non-containerized cargo, which are summarized as follows:

1. Containers can be locked, which means that their contents cannot easily be modified except at origin or destination, which implies increased safety and reduced loss and damage.
2. Due to its standard structure, transfer operations at terminals can also be tailored to process containerized cargo fast and with minimal amount of effort. Container ports are excellent examples where such operations can be seen, using container-specific equipment such as container gantry cranes, implying reduced cargo handling.
3. Containers can be stacked, which implies a more efficient use of storage space within the port.
4. Cargo of various sizes, shapes, dimensions or weight can be carried in a container, making it a flexible enough means of transport to accommodate for a variety of cargo types.

Due to these advantages, intermodal transportation, as it is implemented and used today, heavily relies on containerization, to the point that intermodal transportation is often equated to moving containers over long distances, on multimodal networks. Intermodal transportation is not restricted, however, to containers and intercontinental exchanges.

6.3 Choice of Carrier and Transportation Mode

There are a number of factors that shippers use to decide which particular mode or modes of transportation to use for shipping their freight. Earlier studies on this topic have identified six main factors that influence a shipper's decision:

- Freight rates (including cost and charges)
- Reliability (delivery time)
- Transit times (time-in-transit, speed, delivery time)
- Over, short and damaged shipments (loss, damage, claims processing, tracing)
- Shipper market considerations (customer service, user satisfaction, market competitiveness, market influences)
- Carrier considerations (availability, capability, reputation, special equipment)

Later research identified frequency and flexibility as additional factors in particular industries, and that factors such as international dimension, economies of scale, security, environmental concerns and energy use, integration with the supply chain and information technologies, and particularly the use of the Internet are all relevant to the transportation mode choice but remain as under-researched areas.

Other studies have recognized that the perception of a mode of transport is likely to have a significant effect on the choice of mode. Of the more prominent ones, earlier research has identified communication, quality of service, consistency (in delivery), transit times and rates as being the main factors, with the first two being more significant than the others.

6.4 Shipment Options

There are at least two ways in which shipments can be made, which are described in more detail here.

6.4.1 Direct or Customized Shipments

Direct or customized shipments arise when goods are shipped from the source of supply to the demand point without the use of any intermediate facilities or demand points. Full-load trucking is an example of customized transportation, where, upon the call of a customer wishing to have goods shipped from one (source) point to another (destination) point, a dispatcher assigns a truck to this task. The truck travels to the customer location, is loaded and then moves to the destination, where it is unloaded. Following this, the driver is assigned a new task by the dispatcher, kept waiting until a new demand appears in the near future, or repositioned to a location where a load exists or is expected to be available about the arrival time. The advantages of full-load trucking come from its flexibility in adapting to a highly dynamic environment and uncertain future demands, offering reliability in service and low tariffs compared to other modes of transportation. The fill efficiency of full-load trucking is achieved through the implementation of resource management and allocation strategies that seek to make the best use of the available resources, while maximizing the volume of demand satisfied and the associated profits. Customized services are also offered, for example, by charted sea or river vessels and planes.

6.4.2 Consolidated Shipments

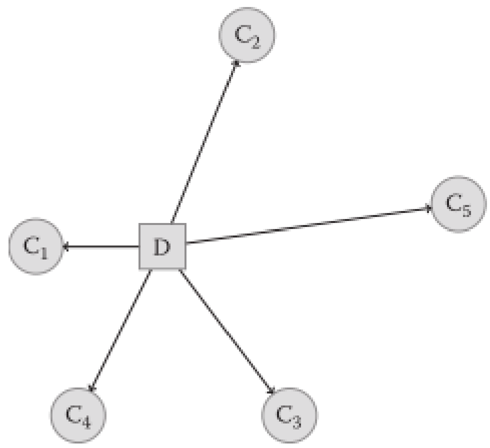
In many cases, trade-offs between volume and frequency of shipping, along with the cost of transportation, render customized services impractical. In such cases, consolidation is used for serving a number of demand points using a particular service. Freight consolidation transportation is performed by Less-Than-Truckload (LTL) motor carriers, railways, ocean shipping lines, regular and express postal services, etc. One example of LTL transportation is when the loads of a given set of customers are consolidated at a given source (e.g. a depot), and the deliveries are multiple destination points. In this case, a truck or a lorry is loaded with all the goods at the depot and is dispatched to visit the customers in a particular order to perform the deliveries. A consolidation-based transportation system can also be structured as a hub-and-spoke network, where shipments for a number of origin-destination points may be transferred via intermediate consolidation facilities, or hubs, such as airports, seaport container terminals, rail yards, truck break-bulk terminals and intermodal platforms.

6.5 Distribution Structures

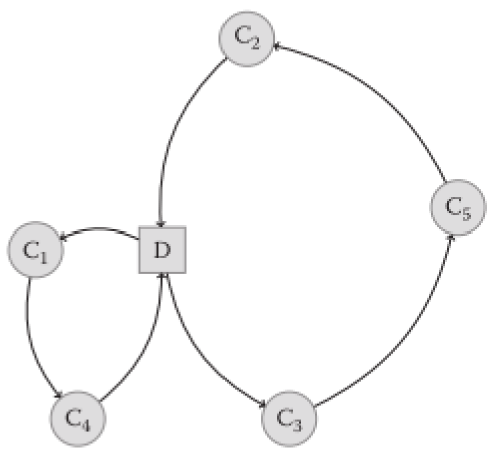
The structure of the way in which goods are distributed is generally in the form of single-echelon (or single-tier) or multi-echelon (or multi-tier) distribution network. The type of structure used depends on a number of factors, such as the type of area (e.g. urban, suburban, rural) or the size of the area (e.g. countries, continents), the type of product(s) being shipped, the types of vehicles used and the demand requirements in terms of volume and time.

6.5.1 Single-echelon

A single-echelon distribution structure does not involve the use of any intermediate facilities between the source(s) of supply and the source(s) of demand and operates on the basis that deliveries are made from the former to the latter. This might be in the form of direct shipments from the supply point to the customer, or, as in the case of consolidation, from one supply point to many customers.

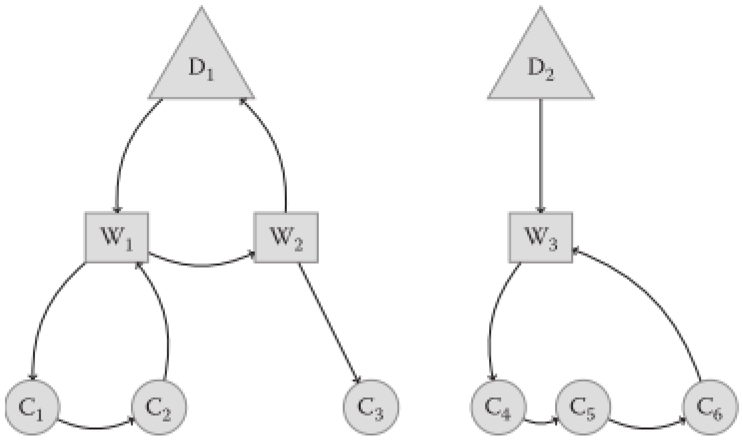


The figure above shows a single echelon distribution structure where a central depot, shown by D, serves five customers labeled C_1–C_5, all using direct shipments. In this example, it is assumed that each link carries at least one full-truckload shipment. The figure below shows the same network in which it is assumed that the deliveries can be consolidated. In particular, one LTL vehicle serves customers C_1 and C_4 in the given order, and the other serves customers C_3,C_5 and C_2 in the given order.



6.5.2 Multi-echelon

If the distribution network includes several layers of intermediate facilities, such as warehouses, consolidation facilities, distribution centres and cross-docks, to which the goods are sent as part of the distribution from their origin(s) to their destination(s), then the network is said to have a multi-echelon structure. Supply chain networks often have such a structure, for which the planning problems entail the management and operation of the facilities in the network as well as the assignment of (aggregated) flow between the facilities. In the context of freight distribution, however, the planning problems are more specific to the management of transport modes and the vehicle fleets used in between the different layers.



A two-echelon network distribution structure is shown in the figure above, where the first echelon includes all movements between the first layer of depots shown by D_1 and D_2 and the second layer of intermediate points (e.g. warehouses) shown by W_1, W_2 and W_3. In this example, consolidated shipments depart from D1 to serve the two warehouses W_1 and W_2. In contrast, a direct shipment using a full-truckload is made between D_2 and W_3. The second echelon consists of all the distribution activities between the second layer and the third layer, where the latter includes six demand points shown by C_1–C_6. It is clear from the figure that this network uses a mix of consolidated and direct shipments at this echelon.

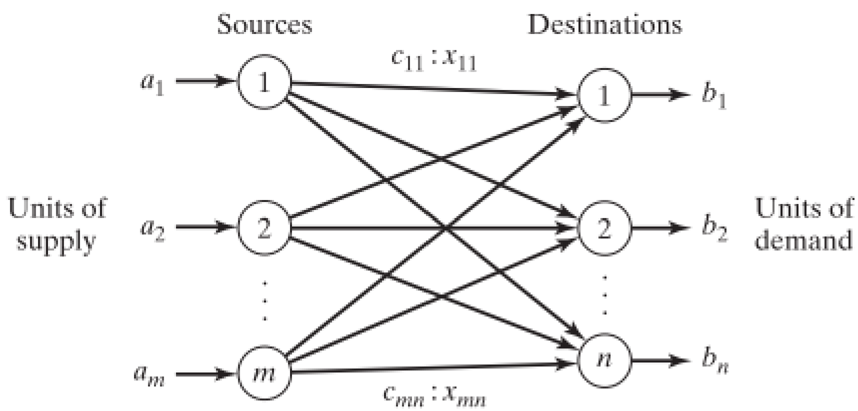
Structures such as those shown in the figure above have been suggested for use within urban areas, where the first layer of depots consists of those that are located outside the boundaries of the urban zone and the second layer consists of intermediate points located on the boundaries. The expectation is that bulk deliveries are made from the first to the second layer using heavy goods vehicles, which would stop here. The distribution activities within the urban zone would then be carried out using smaller vehicles that would operate between the second layer of intermediate points and the third layer of demand points. Such a strategy prevents large and heavy vehicles from entering urban zones.

6.6 Transportation Model

The transportation model deals with a special class of linear programming problem in which the objective is to transport a homogeneous commodity from various origins or factories to different destinations or markets at a total minimum cost.

6.6.1 Definition of the Transportation Model

The problem is represented by the network in the figure below. There are m sources and n destinations, each represented by a node. The arcs represent the routes linking the sources and the destinations. Arc (m, n) joining source m to destination n carries two pieces of information: the transportation cost per unit, c_{mn}, and the amount shipped, x_{mn}. The amount of supply at source m is a_m, and the amount of demand at destination n is b_n. The objective of the model is to minimize the total transportation cost while satisfying all the supply and demand restrictions.



Example #1:

MG Auto has three plants in Los Angeles, Detroit, and New Orleans and two major distribution centers in Denver and Miami. The quarterly capacities of the three plants are 1000, 1500, and 1200 cars, and the demands at the two distribution centers for the same period are 2300 and 1400 cars. The transportation costs per car on the different routes, rounded to the closest dollar is given in the table below. Determine the amount of cars to distributed to each distribution center such that the cost transportation cost is minimum, satisfying all demand and supply constraints.

Transportation cost per car		
From/To	Denver(1)	Miami(2)
LA (1)	\$80	\$215
Detroit (2)	\$100	\$108
New Orleans (1)	\$102	\$68

Solution:

The linear Programming model of the problem is:

Minimize, $z = 80x_{11} + 215x_{12} + 100x_{21} + 108x_{22} + 102x_{31} + 68x_{32}$

Subject to:

$x_{11} + x_{12} + 0 + 0 + 0 + 0 = 1000$ (LA)
$0 + 0 + x_{21} + x_{22} + 0 + 0 = 1500$ (Detroit)
$0 + 0 + 0 + 0 + x_{31} + x_{32} = 1200$ (New Orleans)
$x_{11} + 0 + x_{21} + 0 + x_{31} + 0 = 2300$ (Denver)
$0 + x_{12} + 0 + x_{22} + 0 + x_{32} = 1400$ (Miami)
$x_{ij} \geq 0$, $i = 1, 2, 3$ and $j = 1, 2$

All the constraints are equations because the total supply (= 1000+ 1500 + 1200 = 3700) equals the total demand (= 2300 + 1400 = 3700).

From / To	Denver	Miami	**Supply**
LA	80 x_{11}	215 x_{12}	1000
Detroit	100 x_{21}	108 x_{22}	1500
New Orleans	102 x_{31}	68 x_{32}	1200
Demand	2300	1400	

The special structure of the transportation problem allows a compact representation of the problem using the transportation tableau format in shown above.

The optimal solution of the given transportation problem ships 1000 cars from LA to Denver ($x_{11} = 1000$), 1300 from Detroit to Denver ($x_{21} = 1300$), 200 from Detroit to Miami ($x_{22} = 200$), and 1200 from New Orleans to Miami ($x_{31} = 1200$). The associated minimum transportation cost is computed as $100 \times \$80 + 1300 \times \$100 + 200 \times \$108 + 1200 \times \$68 = \$313,200$.

A transportation problem is said to be balanced if the total supply equals to the total demand. However, if the problem is not balanced then it can be balanced by introducing a dummy row or column having cost value of zero, i.e:

- ***If Total Supply*** $<$ ***Total Demand:*** A dummy row, whose each cell have zero cost value, is added to the table with a corresponding supply value = Total Demand - Total Supply.
- ***If Total Supply*** $>$ ***Total Demand:*** A dummy column, whose each cell have zero cost value, is added to the table with a corresponding demand value = Total Supply - Total Demand.

6.6.2 The Transportation Algorithm

The special transportation algorithm that will be presented in this section was developed early on when hand computations were the norm and shortcuts were warranted. Today, powerful computer codes can solve transportation models of any size as a regular LP. But there is more to the transportation model than the hand computations. First, its historical significance in the evolution of OR techniques is important and must be preserved. Second, the special transportation tableau format can facilitate modeling a number of situations that do not deal directly with transporting goods. Third, the algorithmic hand computations boast such (almost intuitive) simplicity that a beginner can get a “feel” of what optimization is about. Lastly, the transportation algorithm does provide insight into the use of the theoretical primal–dual relationships to achieve a practical end result—that of developing simple rules for hand computations. The exercise is theoretically intriguing.

The basic steps of the Transportation Algorithm are as follows:

- **Step 1**: Balance the given transportation problem.
- **Step 2**: Compute the starting basic feasible solution. A basic feasible solution can be computed using either of these methods:
    1. North - west corner method.
    2. Least - cost cell method. (Or Inspection method Or Matrix minimum - row minimum - column minimum method).

3. Vogel's Approximation Method, generally known as VAM.

- **Step 3**: After getting the basic feasible solution conduct a optimality test to check whether the solution is optimal or not. There are two methods of computing optimality test, they are:

  1. Stepping Stone Method.
  2. Modified Distribution Method, generally known as MODI method.

For this subject, we will be only computing up to the basic feasible solution of the transportation problem since solving for optimality test is a cumbersome task on its own and doing so causes to deviate away from the course domain of the subject. These topics are only introduced for the sake of familiarizing students with the scope of transportation engineering. For an in-depth study about these topics, students are highly encouraged to peruse books related to Operation Research.

6.6.3 Determining the Basic Feasible Solution

a. North-West Corner Method

The following are the steps used for determining the basic feasible solution of a transportation problem using the North-West Corner Method:

- Balance the problem, that is, check whether the total demand is equal to the total demand or not. If not, open a dummy column or dummy row as the case may be and balance the problem.

- Start from the left hand side top corner or cell and make allocations depending on the availability and requirement constraint. If the availability constraint is less than the requirement constraint, then for that cell make allocation in units which is equal to the availability constraint. In general, verify which is the smallest among the availability and requirement and allocate the smallest one to the cell under question. Then proceed allocating either side ways or downward to satisfy the rim requirement. Continue this until all the allocations are over.

- Once all the allocations are over, i.e., both rim requirement (column and row i.e., availability and requirement constraints) are satisfied, write allocations and calculate the cost of transportation.

Numerical Example:(Sunray Transport Problem)
SunRay Transport Company ships truckloads of grain from three silos to four mills. The supply (in truckloads) and the demand (also in truckloads) together with the unit transportation costs per truckload on the different routes are summarized in the table below. The unit transportation costs are in hundreds of dollars. The model seeks the minimum cost shipping schedule between the silos and the mills.

Transportation Cost

Silo / Mill	1	2	3	4	Supply
1	\$10	\$2	\$20	\$11	15
2	\$12	\$7	\$9	\$20	25
3	\$4	\$14	\$16	\$18	10
Demand	5	15	15	15	

Solution:
Total Supply = 15 + 25 + 10 = 50
Total Demand = 5 + 15 + 15 + 15 = 50

∵ Total Supply = Total Demand, the transportation model is balanced.

At first, we choose the cell that is in the extreme north-west position, that is, cell corresponding to transport operation from silo 1 to mill 1. For that cell, we have supply of 15 and demand of 5. ∵ 5 < 15, we allocate the value of 5 this cell. Allocating 5 truckloads from silo 1 to mill 1 leaves silo 1 with a remainder of 10 truckloads of supply. Also demand for mill 1 has been satisfied so we gray out the column corresponding to mill 1 indicating it requires no further action. This is shown in the transportation tableau below:

Silo / Mill	1	2	3	4	**Supply**
1	5 \$10	\$2	\$20	\$11	~~15~~ 10
2	\$12	\$7	\$9	\$20	25
3	\$4	\$14	\$16	\$18	10
Demand	~~5~~ 0	15	15	15	

Similarly, cell corresponding to transport operation from silo 1 to mill 2 is in the north-west position considering the remaining table (Table excludes the grayed out region). For this cell, supply is 10 and demand is 15. ∵ 10 < 15, we allocate value of 10 for this cell. Allocating 10 truckloads from silo 1 leaves the supply capacity of silo 1 to 0 and demand of mill 2 is reduced to 5 truckloads. Since silo 1 can no longer supply any truckloads, we gray out the corresponding row.

Silo / Mill	1	2	3	4	**Supply**
1	5 $10	10 $2	$20	$11	~~15~~ ~~10~~ 0
2	$12	$7	$9	$20	25
3	$4	$14	$16	$18	10
Demand	~~5~~ 0	~~15~~ 5	15	15	

Similarly, cell corresponding to transport operation from silo 2 to mill 2 is in the north-west position considering the remaining table. For this cell, supply is 25 and demand is 5. $\because 5 < 25$, we allocate value of 5 for this cell. Allocating 5 truckloads from silo 2 reduces the supply capacity of silo 2 to 20 and demand of mill 2 is reduced to 0 truckloads. Since mill 2 can no longer receive truckloads, we gray out the corresponding column.

Silo / Mill	1	2	3	4	**Supply**
1	5 $10	10 $2	$20	$11	~~15~~ ~~10~~ 0
2	$12	5 $7	$9	$20	~~25~~ 20
3	$4	$14	$16	$18	10
Demand	~~5~~ 0	~~15~~ ~~5~~ 0	15	15	

Similarly, cell corresponding to transport operation from silo 2 to mill 3 is in the north-west position considering the remaining table. For this cell, supply is 20 and demand is 15. $\because 15 < 20$, we allocate value of 15 for this cell. Allocating 15 truckloads from silo 2 reduces the supply capacity of silo 2 to 5 and demand of mill 3 is reduced to 0 truckloads. Since mill 3 can no longer receive truckloads, we gray out the corresponding column.

Silo / Mill	1	2	3	4	**Supply**
1	5 \$10	10 \$2	\$20	\$11	~~15~~ ~~10~~ 0
2	\$12	5 \$7	15 \$9	\$20	~~25~~ ~~20~~ 5
3	\$4	\$14	\$16	\$18	10
Demand	~~5~~ 0	~~15~~ ~~5~~ 0	~~15~~ 0	15	

Similarly, cell corresponding to transport operation from silo 2 to mill 4 is in the north-west position considering the remaining table. For this cell, supply is 5 and demand is 15. $\because 5 < 15$, we allocate value of 5 for this cell. Allocating 5 truckloads from silo 2 reduces the supply capacity of silo 2 to 0 and demand of mill 3 is reduced to 10 truckloads. Since, silo 2 can no longer supply any truckloads, we gray out the corresponding row.

Silo / Mill	1	2	3	4	**Supply**
1	5 \$10	10 \$2	\$20	\$11	~~15~~ ~~10~~ 0
2	\$12	5 \$7	15 \$9	5 \$20	~~25~~ ~~20~~ ~~5~~ 0
3	\$4	\$14	\$16	\$18	10
Demand	~~5~~ 0	~~15~~ ~~5~~ 0	~~15~~ 0	~~15~~ 10	

Finally, there is only one cell left, cell corresponding to transport operation from silo 3 to mill 4, so we assign the remaining 10 truckloads to this cell and the transportation supply and demand constraints are also met.

From the basic feasible solution obtained using the North-West Corner method, following assignments are obtained: $x_{11} = 5$, $x_{12} = 10$, $x_{22} = 5$, $x_{23} = 15$, $x_{24} = 5$, and $x_{34} = 10$.

Also, the associated cost of the schedule is,
$z = 5 \times \$10 + 10 \times \$2 + 5 \times \$7 + 15 \times \$9 + 5 \times \$20 + 10 \times \$18 = \$520$

b. Least-Cost Cell Method

The least-cost method finds a better starting solution by targeting the cheapest routes. It assigns as much as possible to the cell with the smallest unit cost (ties are broken arbitrarily). Next, the satisfied row or column is crossed out and the amounts of supply and demand are adjusted accordingly. If both a row and a column are satisfied simultaneously, only one is crossed out, the same as in the northwest-corner method. Next, select the uncrossed-out cell with the smallest unit cost and repeat the process until exactly one row or column is left uncrossed out.

The following are the steps used for determining the basic feasible solution of a transportation problem using the Least-Cost Cell Method:

- *Step 1:* Identify the lowest cost cell in the given matrix. When more than one cell has the same cost, then both the cells are competing for allocation. This situation in transportation problem is known as tie. To break the tie, select any one cell of your choice for allocation. Make allocations to this cell either to satisfy availability constraint or requirement constraint. Once one of these is satisfied, then mark crosses ($\times$) in all the cells in the row or column which ever has completely allocated.

- *Step 2:* Are all allocations completed? If yes the basic solution has been obtained, else go to the *Step 1.*

Numerical Example: Determine the basic feasible solution for the Sunray Transport Problem using the Least-Cost Cell method.
Solution:
Total Supply = 15 + 25 + 10 = 50
Total Demand = 5 + 15 + 15 + 15 = 50

$\because$ Total Supply = Total Demand, the transportation model is balanced.

First, we have to choose the cell having the least cost, in this problem the cell corresponding to transport operation from silo 1 to mill 2 has the least cost (\$2). For this cell, the supply is 15 and demand is also 15. So we assign the value of 15 for this cell, which reduces the supply of silo 1 to 0 and demand of mill 2 to 0. Since silo 1 can no longer supply and mill 2 can no longer receive truckloads, we gray out the corresponding row and column.

Silo / Mill	1	2	3	4	**Supply**
1	$10	15 $2	$20	$11	~~15~~ 0
2	$12	$7	$9	$20	25
3	$4	$14	$16	$18	10
Demand	5	~~15~~ 0	15	15	

Similarly for the remaining active cells, the cell corresponding to transport operation from silo 3 to mill 1 has the least cost ($4) among other active cells. For this cell, the supply is 10 and demand is 5. Since $5 < 10$ we assign the value of 5 for this cell, which reduces the supply of silo 3 to 5 and demand of mill 1 to 0. Since mill 1 can no longer receive truckloads, we gray out the corresponding column.

Silo / Mill	1	2	3	4	**Supply**
1	$10	15 $2	$20	$11	~~15~~ 0
2	$12	$7	$9	$20	25
3	5 $4	$14	$16	$18	~~10~~ 5
Demand	~~5~~ 0	~~15~~ 0	15	15	

Similarly for the remaining active cells, the cell corresponding to transport operation from silo 2 to mill 3 has the least cost ($9) among other active cells. For this cell, the supply is 25 and demand is 15. Since $15 < 25$ we assign the value of 15 for this cell, which reduces the supply of silo 2 to 10 and demand of mill 3 to 0. Since mill 3 can no longer receive truckloads, we gray out the corresponding column.

Silo / Mill	1	2	3	4	**Supply**
1	\$10	15 \$2	\$20	\$11	$\cancel{15}$ 0
2	\$12	\$7	15 \$9	\$20	$\cancel{25}$ 10
3	5 \$4	\$14	\$16	\$18	$\cancel{10}$ 5
Demand	$\cancel{5}$ 0	$\cancel{15}$ 0	$\cancel{15}$ 0	15	

Similarly for the remaining active cells, the cell corresponding to transport operation from silo 3 to mill 4 has the least cost (\$18) among other active cells. For this cell, the supply is 5 and demand is 15. Since $5 < 15$ we assign the value of 5 for this cell, which reduces the supply of silo 3 to 0 and demand of mill 4 to 10. Since silo 3 can no longer supply truckloads, we gray out the corresponding row.

Silo / Mill	1	2	3	4	**Supply**
1	\$10	15 \$2	\$20	\$11	$\cancel{15}$ 0
2	\$12	\$7	15 \$9	\$20	$\cancel{25}$ 10
3	5 \$4	\$14	\$16	5 \$18	$\cancel{10}$ $\cancel{5}$ 0
Demand	$\cancel{5}$ 0	$\cancel{15}$ 0	$\cancel{15}$ 0	$\cancel{15}$ 10	

Finally the cell corresponding to transport operation from silo 2 to mill 4 is only left which has a both demand and supply value of 10. Hence we assign the value of 10 to this cell.

From the basic feasible solution obtained using the North-West Corner method, following assignments are obtained: $x_{13} = 5, x_{12} = 15, x_{23} = 15, x_{24} = 10, and\, x_{34} = 5$.

Also, the associated cost of the schedule is,
$z = 15 \times \$2 + 5 \times \$4 + 15 \times \$9 + 10 \times \$20 + 5 \times \$18 = \ \$475.$

c. Vogel's Approximation Method (Opportunity Cost Method)

VAM is an improved version of the least-cost method that generally, but not always, produces better starting solutions. In this method, we use concept of opportunity cost. Opportunity cost is the penalty for not taking correct decision. This is the penalty for not taking correct decision and hence the opportunity cost. The following are the steps used for determining the basic feasible solution of a transportation problem using the Least-Cost Cell Method:

- Step 1: For each row (column), determine a penalty measure by subtracting the smallest unit cost in the row (column) from the next smallest unit cost in the same row (column). This penalty is actually a measure of lost opportunity one forgoes if the smallest unit cost cell is not chosen.
- Identify the row or column with the largest penalty, breaking ties arbitrarily. Allocate as much as possible to the variable with the least unit cost in the selected row or column. Adjust the supply and demand, and cross out the satisfied row or column. If a row and a column are satisfied simultaneously, only one of the two is crossed out, and the remaining row (column) is assigned zero supply (demand).
- 1. If exactly one row or column with zero supply or demand remains uncrossed out, stop.
  2. If one row (column) with positive supply (demand) remains uncrossed out, determine the basic variables in the row (column) by the least-cost method. Stop.
  3. If all the uncrossed-out rows and columns have (remaining) zero supply and demand, determine the zero basic variables by the least-cost method. Stop.
  4. Otherwise, go to step 1.

Numerical Example: Determine the basic feasible solution for the Sunray Transport Problem using the Vogel's Approximation method.
Solution:
Total Supply = 15 + 25 + 10 = 50
Total Demand = 5 + 15 + 15 + 15 = 50

$\because$ Total Supply = Total Demand, the transportation model is balanced.

At first, we compute the values of penalties for each row and column of the transportation model. For first row (supply row of silo 1), 10 and 2 are the two minimum values. Thus, the value of penalty for the corresponding row is 8. For first column (demand column of mill 1), 10 and 4 are the two minimum value. Thus, the value of penalty for the corresponding column is 6. Similarly the value of penalty for all rows and columns are computed as shown in the table below. Comparing the values of penalty of all rows and column, it is found that the value of penalty corresponding to the third row, that is 10, is the maximum value. Since 10 is the max value, we determine the cell to be assigned for the corresponding row using least-cost cell method.

S/M	1	2	3	4	Supply	P_1
1	\$10	\$2	\$20	\$11	15	10 - 2 = 8
2	\$12	\$7	\$9	\$20	25	9 - 7 = 2
3	5 \$4	\$14	\$16	\$18	$\cancel{10}$ 5	14 - 4 = 10 ←
Demand	$\cancel{5}$ 0	15	15	15		
P_1	10 - 4 = 6	5	7	7		

Similarly, we compute the value of penalty for all active rows and columns. Since 9, which is the penalty value for row 1, is the maximum value among all values of penalty, we determine the cell to be assigned for the corresponding row using least-cost cell method.

S/M	1	2	3	4	Sup	P_1	P_2
1	\$10	15 \$2	\$20	\$11	$\cancel{15}$ 0	8	9 ←
2	\$12	\$7	\$9	\$20	25	2	2
3	5 \$4	\$14	\$16	\$18	$\cancel{10}$ 5	10 ←	2
Dem	$\cancel{5}$ 0	$\cancel{15}$ 0	15	15			
P_1	6	5	7	7			
P_2	-	5	7	7			

Similarly,

S/M	1	2	3	4	Sup	P_1	P_2	P_3
1	$10	15 $2	$20	$11	~~15~~ 0	8	9 ←	-
2	$12	$7	15 $9	$20	~~25~~ 10	2	2	11 ←
3	5 $4	$14	$16	$18	~~10~~ 5	10 ←	2	2
Dem	~~5~~ 0	~~15~~ 0	~~15~~ 0	15				
P_1	6	5	7	7				
P_2	-	5	7	7				
P_3	-	-	7	2				

Similarly,

S/M	1	2	3	4	Sup	P_1	P_2	P_3	P_4
1	$10	15 $2	$20	$11	~~15~~ 0	8	9 ←	-	-
2	$12	$7	15 $9	10 $20	~~25~~ ~~10~~ 0	2	2	11 ←	20 ←
3	5 $4	$14	$16	$18	~~10~~ 5	10 ←	2	2	18
Dem	~~5~~ 0	~~15~~ 0	~~15~~ 0	~~15~~ 5					
P_1	6	5	7	7					
P_2	-	5	7	7					
P_3	-	-	7	2					
P_4	-	-	-	2					

The basic feasible solution is shown below:

Silo / Mill	1	2	3	4	Supply
1	10	15 $2	20	11	15
2	12	7	15 $9	10 $20	25
3	5 $4	14	16	5 $18	10
Demand	5	15	15	15	

The associated cost of the schedule is,
$z = 15 \times \$2 + 5 \times \$4 + 15 \times \$9 + 10 \times \$20 + 5 \times \$18 =$ $475.

North-West Corner Method	Least-Cost Cell Method	Vogel's Approximation Method
1. The allocation is made from the left hand side top corner irrespective of the cost of the cell.	The allocations are made depending on the cost of the cell. Lowest cost is first selected and then next highest etc.	The allocations are made depending on the opportunity cost of the cell.
2. As no consideration is given to the cost of the cell, naturally the total transportation cost will be higher than the other methods.	As the cost of the considered while cell is making allocations, the total cost of transportation will be comparatively less.	As the allocations are made depending on the opportunity cost of the cell, the basic feasible solution obtained will be very nearer to optimal solution.
3. It takes less time. This method is suitable to get basic feasible solution quickly.	The basic feasible solution, we get will be very nearer to optimal solution. It takes more time than northwest coroner method.	It takes more time for getting basic Feasible solution. But the solution we get will be very nearer to Optimal solution.
4. When basic feasible solution alone is asked, it is better to go for northwest corner method.	When optimal asked, better solution to go is for least-cost cell method for basic feasible solution and MODI for optimal solution.	VAM and MODI is the best option to get optimal solution.

6.7 Assignment Model

An assignment problem is a particular case of transportation problem where the objective is to assign a number of resources to an equal number of activities so as to minimize total cost or maximize total profit of allocation.

The problem of assignment arises because available resources such as men, machines etc. have varying degrees of efficiency for performing different activities, therefore, cost, profit or loss of performing the different activities is different.

Thus, the problem is "How should the assignments be made so as to optimize the given objective". Some of the problem where the assignment technique may be useful are assignment of workers to machines, salesman to different sales areas.

6.7.1 Definition of the Assignment Problem

Suppose there are n jobs to be performed and n persons are available for doing these jobs. Assume that each person can do each job at a term, though with varying degree of efficiency, let c_{ij} be the cost if the i-th person is assigned to the j-th job. The problem is to find an assignment (which job should be assigned to which person one on-one basis) So that the total cost of performing all jobs is minimum, problem of this kind are known as assignment problem.

For solving the assignment problem we use Assignment technique or Hungarian method or Flood's technique. All are one and the same. Above, it is mentioned that one origin is to be assigned to one destination. This feature implies the existence of two specific characteristics in linear programming problems, which when present, give rise to an assignment problem. The first one being the pay of matrix for a given problem is a square matrix and the second is the optimum solution (or any solution with given constraints) for the problem is such that there can be one and only one assignment in a given row or column of the given payoff matrix. The transportation model is a special case of linear programming model (Resource allocation model) and assignment problem is a special case of transportation model, therefore it is also a special case of linear programming model. Hence it must have all the properties of linear programming model. That is it must have: (i) an objective function, (ii) it must have structural constraints, (iii) It must have non-negativity constraint and (iv) The relationship between variables and constraints must have linear relationship. In our future discussion, we will see that the assignment problem has all the above properties.

There are different types of assignment problems. They are:

1. Assigning the jobs to machines when the problem has square matrix to minimize the time required to complete the jobs. Here the number of rows i.e. jobs are equals to the number of machines i.e. columns.

2. The second type is maximization type of assignment problem. Here we have to assign certain jobs to certain facilities to maximize the returns or maximize the effectiveness.

3. Assignment problem having non-square matrix. Here by adding a dummy row or dummy columns as the case may be, we can convert a non-square matrix into a square matrix and proceed further to solve the problem.

4. Assignment problem with restrictions. Here restrictions such as a job cannot be done on a certain machine or a job cannot be allocated to a certain facility may be specified. In such cases, we should neglect such cell or give a high penalty to that cell to avoid that cell to enter into the program.

5. Traveling sales man problem (cyclic type). Here a salesman must tour certain cities starting from his hometown and come back to his hometown after visiting all cities.

6.7.2 Hungarian Method

The Hungarian method (developed by Hungarian mathematician D. Konig) is an efficient method of finding the optimal solution of an assignment problem without making a direct comparison of every solution. The method works on the principle of reducing the given cost matrix to a matrix of opportunity costs. Opportunity costs show the relative penalties associated with assigning a resource to an activity. Hungarian method reduces the cost matrix to the extent of having at least one zero in each row and column so as to make optimal assignments.

The Hungarian method (minimization case) can be summarized in the following steps:

Step A: Develop the cost matrix from the given problem If the number of rows are not equal to the number of columns, then add required number of dummy rows or columns. The cost element in dummy rows/columns are always zero.

Step B: Compute the opportunity cost matrix 1. Identify the smallest element in each row of cost matrix and then subtract it from each element of that row.

2. In the reduced matrix obtained from 2(a), identify the smallest element in each column and then subtract it from each element of that column. Each row and column now have at least one zero element.

Step C: Make assignments in the opportunity cost matrix The procedure of making assignments is as follows:

1. First round for making assignments:
   (a) Identify rows successively from top to bottom until a row with exactly one zero element is found. Make an assignment to this single zero by enclosing it with a square bracket ,([]), and then cross off, (×), all other zeros in the corresponding column.
   (b) Identify columns successively from left to right hand with exactly one zero element that has not been assigned. Make an assignment to this single zero by enclosing it with a square bracket ,([ ]), and then cross off, (×), all other zeros in the corresponding row.
2. Second round for making assignments:
   (a) If a row and/or column has two or more unmarked zeros and one cannot be chosen by inspection, then choose zero element arbitrarily for assignment.
   (b) Repeat sub-steps (1) and (2) successively until one of the following situations arise

Step D: Optimality Critereon The following are the steps involved in determining the optimality of the solution:

1. If all zero elements in the cost matrix are either marked with square bracket ([ ]) or are crossed off (×) and there is exactly one assignment in each row and column, then it is an optimal solution. The total cost associated with this solution is obtained by adding the original cost elements in the occupied cells.
2. If a zero element in a row or column was chosen arbitrarily for assignment in Step D(1), there exists an alternative optimal solution.
3. If there is no assignment in a row (or column), then this implies that the total number of assignments are less than the number of rows/columns in the square matrix. In such a situation proceed to Step E.

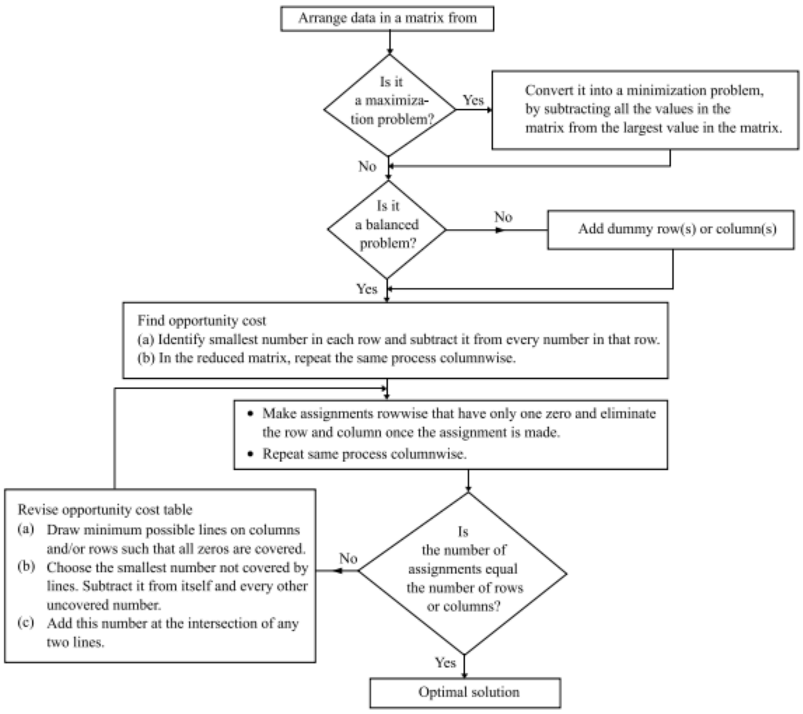
Step E: Revise the opportunity cost matrix Draw a set of horizontal and vertical lines to cover all the zeros in the revised cost matrix obtained from Step C, by using the following procedure:

1. For each row in which no assignment was made, mark a tick (✓)
2. Examine the marked rows. If any zero element is present in these rows, mark a tick (✓) to the respective columns containing zeros.
3. Examine marked columns. If any assigned zero element is present in these columns, tick (✓) the respective rows containing assigned zeros.
4. Repeat this process until no more rows or columns can be marked.
5. Draw a straight line through each marked column and each unmarked row.

Step F: Develop the new revised opportunity cost matrix The steps are as follows:

1. Among the elements in the matrix not covered by any line, choose the smallest element. Call this value k.
2. Subtract k from every element in the matrix that is not covered by a line.
3. Add k to every element in the matrix covered by the two lines, i.e. intersection of two lines.
4. Elements in the matrix covered by one line remain unchanged.

Step G: Repeat Repeat steps C to F until an optimal solution is obtained.



Numerical Example: Time taken by trucks A, B, C, D, and E to complete jobs 1, 2, 3 , 4, and 5 are given in the table below and each truck is to be assigned a job. Determine the job assignment for each truck such that the time taken to complete all the jobs is minimum.

Jobs / Truck	A	B	C	D	E
1	13	8	16	18	19
2	9	15	24	9	12
3	12	9	4	4	4
4	6	12	10	8	13
5	15	17	18	12	20

Solution:

Rows 1, 2, 3, 4, and 5 have minimum value of 8, 9, 4, 6 , and 12 respectively. Using these values we find out the row reduction matrix.

	A	*B*	*C*	*D*	*E*	
1	5	0	8	10	11	
2	0	6	15	0	3	
3	8	5	0	0	0	
4	0	6	4	2	7	
5	3	5	6	0	8	

Similarly, all the columns have minimum value of 0. Using these values we find out the column reduction matrix.

	A	B	C	D	E	
1	5	0	8	10	11	
2	0	6	15	0	3	
3	8	5	0	0	0	
4	0	6	4	2	7	
5	3	5	6	0	8	

Then, rows 1, 4, and 5 have only one 0, so we assign these cells with square bracket and cross out any other value of 0 in the respective columns which have been assigned. Similarly column wise, column C has got only one 0, so we assign that cell and cross out the remaining 0 values in the row corresponding to this assigned cell.

	A	B	C	D	E	
1	5	[0]	8	10	11	
2	0×	6	15	0×	3	
3	8	5	[0]	0×	0×	
4	[0]	6	4	2	7	
5	3	5	6	[0]	8	

Since there are cost matrix is of size 5 × 5, we require 5 assignments for the solution to be optimal. However, we have only got 4 assignments, so the solution is not optimal. Now, we should draw minimum number of horizontal and vertical lines such that it covers all zeros to revise the cost matrix for optimal solution.
Out of all five rows row 2 is not assigned, so we tick mark this row. In the marked row we have two cells having zero which belongs to column A and D, so we mark these columns. In the marked column A and D we have a cell which consists an assigned 0 for each column, so we mark the rows corresponding to these cells (row 4 and 5). Now we draw horizontal lines on

the rows which has not been marked, and draw vertical lines on the columns which has not been marked.

	A	B	C	D	E	
1	5	[0]	8	10	11	
2	0×	6	15	0×	3	✓
3	8	5	[0]	0×	0×	
4	[0]	6	4	2	7	✓
5	3	5	6	[0]	8	✓
	✓			✓		

For the revised cost matrix we require the value of k, which is the minimum value among the cells which remain uncovered by the lines. In this case, the value of k is 3. Then, we add the value of k to the cell values which lies in the intersection of the drawn lines and subtract the value of k from the cell values which remained uncovered by the lines. Now, the cost matrix has been revised.

	A	B	C	D	E	
1	8	0	8	13	11	
2	0	3	12	0	0	
3	11	5	0	3	0	
4	0	3	1	2	4	
5	3	2	3	0	5	

Again rows 1, 4, and 5 have only one cell having the value of 0 (1-B, 4-A,k and 5-D), so we assign these cells with square brackets and cross out the cells having the value of 0 in their respective columns. Similarly column wise column C has only one cell having the value of 0 (3-C), so we assign this cell with square bracket and cross out the cell having the value of 0 in its respective row.

	A	B	C	D	E	
1	8	[0]	8	13	11	
2	$0\times$	3	12	$0\times$	[0]	
3	11	5	[0]	3	$0\times$	
4	[0]	3	1	2	4	
5	3	2	3	[0]	5	

$\because$ the size of the matrix = the number of assignments = 5, the solution obtained is the optimal solution.
The optimum assignment is given below:

Job	Truck
1	B
2	E
3	C
4	A
5	D

Also, minimum time = 8 + 12 + 4 + 6 + 12 = 42.

6.8 Travelling Salesman Problem

The travelling salesman problem asks the following question: ”Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city exactly once and returns to the origin city?”.

Just consider how a postman delivers the post to the addressee. He arranges all the letters in an order and starts from the post office and goes from addressee to addressee and finally back to his post office. If he does not arrange the posts in an order he may have to travel a long distance to clear all the posts. Similarly, a traveling sales man has to plan his visits. Let us say, he starts from his head office and go round the branch offices and come back to his head office. While traveling he will not visit the branch already visited and he will not come back until he visits all the branches.

There are different types of traveling salesman's problems. One is ***cyclic*** problem. In this problem, he starts from his head quarters and after visiting all the branches, he will be back to his head quarters. The second one is ***Acyclic*** problem. In this case, the traveling salesman leaves his head quarters and after visiting the intermediate branches, finally reaches the last branch and stays there. The first type of the problem is solved by Hungarian method or Assignment technique. The second one is solved by Dynamic programming method.

Note: The traveling salesman problem, where we sequence the cities or branches he has to visit is a SEQUENCING PROBLEM. But the solution is solved using the Assignment technique. Hence basically, the traveling salesman problem is a SEQUENCING PROBLEM; the objective is to minimize the total distance traveled.

Numerical Example:
A salesman stationed at city A has to decide his tour plan to visit cities B, C, D, E and back to city A I the order of his choice so that total distance traveled is minimum. No sub touring is permitted. He cannot travel from city A to city A itself. The distance between cities in Kilometers is given below:

Cities	A	B	C	D	E
A	∞	16	18	13	20
B	21	∞	16	27	14
C	12	14	∞	15	21
D	11	18	19	∞	21
E	16	14	17	12	∞

Solution:
Performing row reduction we get the row opportunity cost matrix (ROCM).

	A	B	C	D	E	
A	∞	3	5	0	7	
B	7	∞	2	13	0	
C	0	2	∞	3	9	
D	0	7	8	∞	10	
E	4	2	5	0	∞	

Performing column reduction on the ROCM we get the total opportunity cost matrix (TOCM).

	A	B	C	D	E	
A	∞	1	3	0	7	
B	7	∞	0	13	0	
C	0	0	∞	3	9	
D	0	5	6	∞	10	
E	4	0	3	0	∞	

Now we assign using the same technique used in the Assignment Problem.

	A	B	C	D	E	
A	∞	1	3	[0]	7	
B	7	∞	[0]	13	0×	
C	0×	0×	∞	3	9	
D	[0]	5	6	∞	10	
E	4	[0]	3	0×	∞	

Since we there are only 4 assignments but the matrix is of size 5 × 5, we need to modify the opportunity cost matrix for optimum solution. Now, we will draw the minimum number of vertical and horizontal lines to cover all the 0's.

	A	B	C	D	E	
A	∞	1	3	[0]	7	✓
B	7	∞	[0]	13	0×	
C	0×	0×	∞	3	9	✓
D	[0]	5	6	∞	10	✓
E	4	[0]	3	0×	∞	✓
	✓	✓		✓		

For new opportunity cost matrix, k = min(3, 6, 7, 9, ∞) = 3. The new opportunity cost matrix is:

	A	B	C	D	E	
A	∞	1	0	0	4	
B	10	∞	0	16	0	
C	0	0	∞	3	6	
D	0	5	3	∞	7	
E	4	0	0	0×	∞	

Again we assign the cells for the solution.

	A	B	C	D	E	
A	∞	1	0	0	4	
B	10	∞	$0\times$	16	$[0]\times$	
C	$0\times$	[0]	∞	3	6	
D	[0]	5	3	∞	7	
E	4	$0\times$	0	0	∞	

Here we have a tie situation, that is, more than one solution can be obtained. At first we assign the cell A-D, we get:

	A	B	C	D	E	
A	∞	1	$0\times$	[0]	4	
B	10	∞	$0\times$	16	$[0]\times$	
C	$0\times$	[0]	∞	3	6	
D	[0]	5	3	∞	7	
E	4	$0\times$	[0]	$0\times$	∞	

We have got 5 assignments, however, the solution is not valid for TSP since there are two cycles, that is, (A $\rightarrow$ D, D $\rightarrow$ A) and (B $\rightarrow$ E, E $\rightarrow$ C, C $\rightarrow$ B).

Now, we assign the cell A-C, we get:

	A	B	C	D	E	
A	∞	1	[0]	$0\times$	4	
B	10	∞	$0\times$	16	$[0]\times$	
C	$0\times$	[0]	∞	3	6	
D	[0]	5	3	∞	7	
E	4	$0\times$	$0\times$	[0]	∞	

We have 5 assignments, and the solution is valid for TSP since it makes a one complete cycle travelling all cities starting from city A (A $\rightarrow$ C, C $\rightarrow$ B, B $\rightarrow$ E, E $\rightarrow$ D, D $\rightarrow$ A).
The minimum distance = 18 + 14 + 14 + 12 + 11 = 69 Km.

Numerical Example:
Given the travel costs below, show how to sequence the travel between the cities so as to minimize the total setup cost per cycle.

Cities	A	B	C	D	E
A	∞	2	5	7	1
B	6	∞	3	8	2
C	8	7	∞	4	7
D	12	4	6	∞	5
E	1	3	2	8	∞

Solution:
ROCM:

	A	B	C	D	E	
A	∞	1	4	6	0	
B	4	∞	1	6	0	
C	4	3	∞	0	3	
D	8	0	2	∞	1	
E	0	2	1	7	∞	

TOCM:

	A	B	C	D	E	
A	∞	1	3	6	0	
B	4	∞	0	6	0	
C	4	3	∞	0	3	
D	8	0	1	∞	1	
E	0	2	0	7	∞	

Now, we assign the cells for the solution.

	A	B	C	D	E	
A	∞	1	3	6	[0]	
B	4	∞	[0]	6	$0\times$	
C	4	3	∞	[0]	3	
D	8	[0]	1	∞	1	
E	[0]	2	$0\times$	7	∞	

We have got 5 assignments, however, the solution is not valid for TSP since there are two cycles, that is, (A $\rightarrow$ E, E $\rightarrow$ A) and (B $\rightarrow$ C, C $\rightarrow$ D, D $\rightarrow$ B).

The next best solution is obtained by assigning minimum non-zero cell in the matrix (In this case 1). At first, we have to assign all 1's and then the 0's. Also we have a tie between the cells D-C and D-E. At first, we assign the cell D-C.

	A	B	C	D	E	
A	∞	[1]	3	6	$0\times$	
B	4	∞	$0\times$	6	[0]	
C	4	3	∞	[0]	3	
D	8	$0\times$	[1]	∞	$1\times$	
E	[0]	2	$0\times$	7	∞	

We have got 5 assignments, however, the solution is not valid for TSP since there are two cycles, that is, (C $\rightarrow$ D, D $\rightarrow$ C) and (A $\rightarrow$ B, B $\rightarrow$ E, E $\rightarrow$ A).

Now, we assign the cell D-E.

	A	B	C	D	E	
A	∞	[1]	3	6	$0\times$	
B	4	∞	[0]	6	$0\times$	
C	4	3	∞	[0]	3	
D	8	$0\times$	$1\times$	∞	[1]	
E	[0]	2	$0\times$	7	∞	

We have 5 assignments, and the solution is valid for TSP since it makes a one complete cycle travelling all cities starting from city A (A $\rightarrow$ B, B $\rightarrow$ C, C $\rightarrow$ D, D $\rightarrow$ E, E $\rightarrow$ A).
minimum distance = 2 + 3 + 4 + 5 + 1 = 15

6.9 Review Problems

1. Describe the three main actors involved in freight transportation and distribution.

2. List out the factors that influences the decision of shippers regarding the choice of carrier and transportation mode.

3. Describe the two ways in which shipments can be made.

4. Describe the three distribution structures used in freight transport and distribution.

5. Write the differences among the North-West Corner Method, Least-Cost Cell Method, and Vogel's Approximation Method.

6. A dairy firm has three plants located in a state. The daily milk production at each plant is as follows: Plant 1 : 6 million litres, Plant 2 : 1 million litres, and Plant 3 : 10 million litres.

   Each day, the firm must fulfil the needs of its four distribution centres. The minimum requirement of each centre is as follows: Distribution centre 1 : 7 million litres, Distribution centre 2 : 5 million litres, Distribution centre 3 : 3 million litres, and Distribution centre 4 : 2 million litres.

   Cost (in hundreds of rupees) of shipping one million litre from each plant to each distribution centre is given in the following table:

Plant / Distribution Center	D_1	D_2	D_3	D_4
P_1	2	3	11	7
P_2	1	0	6	1
P_3	5	8	15	9

   Find the initial basic feasible solution for given problem by using following methods:

   (a) North-West Corner Method
   (b) Least-Cost Cell Method
   (c) Vogel's Approximation Method

7. A logistics department of a freight company has five trucks with five jobs to be performed. The time (in hours) that each truck takes to perform each job is given in the effectiveness matrix.

Job / Truck	1	2	3	4	5
A	10	5	13	15	16
B	3	9	18	13	6
C	10	7	2	2	2
D	7	11	9	7	12
E	7	9	10	4	12

How should the jobs be allocated, one per truck, so as to minimize the total truck-hours?

6.9.1 Answers

6 .

a $x_{11} = 6$, $x_{21} = 1$, $x_{32} = 5$, $x_{33} = 3$, $x_{34} = 2$, and Total Cost = Rs. 11600

b $x_{11} = 6$, $x_{22} = 1$, $x_{31} = 1$, $x_{32} = 4$, $x_{33} = 3$, $x_{34} = 2$, and Total Cost = Rs. 11200

c $x_{11} = 1$, $x_{12} = 5$, $x_{24} = 1$, $x_{31} = 6$, $x_{33} = 3$, $x_{34} = 1$, and Total Cost = Rs. 10200

7 .

Job	Truck	Time
A	2	5
B	1	3
C	5	2
D	3	9
E	4	4
		Total = 23

Chapter 7

Environmental Effects of Transportation, Fuel Consumption Models and Emission Standards

7.1 Introduction

Transportation activities, both passenger and freight, is crucial to economic activity, as it allows an increase in the flow of people and goods globally. Investment in transportation is fundamental for economic growth and development; the movement of people and goods from one place to another enables the exchange of goods and manpower, stimulating trade and commerce, and thus economic development. However, transport comes at a price having undesirable environmental impacts, which are generally referred to as *externalities*.

As the environmental impacts of transportation increase with the flow of transportation demand, the need to mitigate these is paramount. There are often several objectives that transportation strives to meet, which relate to economic, social and environmental aspects. However, it is not always possible to meet all the objectives simultaneously, which sometimes results in a compromise in one or more of the objectives, depending on the trade-offs.

This chapter discusses about the various externalities resulted due to extensive transportation activities. Most of these externalities are resulted due to vehicular emissions, therefore we will discuss about the fuel consumption models to quantify the fuel consumption contributed by vehicles. Lastly, this chapter also discusses about the various emission standards incorporated globally.

7.2 Environmental Problems due to Transportation

7.2.1 Vehicular Emissions and Air Pollution

Being arguably the most prominent of externalities, emissions are a consequence of use of fuel within the transport sector, such as liquid fuel, most of which is produced from petroleum. Gasoline and diesel are two of the most widely used types of liquid fuels, both of which are composed of many different types of hydrogen and carbon compounds, known as hydrocarbons, in which hydrogen and carbon atoms are chemically bound. Hydrocarbons include, for example, methane (CH_4) and ethane (C_2H_6).

Engines used in a majority of transportation vehicles, and particularly in road transport, produce energy by mixing hydrocarbons with air. Volumewise, a large fraction of air contains oxygen (O_2) and nitrogen (N_2). The process by which energy is produced by engines is known as internal combustion, which, under ideal conditions, should only result in carbon dioxide (CO_2) and water vapour (H_2O). However, the process is incomplete, meaning that not all of the fuel is consumed during internal combustion. As a result, a number of by-products are generated, such as carbon monoxide (CO) as a result of insufficient oxygen, nitrogen oxide (NO_x) as a result of nitrogen reacting with oxygen and particulate matter (PM) which are small particles of dust, soot and organic matter suspended in the atmosphere. In addition, small amounts of methane (CH_4) and nitrous oxide (N_2O) are also produced during the combustion process. Furthermore, when sunlight reacts with air that contains hydrocarbons and NOx, it produces another gas called ozone (O_3). In maritime transportation, ships running on diesel engines burn fuel that contains high amounts of sulphur content, as a result of which sulphur oxide as well as nitrogen oxide emissions (also known as NO_x and SO_x) are significant. The existence of these gases in the atmosphere gives rise to air pollution.

Some of the environmental impacts of vehicular emissions and air pollution are discussed below.

Particulate Matter (PM)

Particulate matter is the general term for the mixture of solid particles and liquid droplets found in the air. Particulate matter includes dust, dirt, soot, smoke and liquid droplets. It can be emitted into the air from natural and man made sources, such as windblown dust, motor vehicles, construction factories, and fires. Particles are also formed in the atmosphere by condensation or the transformation of emitted gases such as sulphur dioxide, nitrogen oxides, and volatile organic compounds. sites, factories, and fires. Particles are also formed in the atmosphere by condensation or the transformation of emitted gases such as sulphur dioxide, nitrogen oxides, and volatile organic compounds.

Scientific studies show a link between particulate matter (alone or in combination with other pollutants in the air) and a series of health effects. Studies of human populations and laboratory studies of animals and humans have established linkages to major human health impacts including respiratory symptoms; aggravation of existing respiratory and cardiovascular disease; alterations in the body's defence systems against foreign materials; damage to lung tissue; carcinogenesis, and premature mortality.

PM also causes damage to materials and soiling; it is a major cause of substantial visibility impairment in many parts of the world.

Motor vehicle particle emissions and the particles formed by the transformation of motor vehicle gaseous emissions tend to be in the fine particle range. Fine particles (those less than 2.5 micrometers in diameter) are of health concern because they easily reach the deepest recesses of the lungs. Scientific studies have linked fine particles (alone or in combination with other air pollutants), with a series of significant health problems, including premature death; respiratory related hospital admissions and emergency room visits; aggravated asthma; acute respiratory symptoms, including aggravated coughing and difficult or painful breathing; chronic bronchitis; and decreased lung function that can be experienced as shortness of breath.

The US Environment Protection Agency (EPA) recently tightened the air quality standards for particulates. The standard for coarse particles (10 microns or less) remains essentially unchanged, while a new standard for fine particles (2.5 microns or less) was set at an annual limit of 15 micrograms per cubic meter and a 24-hour limit of 65 micrograms per cubic meter.

The California Air Resources Board (CARB) has evaluated diesel exhaust as a candidate toxic air contaminant under the State's air toxics identification programme. To evaluate whether or not diesel exhaust causes cancer, the Office of Environmental Health Hazard Assessment (OEHHA) reviewed all controlled animal and mutagenicity studies as well as studies of worker populations exposed to diesel exhaust. They analysed over 30 human studies concerning lung cancer risk and workplace exposure to diesel exhaust. They found that workers who were exposed to diesel exhaust were consistently more likely than others to develop lung cancer. The consistent results are unlikely to be due to chance, confounding, or bias, according to CARB.

As a result, CARB concluded that a reasonable and likely explanation for the increased rates of lung cancer observed in the epidemiological studies is a causal association between diesel exhaust exposure and lung cancer.

Carbon Monoxide (CO)

Carbon monoxide (CO) is a tasteless, odourless, and colourless gas produced though the incomplete combustion of carbon-based fuels. CO enters the bloodstream through the lungs and reduces the delivery of oxygen to the body's organs and tissues. The health threat from exposure to low concentrations of CO is most serious for those who suffer from cardiovascular disease, particularly those with angina or peripheral vascular disease. Healthy individuals also are affected, but only at higher exposure levels. Exposure to elevated CO levels is associated with impairment of visual perception, work capacity, manual dexterity, learning ability and performance of complex tasks.

Nitrogen Oxides (NO_x)

NO_x emissions produce a wide variety of health and welfare effects. Nitrogen dioxide can irritate the lungs and lower resistance to respiratory infection (such as influenza). NO_x emissions are an important precursor to acid rain that may affect both terrestrial and aquatic ecosystems. Atmospheric deposition of nitrogen leads to excess nutrient enrichment problems ("eutrophication") a problem, for example, in the Chesapeake Bay and several other nationally important estuaries along the East and Gulf Coasts of the USA. Eutrophication can produce multiple adverse effects on water quality and the aquatic environment, including increased nuisance and toxic algal blooms, excessive phytoplankton growth, low or no dissolved oxygen in bottom waters, and reduced sunlight causing losses in submerged aquatic vegetation critical for healthy estuarine ecosystems. Nitrogen dioxide and airborne nitrate also contribute to pollutant haze, which impairs visibility and can reduce residential property values and revenues from tourism.

Photo-chemical Oxidants (Ozone)

Ozone is not emitted directly into the atmosphere, but is formed by a reaction of volatile organic compounds (VOC) — for vehicles mainly hydrocarbons (HC) — and NO_x in the presence of heat and sunlight. Ground-level ozone forms readily in the atmosphere, usually during hot summer weather. VOCs are emitted from a variety of sources, including motor vehicles, chemical plants, refineries, factories, consumer and commercial products, and other industrial sources. VOCs are also emitted by natural sources such as vegetation. NO_x is emitted from motor vehicles, power plants and other source of combustion. Changing weather patterns contribute to yearly differences in ozone concentrations and differences from city to city. Ozone can also be transported into an area from pollution sources found hundreds of miles upwind.

Ground-level ozone is the prime ingredient of smog, the pollution that blankets many areas during the summer4. Short-term exposures (1-3 hours) to high ambient ozone concentrations have been linked to increased hospital admissions and emergency room visits for respiratory problems. Repeated exposures to ozone can exacerbate symptoms and the frequency of episodes for people with respiratory diseases such as asthma. Other health effects attributed to short-term exposures include significant decreases in lung function and increased respiratory symptoms such as chest pain and cough. These effects are generally associated with moderate or heavy exercise or exertion. Those most at risk include children who are active outdoors during the summer, outdoor workers, and people with pre-existing respiratory diseases like asthma. In addition, long-term exposures to ozone may cause irreversible changes in the lungs, which can lead to chronic ageing of the lungs or chronic respiratory disease.

Ambient ozone also affects crop yield, forest growth, and the durability of materials. Because ground-level ozone interferes with the ability of a plant to produce and store food, plants become more susceptible to disease, insect attack, harsh weather and other environmental stresses. Ozone chemically attacks elastomers (natural rubber and certain synthetic polymers), textile fibres and dyes, and, to a lesser extent, paints. For example, elastomers become brittle and crack, and dyes fade after exposure to ozone.

Ozone is also an effective greenhouse gas, both in the stratosphere and the troposphere5. That is, ozone absorbs infrared radiation emitting from the earth, captures it before it escapes into space, and re-emits a portion of it back toward the earth's surface. The specific role of ozone in climate change is very complex and not yet well understood. Ozone concentrations in the atmosphere vary spatially, both regionally and vertically, and are most significant in urban areas where precursor gases are abundant. This variability makes assessment of global, long-term trends difficult.

Lead

Over the past century, a range of clinical, epidemiological and toxicological studies have defined the nature of lead toxicity and investigated its mechanisms of action, identifying young children as a critically susceptible population. As noted by the 1995 Environmental Health Criteria Document for Lead, published by the International Programme on Chemical

Safety (IPCS), lead affects many organs and organ systems in the human body with sub-cellular changes and neuro-developmental effects appearing to be the most sensitive. The most substantial evidence from cross sectional and prospective studies of populations with lead levels generally below 25 µg/decilitre of blood relates to reductions in intelligence quotient (IQ).

As noted by the IPCS, existing epidemiological studies do not provide definitive evidence of a threshold. Below the range of about 10-15 µg /decilitre of blood, the effects of confounding variables and limits in the precision in analytical and psychometric measurements increase the uncertainty attached to any estimate of effect. However, there is some evidence of an association below this range. Animal studies provide support for a causal relationship between lead and nervous system effects, reporting deficits in cognitive functions at lead levels as low as 11-15 µg/decilitre of blood which can persist well beyond the termination of lead exposure. Other effects, which may occur, include:

- impaired sensory motor function;
- impaired renal function;
- a small increase in blood pressure has been associated with lead exposure;
- some but not all epidemiological studies show a dose dependent association of pre-term delivery and some indices of foetal growth and maturation at of 15 µg/decilitre or more.

Lead and its compounds may enter the environment at any point during mining, smelting, processing, use, recycling or disposal. In countries where leaded gasoline is still used, the major air emission is from mobile and stationary combustion of gasoline. Areas in the vicinity of lead mines and smelters are subject to high levels of air emissions.

Airborne lead can be deposited on soil and water, thus reaching humans through the food chain and in drinking water. Atmospheric lead is also a major source of lead in household dust.

Because of the concerns highlighted above, a global consensus has emerged to phase out the use of lead in gasoline.

In December 1994, at the Summit of the Americas, heads of state from a number of countries pledged to develop national action plans for the phase out of leaded gasoline in the Western Hemisphere.

In May 1996, the World Bank called for a global phase out of leaded gasoline and offered to help countries design feasible phase out schedules and incentive frameworks.

A key recommendation of the Third "Environment for Europe" Ministerial Conference held in Sofia, Bulgaria in October 1995 called for the reduction and ultimate phase out of lead in gasoline.

In June 1996, the second United Nations Conference on Human Settlements, called Habitat II, included the elimination of lead from gasoline as a goal in its agenda.

In May 1997, environmental ministers from the Group of Seven plus Russia endorsed the phase out of leaded gasoline in the 1997 Declaration of Environmental Leaders of the Eight on Children's Environmental Health.

Climate Change

Beyond direct adverse health effects, there are other concerns with vehicle emissions. Among these is global warming or the greenhouse effect. Greenhouse warming occurs when certain gases allow sunlight to penetrate to the earth but partially trap the planet's radiated infrared heat in the atmosphere. Some such warming is natural and necessary. If there were no water vapour, carbon dioxide, methane, and other infrared absorbing (greenhouse) gases in the atmosphere trapping the earth's radiant heat, our planet would be about 60°F (33°C) colder, and life as we know it would not be possible. Naturally occurring greenhouse gases include water vapour, carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), and ozone (O_3).

Several classes of halogenated substances that contain fluorine, chlorine, or bromine are also greenhouse gases, but they are, for the most part, solely a product of industrial activities. Chlorofluorocarbons (CFCs) and hydrochlorofluorocarbons (HCFCs) are halocarbons that contain chlorine, while halocarbons that contain bromine are referred to as halons. Other fluorine containing halogenated substances include hydrofluorocarbons (HFCs), perfluorocarbons (PFCs), and sulphur hexafluoride (SF_6).

There are also several gases that, although they do not have a direct global warming effect, do influence the formation and destruction of ozone, which does have such a terrestrial radiation absorbing effect. These gases include carbon monoxide (CO), oxides of nitrogen (NO_X), and non-methane volatile organic compounds (NMVOCs).

Aerosols, extremely small particles or liquid droplets often produced by emissions of sulphur dioxide (SO_2), can also affect the absorptive characteristics of the atmosphere.

Although CO2, CH4, and N2O occur naturally in the atmosphere, the atmospheric concentration of each of them has risen, largely as a result of human activities. Since 1800, atmospheric concentrations of these greenhouse gases have increased by 30, 145, and 15%, respectively (IPCC 1996). This build-up has altered the composition of the earth's atmosphere, and may affect the global climate system.

Beginning in the 1950s, the use of CFCs and other ozone depleting substances (ODSs) increased by nearly 10% a year, until the mid-1980s when international concern about ozone depletion led to the signing of the Montreal Protocol. Since then, the consumption of ODSs has rapidly declined, as they are phased-out. In contrast, use of ODS substitutes such as HFCs, PFCs, and SF_6 has grown significantly and all have strong greenhouse forcing effects.

7.2.2 Noise Pollution

Transport gives rise to noise pollution, primarily in urban areas, with undesirable effects on human health. These effects can range from short-term annoyance, sleep disturbance and impaired cognitive functioning to significant long-term physiological problems such as cardiovascular diseases, loss of hearing and mental health problems.

7.2.3 Land and resource consumption

Natural resources are depleted for use in the form of fuel, as well as by extraction of the materials required for building or construction. In addition, the infrastructure itself necessary for transportation activities (roads or highways, railway tracks, ports or terminals, storage or distribution facilities) damages land and local ecosystems.

7.3 Mitigating measures

Various approaches have been suggested to reduce the environmental impacts of transportation, some of which are described here:

- ***Technological improvements*** are aimed at developing new technology to improve the environmental performance of vehicles, including improved design of vehicles that allow for additional carrying capacity without compromising the fuel or energy efficiency, new vehicles that run on alternative sources of energy (such as electric vehicles or alternative fuels such as bio-diesel) and new design or redesign of engines that have better energy efficiency.

- ***Strategic or tactical approaches*** include changing the so-called modal split of freight and transferring freight from energy-intensive modes of transport, such as road, to more environmentally friendly modes of transport, for example rail, coastal shipping, waterways, or use of any of these in combination with road transport. Tactical approaches also include fleet management involving the selection of the right vehicles to purchase and maintain, as well as regular vehicle maintenance to ensure that they operate at optimum efficiency. In road transportation, for example underinflated tires, fuel leaks or poor combustion can lead to loss in fuel efficiency of a vehicle.

- ***Operational strategies*** include reducing the actual numbers of vehicles running, vehicle kilometers and tonne kilometers by increasing load factors (reducing empty or partly loaded running of lorries), utilizing new information technology to improving the routing and scheduling of vehicles, consolidating deliveries, sharing loads and pick up deliveries with other companies and improving driver training and behavior.

It is important that these approaches are implemented to produce win-win solutions, in which the environmental benefits should be at least as much as the gains reaped by the organizations or companies who adopt these practices.

Kyoto Protocol

In late November 1995, the IPCC Working Group 1 concluded, "the balance of evidence suggests that there is a discernible human influence on global climate." In December 1997, acting on this consensus, countries around the world approved the Kyoto Protocol to the 1992 Climate Change Treaty. Key aspects of the agreement include:

Reductions: Thirty-eight industrialized nations are required to reduce their "greenhouse" gas emissions from 1990 levels between 2008 and 2012. The European Union would reduce them by 8%, the United States by 7% and Japan by 6%. Some would face smaller reductions, and a few would not face any now. As a group, the industrialized nations would cut back on the emissions of such gases by just more than 5%.

Gases Involved: Emissions of six gases would be affected: carbon dioxide, methane, nitrous oxide, and three halo-carbons used as substitutes for ozone-damaging chlorofluorocarbons.

'Offshore' Reductions: Countries that do not meet their own emission targets can strike deals with nations that do better than required, to buy the excess "quota". This may encourage reductions to be made where most cost-effective.

Enforecement: A later meeting of the treaty parties will decide on "appropriate and effective" ways to deal with non-compliance.

Developing Countries: Developing countries, including major greenhouse gas emitters such as China and India, are asked to set voluntary reduction targets.

Next Step: The accord approved by the Kyoto conference takes effect once it is ratified by 55 nations, representing 55% of 1990 carbon dioxide emissions. It is binding on individual countries only after their governments' complete ratification.

Implementing this agreement requires significant improvements in fuel economy and carbon dioxide emissions from vehicles.

7.4 Fuel Consumption Models

Fuel consumption models are mathematical relationships giving fuel consumption as a function a function of other measurable or calculable quantities that are usually aggregated over the portion of the transportation system being analyzed. The quantities measured or calculated may be vehicle-miles traveled, number of stops per unit of time, total vehicular delay per unit of time, average speed, and so on. Two forms that have found the most use in practice to date are drive-mode elemental models and average-speed models.

7.4.1 Drive-Mode Elemental Models

As the name implies, drive-mode elemental models are made up of the elements that contribute to fuel consumption while driving (i.e., fuel used in cruising, idling, and accelerating). The basic assumptions in an elemental model are that the elements are independent and their sum equals the total fuel consumed.

The simplest form of drive-mode elemental model is:

$$G = f_1 L + f_2 D + f_3 S \tag{7.1}$$

Where,

G = fuel consumed per vehicle over a measured distance (total section distance)

L = total section distance traveled

D = stopped delay per vehicle (i.e., time spent in idling)
S = number of stops
f_1 = fuel consumption rate per unit distance traveled while cruising
f_2 = fuel consumption rate per unit time while idling
f_3 = excess fuel used in decelerating to stop and accelerating back to cruise speed

7.4.2 Average-Speed Models

The model relating fuel consumption to trip time or its inverse, average speed is:

$$F = k_1 + k_2 T = k_1 + \frac{k_2}{v} \tag{7.2}$$

$$for\ 10 \leq v \leq 56\ kmph$$

Where,
F = fuel consumed per vehicle per unit distance (e.g., liters/km or gal/mi)
T = travel time per unit distance, including stops and speed changes (e.g., min/km or min/mi)
v = average speed measured over a distance, including stops and speed changes
k_1 = parameter associated with fuel consumed to overcome rolling resistance, approximately proportional to vehicle weight (gal/veh-mi)
k_2 = parameter approximately proportional to fuel consumption while idling (gal/hr)

This form of the average speed model is not valid at speeds higher than 56 km/hr (35 mph) because at higher speeds the effects of air resistance become increasingly stronger. For speeds lower than 15 km/hr (9 mph), fuel consumption increases rapidly and a better fit to the data is:

$$F = F_R + \frac{800 F_0}{v(v+8)} \tag{7.3}$$

Where,

F_R = fuel consumed while in motion (liters/100 km)
F_0) = fuel consumed during stopped time while idling (liters/hr)
v = average speed (km/hr)

The validity of the average-speed model was extended to suburban areas for speeds up to 88 km/hr (55 mph) by the introduction of a term involving v^2.

$$F = k_1 + \frac{k_2}{v} + k_3 v^2 \tag{7.4}$$

The parameters k_1 and k_2 have been estimated from regression analysis of fuel consumption data in several countries. Data on 37 U.S. passenger cars on 1976 model year were found to be closely correlated with vehicle weight and fuel rate at idle, giving the relationships:

$$\because k_1 = 9.5 \times 10^{-6}\, W\, (gal/mi) \tag{7.5}$$
$$= 49.57 \times 10^{-6} Wm\, (liter/km) \tag{7.6}$$
$$\because k_2 = 0.998I = I \times f_2 \tag{7.7}$$

Where,

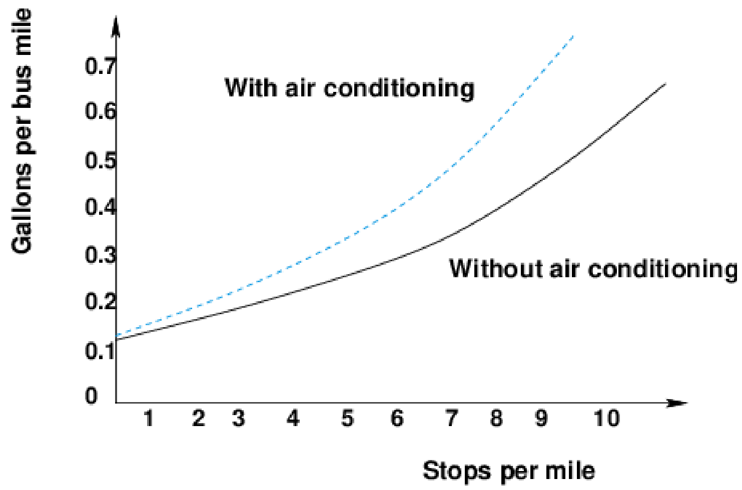
W = vehicle weight (lb)
Wm = vehicle weight (kg)
I = fuel consumption rate while idling

Depending on vehicle weight, engine size, and technological features, the -values of k_1 and k_2 can vary widely. Data from field tests (1973-1976) on nine passenger cars ranging in weight from 2270 to 5474 lb (1035 to 2488 kg) gave values of k_1 ranging from 0.0194 to 0.0518 gal/mi (0.0456 to 0.1218 liter/km) and values of k_2 ranging from 0.418 to 1.031 gal/hr (For use in metric version of the model this would be equivalent to k_2 ranging from 0.983 to 2.42 in order to yield 1/km). The average values were incorporated into a version of model published by FHWA:

$$F = 0.0362 + \frac{0.746}{v}(gal/veh - mi) \tag{7.8}$$

By using the relationship $k_1 = 9.58 \times 10^{-6}\, W$, the value of $k_1 = 0.0362$ represents a vehicle weighing 3800 lb (1718 kg).

Buses operating in an urban environment may present a unique case because of their schedule of frequent stops. In this case, bus fuel consumption may be calculated by using graphs such as the graph given below calibrated with average number of stops per unit distance rather than average speed. The estimate for stops should include both the scheduled stops and the stops resulting from traffic congestion and traffic controls.



7.5 Emissions Control Programmes

7.5.1 European Union

The European Commission set up the European Auto-Oil Programme in 1994 with the European associations for the car industry (ACEA) and the oil industry (Europia), to provide the technical basis for the development of EU policy towards road vehicle emissions. The first phase of the programme was completed in 1996 and following two years of refinement, a three-part dossier to further reduce pollution from road transport in the Community was agreed, in 1998, by the European Parliament and the Council of Environment Ministers. The three dossiers cover:

- measures to be taken against air pollution by emissions from passenger cars;
- measures to be taken against air pollution by emissions from light commercial vehicles (pick up trucks, delivery vans etc.);
- the quality of petrol and diesel fuels.

Overall the Auto Oil programme has five sections:

- fuel quality;
- emissions from private cars;
- emissions from light commercial vehicles;
- emissions from heavy goods vehicles;
- adaptation of provisions relating to roadworthiness testing.

More specifically, the directives aim at:

- Controlling those parameters in the composition of petrol and diesel that influence the level of atmospheric emissions, in particular sulphur, benzene, aromatics and lead;
- Reducing the limit values for certain pollutants in new vehicle models being brought onto the market;
- In addition, improved control of the day-to-day emissions of vehicles in use is to be achieved by the mandatory fitting of on-board diagnostic systems, the introduction of a new testing procedure and a new test to limit evaporative emissions.

EU has been regulating these standards gradually imposing more stringent standards to eventually reduce emissions. The following is a summary list of the standards, and when they come into force:

- Euro 1 (1992)
- Euro 2 (1996)
- Euro 3 (2000)
- Euro 4 (2005)
- Euro 5 (2009)
- Euro 6 (2014)

7.5.2 India

Bharat stage emission standards (BSES) are emission standards instituted by the Government of India to regulate the output of air pollutants from internal combustion engines and Spark-ignition engines equipment, including motor vehicles. The standards and the timeline for implementation are set by the Central Pollution Control Board under the Ministry of Environment & Forests and climate change. These standards, based on European regulations, were first introduced in 2000. n 2014, Saumitra Chaudhary committee gave recommendations on Auto Fuel Vision Policy 2025 which had recommended implementation of BS-IV (2017), BS-V (2019) and BS-VI (2024) standards. In 2016, the Indian government announced that the country would skip the BS-V norms altogether and adopt BS-VI norms by 2020. While the norms help in bringing down pollution levels, it invariably results in increased vehicle cost due to the improved technology & higher fuel prices. Currently, BS IV norms have been enforced across the country since April 2017. However, recently the Supreme Court of India ordered barring of sale of Bharat Stage IV vehicles from April 1, 2020.

Chapter 8

Accident Studies

8.1 Introduction

The problem of accident is a very acute in highway transportation due to complex flow pattern of vehicular traffic, presence of mixed traffic along with pedestrians. Traffic accident leads to loss of life and property. Thus the traffic engineers have to undertake a big responsibility of providing safe traffic movements to the road users and ensure their safety. Road accidents cannot be totally prevented but by suitable traffic engineering and management the accident rate can be reduced to a certain extent. For this reason systematic study of traffic accidents are required to be carried out. Proper investigation of the cause of accident will help to propose preventive measures in terms of design and control.

8.2 Objectives of Accident Studies

Some objectives of accident studies are listed below:

- To study the causes of accidents and suggest corrective measures at potential location;
- To evaluate existing design;
- To compute the financial losses incurred;
- To support the proposed design and provide economic justification to the improvement suggested by the traffic engineer;
- To carry out before and after studies and to demonstrate the improvement in the problem.

8.3 Causes of Road Accidents

The various causes of road accidents are:

1. ***Road Users*** - Excessive speed and rash driving, violation of traffic rules, failure to perceive traffic situation or sign or signal in adequate time, carelessness, fatigue, alcohol,sleep etc.

2. ***Vehicle*** - Defects such as failure of brakes, steering system, tyre burst,lighting system.

3. ***Road Condition*** - Skidding road surface, pot holes, ruts.

4. ***Road design*** - Defective geometric design like inadequate sight distance, inadequate width of shoulders, improper curve design, improper traffic control devices and improper lighting.

5. ***Environmental factors*** -unfavorable weather conditions like mist, snow, smoke and heavy rainfall which restrict normal visibility and and makes driving unsafe.

6. ***Other causes*** -improper location of advertisement boards, gate of level crossing not closed when required etc..

8.4 Accident Analysis

8.4.1 Accident Data Collection

The accident data collection is the first step in the accident study. The data collection of the accidents is primarily done by the police. Motorist accident reports are secondary data which are filed by motorists themselves. The data to be collected should comprise all of these parameters:

- ***General*** - Date, time, person involved in accident, classification of accident like fatal, serious, minor e.t.c.

- ***Location*** - Description and detail of location of accident.

- ***Details of vehicle involved*** - Registration number, description of vehicle, loading detail, vehicular defects e.t.c.

- ***Nature of accident*** - Details of collision, damages, injury and casualty e.t.c.

- ***Road and traffic condition*** - Details of road geometry, surface characteristics,type of traffic, traffic density e.t.c.

- ***Primary causes of accident*** - Details of various possible cases (already mentioned) which are the main causes of accident.

- ***Accident cost*** - Financial losses incurred due to property damage, personal injury and casualty.

These data collected need proper storing and retrieving for the following purpose. The purposes are as follows:

1. Identification of location of points at which unusually high number of accident occur.

2. Detailed functional evaluation of critical accident location to identify the causes of accidents.

3. Development of procedure that allows identification of hazards before large number of accidents occurs.

4. Development of different statistical measures of various accident related factors to give insight into general trends, common casual factors, driver profiles, e.t.c.

8.4.2 Accident Investigation

The accident data collection involves extensive investigation which involves the following procedure:

1. ***Reporting***: It involves basic data collection in form of two methods:

    (a) ***Motorist accident report*** - It is filed by the involved motorist involved in all accidents fatal or injurious.

    (b) ***Police accident report*** - It is filed by the attendant police officer for all accidents at which an officer is present. This generally includes fatal accidents or mostly accidents involving serious injury required emergency or hospital treatment or which have incurred heavy property damage.

2. ***At Scene-Investigation***: It involves obtaining information at scene such as measurement of skid marks, examination of damage of vehicles, photograph of final position of vehicles, examination of condition and functioning of traffic control devices and other road equipment.

3. ***Technical Preparation***: This data collection step is needed for organization and interpretation of the study made. In this step measurement of grades, sight distance, preparing drawing of after accident situation, determination of critical and design speed for curves is done.

4. ***Professional Reconstruction***: In this step effort is made to determine from whatever data is available how the accident occurs from the available data. This involves accident reconstruction which has been discussed in the upcoming section in detail. It is professionally referred as determining behavioral or mediate causes of accident.

5. ***Cause Analysis***: It is the effort made to determine why the accident occurred from the data available and the analysis of accident reconstruction studies.

8.4.3 Accident Data Analysis

The purpose is to find the possible causes of accident related to driver, vehicle, and roadway. Accident analyses are made to develop information such as:

1. ***Driver and Pedestrian*** - Accident occurrence by age groups and relationships of accidents to physical capacities and to psychological test results.

2. ***Vehicle*** - Accident occurrence related to characteristic of vehicle, severity, location and extent of damage related to vehicles.

3. ***Roadway conditions*** - Relationships of accident occurrence and severity to characteristics of the roadway and roadway condition and relative values of changes related to roadways.

It is important to compute accident rate which reflect accident involvement by type of highway. These rates provide a means of comparing the relative safety of different highway and street system and traffic controls. Another is accident involvement by the type of drivers and vehicles associated with accidents.

Accident Rate per Kilometer:

On this basis the total accident hazard is expressed as the number of accidents of all types per km of each highway and street classification.

$$R = \frac{A}{L} \tag{8.1}$$

Where, R = total accident rate per km for one year, A = total number of accident occurring in one year, L = length of control section in km.

Accident Involvement Rate:

It is expressed as numbers of drivers of vehicles with certain characteristics who were involved in accidents per 100 million vehicle-km of travel.

$$R = \frac{N \times 10^8}{V} \tag{8.2}$$

Where, R = accident involvement per 100 million vehicle-km of travel, N = total number of drivers of vehicles involved in accidents during the period of investigation and V = vehicle-km of travel on road section during the period of investigation.

Death Rate Based on Population:

The traffic hazard to life in a community is expressed as the number of traffic fatalities per 100,000 populations. This rate reflects the accident exposure for entire area.

$$R = \frac{B \times 10^5}{P} \tag{8.3}$$

Where, R = death rate per 100,000 population, B = total number of traffic death in one year and P = population of area.

Death Rate Based on Registration:

The traffic hazard to life in a community can also be expressed as the number of traffic fatalities per 10,000 vehicles registered. This rate reflects the accident exposure for entire area and is similar to death rate based on population.

$$R = \frac{B \times 10^4}{M} \tag{8.4}$$

Where, R = death rate per 10,000 vehicles registered, B = total number of traffic death in one year and M = number of motor vehicles registered in the area.

Accident Based on Vehicle-Km of Travel:

The accident hazard is expressed as the number of accidents per 100 million vehicle km of travel. The true exposure to accident is nearly approximated by the miles of travel of the motor vehicle than the population or registration.

$$R = \frac{C \times 10^8}{V} \tag{8.5}$$

Where, R = accident rate per 100 million vehicle km of travel, C = number of total accidents in one year and V = vehicle km of travel in one year.

Numerical Example: The Motor vehicle consumption in a city is 5.082 million liters, there were 3114 motor vehicle fatalities, 355,799 motor vehicle injuries, 6,721,049 motor vehicle registrations and an estimated population of 18,190,238. Kilometer of travel per liter of fuel is 12.42 km/liter. Calculate registration death rate, population death rate and accident rate per vehicle km.
Solution:

Approximate vehicle km of travel = Total consumption o fuel × kilometer of travel per liter of fuel
$\therefore$ Approximate vehicle km of travel $= 5.08 \times 10^9 \times 12.42 = 63.1 \times 10^9$ Km

now,
Registration Death Rate (R) = $\frac{B\times10^4}{M}$ = $\frac{3114\times10^4}{6.72\times10^6}$ = 4.63

also,
Population Death Rate (R) = $\frac{B\times10^5}{P}$ = $\frac{3144\times10^5}{18.2\times10^6}$ = 17.1

also, (total no. of accidents = no. of vehicle fatalities)
Accident Rate per Vehicle (R) = $\frac{C\times10^8}{V}$ = $\frac{3114\times10^8}{63.1\times10^9}$ = 4.93

8.5 Accident Reconstruction

Accident reconstruction deals with representing the accidents occurred in schematic diagram to determine the pre-collision speed which helps in regulating or enforcing rules to control or check movement of vehicles on road at high speed. The following data are required to determine the pre-collision speed:

1. Mass of the vehicle

2. Velocities after collision

3. Path of each vehicle as it approaches collision point

The collision may be of two types collinear impact or angular collision.

8.5.1 Collinear Impact

A collision is said to be a collinear impact/ collision if the colliding bodies are traveling along the same line of action. Collinear impact can be again divided into two types:

1. Rear end collision

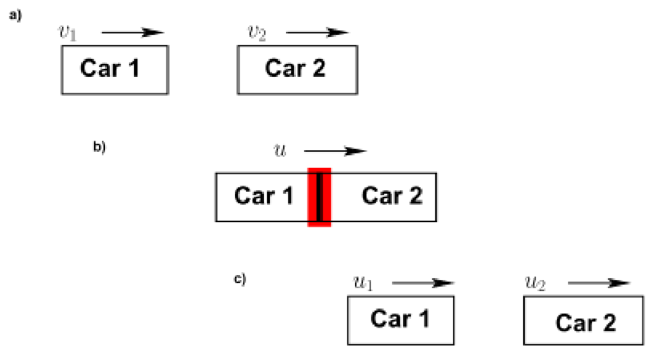
2. Head-on collision

Collinear impact can be defined by two theories, they are:

1. Poisson Impact Theory

2. Energy Theory

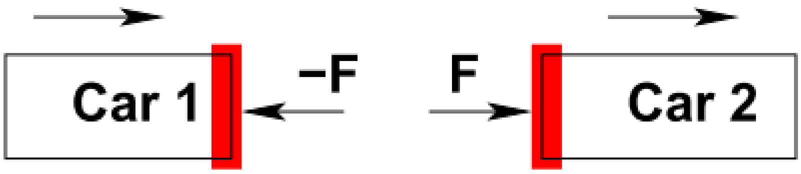
Poisson Impact Theory



Poisson impact theory, divides the impact in two parts - compression and restitution. The Figure above shows two vehicles traveling at an initial speed of v_1 and v_2 collide and obtain a uniform speed say u at the compression stage. And after the compression stage is over the final speed is u_1 and u_2. The compression phase is cited by the deformation of the cars. From the Newtons law $F = ma$,

$$m_1 \frac{dv_1}{dt} \; and \; m_2 \frac{dv_2}{dt} \tag{8.6}$$

where, m_1 and m_2 are the masses of the cars and F is the contact force. We know that every reaction has equal and opposite action. So as the rear vehicle pushes the vehicle ahead with force F. The vehicle ahead will also push the rear vehicle with same magnitude of force but has different direction. The action force is represented by F, whereas the reaction force is represented by $-F$ as shown in the figure below. In the compression phase cars are deformed. The compression phase terminates when the cars have equal velocity. Thus the cars obtain equal velocity which generates the following equation:



$$m_1(u - v_1) = -P_c; \;\; m_2(u - v_2) = P_c \tag{8.7}$$

Where, $P_c \equiv \int_0^{\tau_c} F dt$ which is the compression impulse and τ_c is the compression time. Thus, the velocity after collision is obtained as:

$$u = \frac{m_1 v_1 + m_2 v_2}{m_1 + m_2} \tag{8.8}$$

The compression impulse is given by:

$$P_c = \frac{m_1 m_2}{m_1 + m_2}(v_1 - v_2) \tag{8.9}$$

In the restitution phase the elastic part of internal energy is released,

$$m_1(u_1 - u) = -P_r \tag{8.10}$$
$$m_2(u_2 - u) = P_r \tag{8.11}$$

Where, $P_r \equiv \int_0^{\tau_r} F dt$ is the restitution impulse and τ_r is the restitution time. According to Poisson's hypothesis restitution impulse is proportional to compression impulse.

$$P_r = e\, P_c \tag{8.12}$$

Restitution impulse e is given by:

$$e = \frac{u_2 - u_1}{v_1 - v_2} \tag{8.13}$$

The total impulse is $P = P_1 + P_2$,

$$P = (1 + e)\frac{m_1 m_2}{m_1 + m_2}\Delta v \tag{8.14}$$

The post impact velocities are given by:

$$u_1 = u - e\frac{m_2}{m_1 + m_2}\Delta v \;=\; v_1 - (1 + e)\frac{m_2}{m_1 + m_2}\Delta v \tag{8.15}$$
$$u_2 = u + e\frac{m_1}{m_1 + m_2}\Delta v \;=\; v_2 + (1 + e)\frac{m_1}{m_1 + m_2}\Delta v \tag{8.16}$$

Where. $\Delta v = v_1 - v_2$. But we are required to determine the pre-collision speed according to which the safety on the road can be designed. So we will determine v_1 and v_2 from the given value of u_1 and u_2 .

Numerical Example:
Two vehicles traveling in the same lane have masses 3000 kg (following) and 2500 kg (leading). Also the velocity of the vehicles after collision is 25 kmph and 56 kmph respectively. The coefficient of restitution of the two vehicle system is assumed to be 0.6. Determine the pre-collision speed of the two vehicles.
Solution: mass of the vehicle 1 (m_1) = 3000 Kg
mass of the vehicle 2 (m_2) = 2500 Kg
final speed of the vehicle 1 (u_1) = 25 kmph
final speed of the vehicle 2 (u_1) = 56 kmph
restitution impulse (e) = 0.6

We have,

$u_1 = v_1 - (1+e)\frac{m_2}{m_1+m_2}(v_1 - v_2)$

$25 = v_1 - (1+0.6)\frac{2.5}{3+2.5}(v_1 - v_2)$

$-1.5v_1 + 4v_2 = 137.5$———(1)

Similarly,

$u_2 = v_2 + (1+e)\frac{m_1}{m_1+m_2}(v_1 - v_2)$

$56 = v_2 + (1+0.6)\frac{3}{3+2.5}(v_1 - v_2)$

$4.8v_1 - 0.7v_2 = 308$———(2)

Solving eqn (1) and (2) we get,
$v_1 = 73$ kmph, and $v_2 = 62$ kmph

Thus from the result we can infer that the follower vehicle was traveling at quite high speed which may have resulted in the collision. The solution to the problem may be speed restriction in that particular stretch of road where accident occurred.

Energy Theory

Applying principle of conservation of energy or conservation of momentum also the initial speed of the vehicle can be computed if the skid marks are known. It is based on the concept that there is reduction in kinetic energy with the work done against the skid resistance. So if the vehicle of weight W slow down from speed v_1 to v_2, then the loss in kinetic energy will be equal to the work done against skid resistance, where work done is weight of the vehicle multiplied by the skid distance and the skid resistance coefficient.

$$\frac{W \times (v_1^2 - v_2^2)}{2g} = W \times f \times S \tag{8.17}$$

Where, f is the skid resistance coefficient and S is the skid distance. It also follows the law of conservation of momentum (m_1, v_1 are the mass and velocity of first vehicle colliding with another vehicle of mass and velocity m_2, v_2 respectively).

$$m_1 v_1 = m_2 v_2 \tag{8.18}$$

Numerical Example:
A vehicle of 2000 kg skids a distance of 36 m before colliding with a stationary vehicle of 1500 kg weight. After collision both vehicle skid a distance of

14 m. Assuming coefficient of friction 0.5, determine the initial speed of the vehicle.

Solution:

Let the weight of the moving vehicle is W_A, let the weight of the stationary vehicle is W_B, skid distance before and after collision is s_1 and s_2 respectively, initial speed is v_1, speed after applying brakes before collision is v_2 and the speed of both the vehicles A and B after collision is v_3, and the final speed v_4 is 0. Then:

After Collision:

Loss in kinetic energy of both cars = Work done against skid resistance

$\frac{(W_A+W_B)\times(v_3^2-v_4^2)}{2g} = (W_A + W_B) \times f \times s_2$

$\frac{v_3^2}{2g} = 0.5 \times 14$

$\therefore v_3 = 11.71$ m/s

During Collision:

Momentum before impact = momentum after impact

$\frac{W_A \times v_2}{g} = \frac{(W_A+W_B)\times v_3}{g}$

$v_2 = \frac{(2000+1500)\times 11.71}{2000}$

$\therefore v_2 = 20.5$ m/s

Before Collision:

Loss in kinetic energy of moving vehicle = work done against braking force in reducing the speed

$\frac{W_A\times(v_1^2-v_2^2)}{2g} = W_A \times f \times s_1$

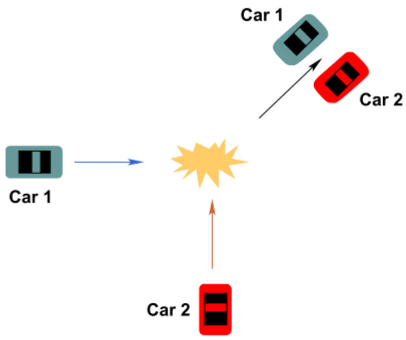
$\frac{(v_1^2-20.5^2)}{2\times 9.81} = 0.5 \times 36$

$\therefore v_1 = 27.8$ m/s $= 100$ kmph

$\therefore$ The pre-collision speed of the moving vehicle is 100 kmph.

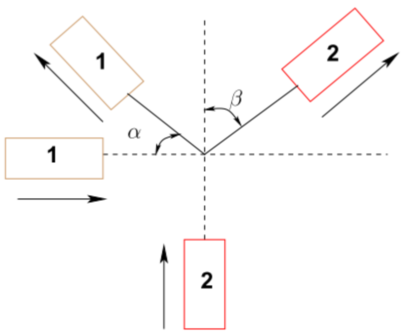
8.5.2 Angular Collision

Angular collision occurs when two vehicles coming at right angles collies with each other and bifurcates in different direction. The direction of the vehicles after collision in this case depends on the initial speeds of the two vehicles and their weights. One general case is that two vehicles coming from south and west direction after colliding move in its resultant direction as shown in the figure below.

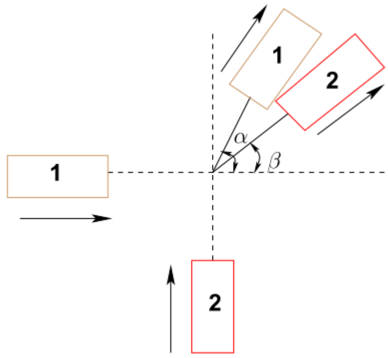


Angular collision of two vehicles resulting in movement in resultant direction

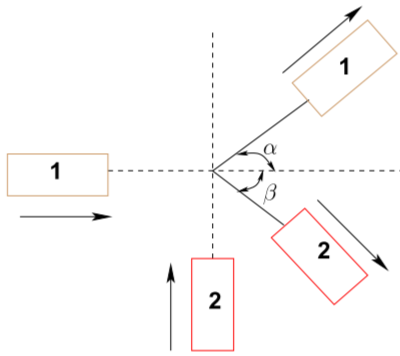
The mass of the car 1 is m_1 kg and the car 2 is m_2 kg and the initial velocity is v_1 m/s and v_2 m/s respectively. So as the momentum is the product of mass and velocity. The momentum of the car 1 and car 2 is m_1v_1 kg m/s and m_2v_2 kg m/s respectively. By the law of conservation of momentum the final momentum should be equal to the initial momentum. But as the car are approaching each other at an angle the final momentum should not be just mere summation of both the momentum but the resultant of the two, Resultant momentum $= \sqrt{(m_1v_1)^2 + (m_2v_2)^2}$ kg m/s. The angle at which they are bifurcated after collision is given by $\arctan(\frac{h}{b})$ where h is the hypotenuse and b is the base. Therefore, the cars are inclined at an angle. Inclined at an angle $= \arctan(\frac{m_2v_2}{m_1v_1})$. Now, since the mass of the two vehicles are same the final velocity will proportionally be changed. The general schematic diagrams of collision are shown in the figures below.



After collision movement of car 1 north of west and car 2 in east of north



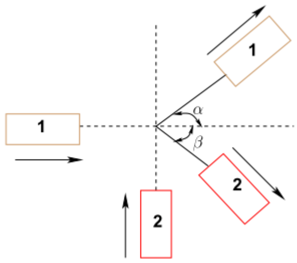
After collision movement of car 1 and car 2 in north of east



After collision movement of car 1 north of east and car 2 in south of east

Numerical Example: Vehicle A is approaching from west and vehicle B from south. After collision A skids 600 north of east and B skids 300 south of east. Skid distance before collision for A is 18 m and B is 26 m. The skid distances after collision are 30m and 15 m respectively. Weight of A and B are 4500 and 6000 respectively. Skid resistance of pavement is 0.55 m. Determine the pre-collision speed.

Solution: Let: initial speed is v_{A1} and v_{B1}, speed after skidding before collision is v_{A2} and v_{B2}, speed of both the vehicles A and B after collision is v_{A3} and v_{B3}, final speed is v_{A4} and v_{B4} is 0, initial skid distance for A and B is s_{A1} and s_{B1}, final skid distance for A and B is s_{A2} and s_{B2}, and weight of vehicle A is W_A and Weight of vehicle B is W_B.



After collision:

Loss in kinetic energy of each cars = Work done against skid resistance

$\frac{W_A(v_{A3}^2 - v_{A4}^2)}{2g} = W_A \times f \times s_{A2}$

$\frac{(v_{A3}^2 - 0)}{2 \times 9.81} = 0.55 \times 30$

$\therefore v_{A3} = 18$ m/s

Similarly, we calculate vB3,

$\therefore v_{B3} = 12.7$ m/s

During collision:

Momentum before impact = momentum after impact (resolving along west-east direction)

$\frac{W_A}{g} \times v_{A2} + 0 = \frac{W_A}{g} \times v_{A3} \cos\alpha + \frac{W_B}{g} \times v_{B3} \cos\beta$

$v_{A2} = v_{A3} \cos\alpha + \frac{W_B}{W_A} \times v_{B3} \cos\beta$

$v_{A2} = 18 \times \cos(60) + \frac{6}{4.5} \times 12.7 \times \cos(30)$

$\therefore v_{A2} = 23.66$ m/s

Now, resolving along the south- north direction, we get,

$\frac{W_B}{g} \times v_{B2} + 0 = \frac{W_A}{g} \times v_{A3} \sin\alpha - \frac{W_B}{g} \times v_{B3} \sin\beta$

$v_{B2} = \frac{W_A}{W_B} \times v_{A3} \sin\alpha - v_{B3} \sin\beta$

$v_{B2} = \frac{4.5}{6} \times 18 \times \sin(60) - 12.7 \times \sin(30)$

$\therefore v_{B2} = 5.34$ m/s

Before Collision:

Loss in kinetic energy of each cars= Work done against skid resistance

$\frac{W_A(v_{A1}^2 - v_{A2}^2)}{2g} = W_A \times f \times s_{A1}$

$v_{A1} = \sqrt{2 \times g \times f \times s_{A1} + v_{A2}^2}$

$v_{A1} = \sqrt{2 \times 9.81 \times 0.55 \times 18 + 23.66^2}$

$\therefore v_{A1} = 27.45$ m/s $= 99$ Km/hr

Similarly, $\therefore v_{B1} = 17.57$ m/s $= 63.26$ Km/hr

8.6 Safety Measures

The ultimate goal is to develop certain improvement measures to mitigate the circumstances leading to the accidents. The measures to decrease the accident rates are generally divided into three groups engineering, enforcement and education. Some safety measures are described below:

8.6.1 Road Signs

Road signs are integral part of safety as they ensure safety of the driver himself (warning signs) and safety of the other vehicles and pedestrians on road (regulatory signs). Driver should be able to read the sign from a distance so that he has enough time to understand and respond. It is essential that they are installed and have correct shape, color, size and location. It is required to maintain them as well, without maintenance in sound condition just their installment would not be beneficial. According to British investigation height of text in road sign should be:

$$H = \frac{(N+6)v}{64} + \frac{3}{4}L \tag{8.19}$$

Where, N = No. of words on the sign, v = speed of vehicle (kmph), L = distance from which inscription should be discernible (m).

8.6.2 Street Lighting

Street lightning of appropriate standard contributes to safety in urban area during night time due to poor visibility. Installation of good lighting results in 21% reduction in all accidents, 29% reduction in "all casualty" accidents, 21% reduction in "non pedestrian casualty" accidents, and 57% reduction in "pedestrian casualty" accidents.

8.6.3 Improvement in Skid Resistance

If road is very smooth then skidding of the vehicles may occur or if the pavement is wet then wet weather crashes occur which account about 20-30%. Thus it is important to improve the skid resistance of the road. Various ways of increasing the skid resistance of road are by constructing high-friction overlay or cutting of grooves into the pavement.

8.6.4 Road Markings

Road markings ensure proper guidance and control to the traffic on a highway. They serve as supplementary function of road sign. They serve as psychological barrier and delineation of traffic path and its lateral clearance from traffic hazards for the safe movement of traffic. Thus their purpose is to provide smooth and safe traffic flow.

8.6.5 Guide Posts With or Without Reflector

They are provided at the edge of the roadway to prevent the vehicles from being off tracked from the roadway. Their provision is very essential in hilly road to prevent the vehicle from sliding from top. Guide posts with reflector guide the movement of vehicle during night.

8.6.6 Guard Rail

Guard rail have similar function as of guide post. On high embankments, hilly roads, road running parallel to the bank of river, shores of lake, near rock protrusion, trees, bridge, abutments a collision with which is a great hazard for a vehicle. It is required to retain the vehicle on the roadway which has accidentally left the road because of fault or improper operation on the part of the driver. Driver who has lost control create a major problem which can be curbed by this measure.

8.6.7 Driver Reviver Stop

Driver reviver stop are generally in use in countries like U.S.A where driver can stop and refresh himself with food, recreation and rest. They play a very important part in traffic safety as they relieve the driver from the mental tension of constant driving. These stops are required to be provided after every 2 hour travel time.

8.6.8 Constructing Flyovers and Bypasses

n areas where local traffic is high bypasses are required to separate through traffic from local traffic to decrease the accident rate. To minimize conflicts at major intersections flyovers are required for better safety and less accident rate.

8.6.9 Regular Accident Studies

Based on the previous records of accidents the preventive measures are taken and after that the data related to accidents are again collected to check the efficiency of the measures and for future implementation of further preventive measures.

8.6.10 Speed Control

Checks on spot speed of all vehicles should be done at different locations and timings and legal actions on those who violate the speed limit should be taken.

8.6.11 Training and Supervision

The transport authorities should be strict while issuing licence to drivers of public service vehicles and taxis. Driving licence of the driver may be renewed after specified period, only after conducting some tests to check whether the driver is fit.

Chapter 9

Vehicle Operating Cost Impacts

9.1 Introduction

Vehicle costs are direct expenses that comprise the costs of vehicle ownership (fixed) and vehicle operation (variable). The latter category, typically referred to as vehicle operating costs (VOCs), varies with vehicle use and is typically expressed in cents per mile traveled by a vehicle. For most transportation modes, VOC involves energy use, tires, maintenance, repairs, and mileage-dependent depreciation. Fixed vehicle costs are those that are largely independent of vehicle use and are generally unaffected by transportation improvements; examples are insurance costs, time-dependent depreciation, financing, and storage. Such costs are therefore typically excluded from VOC impact evaluation of projects.

VOC savings or benefits of a transportation improvement or intervention simply refer to the reduction in vehicle operating costs compared to an existing situation or a base-case alternative.

For areawide or corridor-level projects involving multimodal systems, an improvement in any part of the system can affect VOCs of the other parts or other modes. For example, service improvement in commuter rail or provision of a bus rapid transit along a corridor can affect the level of service on highway facilities in the same corridor because the shift of some travelers from automobile to transit would lead to improved highway level of service due to reduced congestion and thus, lower vehicle operating costs at the highway section.

In this chapter we identify VOC components and factors and present a procedural framework for assessing the VOC impacts of transportation improvements.

9.2 Components of Vehicle Operating Cost

The components of vehicle operating cost are the individual items associated with vehicle operation on which expenses are directly incurred. These include the costs of energy needed to propel the vehicle, fluids, and other light consumables associated with mechanical working of the drivetrain, occasional replacement of the vehicle's contact surfaces with the guideway, vehicle repair and maintenance, and vehicle depreciation.

9.2.1 Fuel

Fuel is a key component of vehicle operating costs. For highway vehicles for instance, fuel costs can account for 50 to 75% of usage-related costs. Fuel cost can be estimated on the basis of fuel efficiency and unit fuel price. Fuel efficiency, in turn, depends primarily on vehicle class, type, age, and speed. Automobile associations, petroleum institutes, and government energy agencies publish fuel prices (dollars per gallon) on a regular basis. In the United States, the average prices of gasoline and diesel in 2005 were $2.2 and $2.4, respectively (USDOE, 2005b). Fuel prices for VOC computation purposes should be derived by subtracting the federal and state gasoline taxes from retail prices. On a mileage basis, the unit costs of fuel (including oil) in 2003–2004 ranged from approximately 7 cents per vehicle-mile for small autos to over 21 cents per vehicle-mile for large trucks (Barnes and Langworthy, 2003; AAA, 2005). Generally, very low speeds, steep uphill grades, and curves lead to higher fuel consumption rates and hence higher overall fuel costs. In the Highway Economic Requirements System (HERS) model (FHWA, 2002), the change in vehicle fuel efficiencies across the years is accounted for in VOC estimation using an adjustment factor.

9.2.2 Shipping Inventory

The inventory cost of cargo (freight transportation) is a special category of user cost. The entity that ships the cargo (the client) is a user of a shipping service made available by a carrier. In the course of transporting perishable or valuable cargo, the client incurs holding costs that represent an opportunity cost: If at the beginning of the shipment, the client had a cash amount worth the cargo being shipped, such an amount would have earned some interest by the time the cargo reaches its destination. So by having the cargo transported, the client is foregoing some benefits. Higher inventory costs are generally directly related to cargo value, greater cargo perishability, higher prevailing opportunity cost of money, and slower speed of the shipping vehicle. To compute the inventory cost for a given vehicle class, an hourly discount rate is typically determined and multiplied by the average value of shipments undertaken by that vehicle class (FHWA, 2002). AASHTO (2003) recommends that the inventory costs of cargo per vehicle-mile should be applied to the unit user cost attributed to cargo-carrying transportation vehicles. The most significant VOC factors that affect the shipping inventory costs are speed and delay, but cargo value and interest rate also can be influential. Higher cargo value and interest rates and greater travel or transfer delay translate to higher unit costs of shipping inventory, and higher speeds lead to lower inventory costs. For example, at a 10% interest rate, two trucks each shipping $100,000 cargo, one traveling at 60 mph and the other at 50 mph, incur inventory costs of approximately 2.5 and 6 cents per mile, respectively (AASHTO, 2003).

9.2.3 Lubricating Oils for Mechanical Working of the Drivetrain

The lubricating oil cost includes the cost of engine oil, transmission fluids, brake fluids, and other similar consumables associated with the operation of vehicle engine and drive train. Oil cost is a product of unit price (dollars/quart) and consumption rates (quarts/mile). The consumption rates depend on the amount of use as well as characteristics of the guideway and vehicle, and operational conditions such as speed, delay, grade, and curves. Typically, the cost of this set of VOC components is reported together with fuel costs, but some sources report them separately.

9.2.4 Preservation of the Vehicle-Guideway Contact Surface

At their points of contact, both the vehicle and guideway experience deterioration due to wear and tear. For highways and runways, the vehicle contact is a tire; for railways, the contact is typically a steel wheel. Updated tire costs (2005 dollars) from the HERS technical report, are as follows: $54.71 per tire for small autos, $86.54 for medium-sized to large autos, $95.39 for four-tire single-unit trucks, $95.38 for sixtire single-unit trucks, $230.10 for single-unit trucks of three or more axles, and $569.74 for combination trucks. Of the various VOC factors, pavement condition, grade, curvature, and speed changes are those that most influence the rate of wear of contact surfaces.

9.2.5 Vehicle Repair and Maintenance

Repair and maintenance costs are incurred on vehicle parts that need replacement or replenishment after some amount of use. For gasoline-powered vehicles, these include the cost of batteries, alternators, fuel pumps, air pump, tire rims, electrical parts such as bulbs and fuses, and so on. These costs also include costs of replacing parts due to crashes, misuse, or other adversarial factors. In some methodologies, the cost of vehicle repair and maintenance is not reported separately but is added to other nonfuel costs. In Year 2005 dollars, the unit cost of vehicle repair and maintenance generally ranged from 4.7 cents per vehicle-mile for small to medium-sized vehicles to 9.3 cents per vehicle-mile for trucks (AAA, 2005). Vehicle repair and maintenance are influenced by pavement condition, curvature, and to a lesser extent, speed, grade, and speed change.

9.2.6 Depreciation

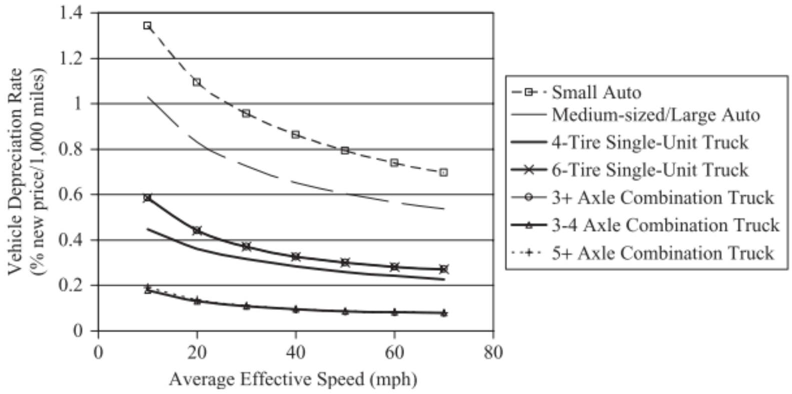
Vehicle depreciation is a function of vehicle usage (miles of travel) and vehicle age (years since manufacture). The table below presents the depreciation costs of selected vehicle classes and types. It can be seen that mileagebased depreciation rates are similar across vehicle classes: This seems reasonable because the lower initial cost of cars is balanced by their shorter service lives compared with trucks, so the net effect is that rates of mileage-based depreciation are similar across vehicle types (Barnes and Langworthy, 2003). Mileage-based depreciation costs can account for a significant fraction of overall vehicle operating costs. In some literature, the cost of vehicle depreciation is reported together with other nonfuel costs.

Average Vehicle Depreciation Costs (2005 Dollars)

	Total Depreciation (cents/h)	Average Travel (mi/y)	Mileage-Related Depreciation		Time-Related Depreciation (cents/h)
			(cents/mi)	(cents/h)	
Small autos	219	11,575	14	80	139
Medium-sized to large autos	257	11,575	12	73	185
Four-tire Single-unit trucks	278	12,371	6	36	242
Six-tire	393	10,952	10	55	338
3+ axles Combination trucks	1,122	15,025	22	209	913
3 or 4 axles	946	35,274	7	129	817
5+ axles	1,017	66,710	8	232	785

Source: Cost values are updated from their 1995 values in FHWA (2002).

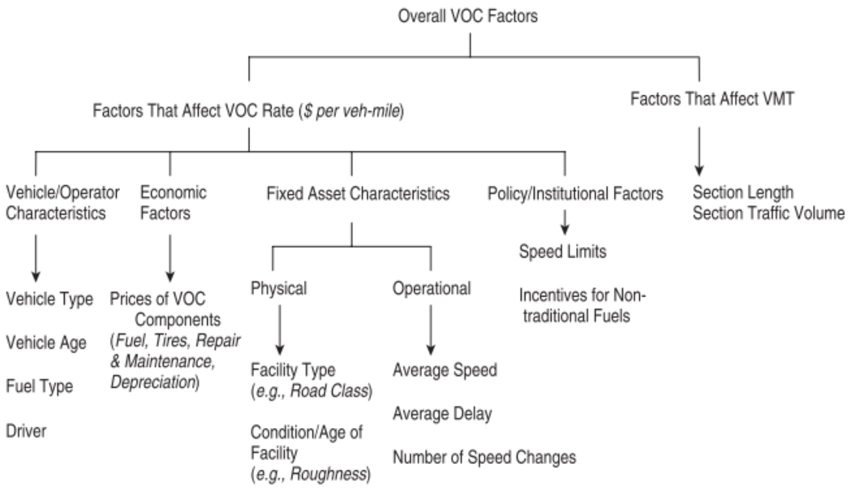
The values presented in the above table are average values. Depreciation rates actually vary by factors such as grade, curves, surface condition, and speed. An improvement in the transportation facility can produce a smoother pavement and improved driving conditions (through reduced stop-and-go situations). Also, all other factors remaining the same, increased speed can lead to reduced depreciation rates, as illustrated in the figure below, for straight constant-speed sections (FHWA, 2002).



Depreciation rate by speed for straight sections (from FHWA, 2002.).

9.3 Factors That Affect Vehicle Operating Cost

For all modes of transportation, vehicle operating costs are affected by factors such as vehicle–operator characteristics, economic factors, condition and other characteristics of the fixed transportation facility, and policy–institutional factors. Although we focus on highway transportation in this section, the principles and concepts can be adapted to other transportation modes. The figure below shows the categories of highway VOC factors.



9.3.1 Vehicle Type

Vehicle operating costs are influenced by size, class, and other vehicle characteristics. Trucks and buses generally have higher operating costs than automobiles, as they consume more fuel and oil and have higher prices for their vehicle parts. Even for a given vehicle type, there could be changes in

VOC over time due to improved vehicle technology and fuel efficiency. If the analyst seeks to carry out long-term VOC impact evaluation, future levels of fuel efficiency could be extrapolated from past trends and duly factored in the VOC computation process.

In some cases, analysts may seek the operating costs associated with bicycling and walking to facilitate a more comprehensive comparison of transportation alternatives that include these modes. A standard bicycle with basic accessories can cost $100 to $500 with annualized maintenance costs of $20 to $40 for tire replacement, tire pumping, and security; for walking, the main consumable is that of footwear, which typically lasts 500 to 5000 miles of walking distance (VTPI, 2004). The human energy use associated with walking and cycling may be considered a benefit rather than a cost, particularly if traveling using these transportation modes substitute for other exercise activities.

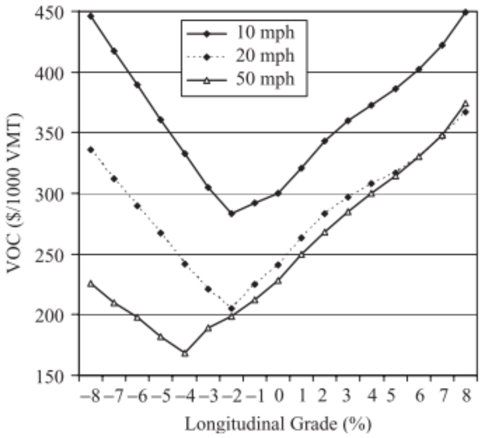
9.3.2 Fuel Type

The uncertainties in supply and increasing costs of fossil fuels coupled with their adverse environmental effects have led to growing use of alternative energy sources for transportation. In evaluating the impacts of transportation improvements, therefore, analysts need to account for the increasing percentage of alternative-fuel vehicles in the traffic stream. At the current time, electric and hybrid vehicles have relatively high purchase costs (150 to 200% of the price of a comparable gasoline car). Electric cars require new battery sets every 20,000 to 30,000 miles costing $2000 to $3000 (averaging 6 to 15 cents per vehicle-mile), and consume 0.25 to 0.5 kWh per mile, so energy costs average 2 to 5 cents per kWh based on typical residential energy rates (USDOE, 2005a). The maintenance costs, including battery replacements, are significantly higher for electric cars (over four fold) compared to hybrid or conventional cars (VTPI, 2005). Even with traditional fuels, there are differences in cost across fuel types: in 2005, the average price of diesel was approximately 10% higher than that of regular leaded gasoline. Also, there are price differences across the three standard grades of gasoline.

9.3.3 Longitudinal Grade

Uphill movements impose additional loads on vehicle engines and therefore require greater consumption of energy compared to downhill or level movements. For downhill trips, fuel consumption is lower than for uphill or level trips, but increased brake applications may lead to increased wear and tear of brake linings and therefore to increased cost of the brake maintenance component of VOC. The figure below illustrates the general relationships between grade and VOC at various speeds for medium-sized automobile.

Generally, overall VOC is lowest for sections with gentle downward slopes (0 to -4%).



Numerical Example:

A 2.15-mile section of State Road 25 on rolling terrain received major improvements in vertical alignment. The average grade of the section was reduced from 3.2% to 2.5%. Traffic volume and composition, and speed were the same after the improvement. Assume that the traffic stream has a 50:50 directional split and is composed primarily of medium-sized automobiles, and the traffic volume is 43,340 vpd (vehicles per day). In both cases, the average speed is 50 mph. What is the first year user benefit in terms of VOC?
Solution:

Section Length (L) = 2.15 mile
Average Daily Traffic (ADT) = 43340 vpd
Average speed of the vehicle (v) = 50 mph

Before Improvement:
uphill: VOC at +3.2% grade (VOC_1) = \$ 275/1000 VMT (using graph for medium-sized automobile)
downhill: VOC at -3.2% grade (VOC_2) = \$ 190/1000 VMT (using graph for medium-sized automobile)
Average VOC before improvement (U_1) = $\frac{275+190}{2}$ = \$232.5/1000 VMT
Vehicle Miles Travelled per year (VMT_1) = $L \times ADT \times 360 = 2.15 \times 43340 \times 360$ = 34011035 vehicle miles per year

After Improvement:
uphill: VOC at +2.5% grade (VOC_1') = \$ 260/1000 VMT (using graph for medium-sized automobile)

downhill: VOC at -2.5% grade ($VOC_2^{'}$) = \$ 200/1000 VMT (using graph for medium-sized automobile)
Average VOC after improvement (U_2) = $\frac{260+200}{2}$ = \$230/1000 VMT
$\because$ the traffic volume remains unchanged after improvement in longitudinal grade,
$VMT_2 = VMT_1$

Now,
Change in unit costs ($U_1 - U_2$) = 232.5 - 230 = 2.5/1000 VMT = 0.0025 VMT

$\therefore$ First year user benefits = $0.5 \times (U_1 - U_2) \times (VMT_1 + VMT_2)$
= $0.5 \times 0.0025 \times 2 \times 34011065$ = \$ 85028

9.3.4 Vehicle Speed

Vehicle operating speed is the dominant factor in determining VOC (Bennett, 1991; Thoresen and Roper, 1996; Bennett and Greenwood, 2001; FHWA, 2002). Transportation improvements influence travel speeds and therefore can profoundly affect VOC. For some vehicles, fuel consumption decreases with increasing speed to a certain point, after which there is little significant change (or sometimes, an increase) in fuel consumption with increasing speed. Factors that affect operating speeds, and subsequently influence fuel VOC, are speed limits (set by policy) and traffic conditions (which vary by the time of day—peak vs. nonpeak). In this section we discuss the impact of speed on shipping inventory costs and present some VOC models based on speed and other factors.

a) Inventory Shipping

Inventory cost is affected by vehicle speed and is calculated as follows:

$$U_{IC} = 100 \times \frac{r}{365 \times 24} \times \frac{1}{S} \times P \tag{9.1}$$

where U_{IC} is the user inventory cost in cents per vehicle mile, r the annual interest rate, P the cargo value in dollars, and S the vehicle speed in miles per hour.

Numerical Example:
Due to a new speed limit policy, the average truck operating speed on a certain interstate freeway increased from 56.5 mph to 61.2 mph. Find the decrease in shipping inventory costs per year for trucks that comprise 22% of the overall traffic stream of 82,500 vehicles per day (vpd). Each truck hauls an average of \$1.5 million worth of goods daily. Assume an 8% interest rate.
Solution:

The daily changes in inventory costs per truck due to the change in travel speed, ΔU_{IC}, can be estimated as follows:
$\Delta U_{IC} = 100 \times \frac{r}{365 \times 24} \times \left(\frac{1}{S_0} - \frac{1}{S_1}\right) \times P$
$100 \times \frac{0.08}{365 \times 24} \times \left(\frac{1}{56.5} - \frac{1}{61.2}\right) \times 1500000$
$\therefore \Delta U_{IC} = 1.9178$ cents/vehicle-mike

Then,
Number of trucks per year = (0.22)(82,500)(365) = 6,624,750
$\therefore$ total reduction in inventory costs for all trucks per year = (1.9178/100)(6,624,750) = \$127,050 per mile

b) VOC Models and Look-up Table Based on Speed and Vehicle Class

Hepburn (1994) developed a VOC model for urban roadways that considers the sum of four VOC components (tires, vehicle depreciation, maintenance, and fuel) as a function of two VOC factors: speed and vehicle class. The model is particularly useful for evaluating VOC impacts of transportation interventions that mostly yield a change in average operating speeds or policies that cause a shift in vehicle class distribution. The Hepburn function is as follows:

For "low" average travel: speeds (< 50 mph)

$$VOC = C + \frac{D}{S} \tag{9.2}$$

For "high" average travel: speeds (> 50 mph)

$$VOC = a_0 - a_1 \times S + a_2 \times S^2 \tag{9.3}$$

where VOC is in cents/mile, S is speed (mph) and C, D, a_0, a_1,and a_2 are coefficients that are functions of vehicle class. The coefficient values are given in the table below:

Vehicle Type	C	D	a_0	a_1	a_2
small automobile	24.8	45.5	27.2	0.035	0.00021
medium-sized automobile	28.5	95.3	33.5	0.058	0.00029
large automobile	29.8	163.4	38.1	0.093	0.00033

The Hepburn model assumes that depreciation depends entirely on vehicle use and that the depreciation rate is constant throughout vehicle life. Furthermore, the model is for tangent, level, and urban road sections with pavement roughness assumed to remain constant over time, and all VOC component costs assumed to vary with distance, with the exception of fuel cost, which varies with speed. It does not explicitly consider the consumption rates and prices of individual VOC components for each vehicle class but is nevertheless useful for quick estimation of VOC.

Numerical Example:
A straight and level urban arterial has an average operating speed of 35 mph. What is the unit VOC of medium-sized automobiles that use this highway?
Solution:

Knowing the values of C and D from Table:

$VOC = C + \frac{D}{S} = 28.5 + \frac{95.3}{35} = 31.22$ cents/vehicle-mile

c) VOC model based on speed, grade, and vehicle class

Zaniewski (1982) provided a VOC model as a function of speed, grade, and vehicle class. Table 7.4 presents the VOCs for medium-sized autos, with updated cost values. If the project section consists of several segments with different grades or VMTs, the unit VOC (dollars/vehicle-mile) is estimated separately for each segment. It should be noted, however, that the vehicles at the time of the Zaniewski study (ca. 1980) had 17% lower fuel efficiency than vehicles in 1997 (FHWA, 2002), and even lower compared to vehicles in 2005. As such, Table given below should be used after stating the necessary assumptions regarding fuel efficiency, or after making due adjustments for fuel efficiency changes over the years.

Fuel Consumption (Gallons) per Minute of Delay by Vehicle Type

Free-Flow Speed (mph)	Small Automobile	Large Automobile	SUV	Two-Axle Single-Unit Truck	Three-Axle Single-Unit Truck	Multiple-Unit Truck
20	0.011	0.022	0.023	0.074	0.102	0.198
25	0.013	0.026	0.027	0.097	0.133	0.242
30	0.015	0.030	0.032	0.122	0.167	0.284
35	0.018	0.034	0.037	0.149	0.203	0.327
40	0.021	0.038	0.043	0.177	0.241	0.369
45	0.025	0.043	0.049	0.206	0.280	0.411
50	0.028	0.048	0.057	0.235	0.321	0.453
55	0.032	0.054	0.065	0.266	0.362	0.495
60	0.037	0.060	0.073	0.297	0.404	0.537
65	0.042	0.066	0.083	0.328	0.447	0.578
70	0.047	0.073	0.094	0.360	0.490	0.620
75	0.053	0.080	0.105	0.392	0.534	0.661

Source: Adapted from AASHTO (2003).

Numerical Example:
A highway section consists of two segments A and B that have the charac-

teristics listed in table below. Determine the total vehicle operating costs for each segment. Assume that all vehicles are medium sized automobiles, and assume further that the values in the table reflect current fuel consumption rates.

	Segment A	Segment B
Traffic volume (ADT)	5320	8580
Average grade (%)	+4	+1.5
Speed (mph)	30	50
length (miles)	5.7	2.6
Directional Split	68% (upward slope); 32% (downward slope)	45% (upward slope); 55% (downward slope)

Solution:
Segment A:

Total Unit VOC = Unit VOC for uphill + unit VOC for downhill
$= unitVOC_{vehiclespeed,grade,vehicleclass} \times$ uphill directional split $+ unitVOC_{vehiclespeed,grade,vehicleclass} \times$ downhill directional split
$= 319 \times 0.68 + 227 \times 0.32 = \$289.56/1000$ VMT

VMT = section length *times* ADT = $5.7 \times 5320 = 30325$ vehicle-miles per day

Overall VOC = Total unit VOC $\times$ VMT = $\frac{289.56}{1000} \times 30325 = \$8781$ per day

Segment B:

Total Unit VOC = Unit VOC for uphill + unit VOC for downhill
= uphill grade $\times$ uphill directional split + downhill grade $\times$ downhill directional split
$= 292 \times 0.45 + 232 \times 0.55 = \$259/1000$ VMT

VMT = section length *times* ADT = $2.6 \times 8580 = 22308$ vehicle-miles per day

Overall VOC = Total unit VOC $\times$ VMT = $\frac{259}{1000} \times 22308 = \$5778$ per day

d) VOC Models Based on Speed, Gradient, Curvature, and Pavement Condition

Some VOC models, such as the World Bank's HDM (Bennett and Greenwood, 2001) and the HERS model (FHWA, 2002), estimate the unit cost of each VOC component as a function of speed, grade, and pavement condition. This is done for basic sections (straight sections with constant speed), and then excess vehicle operating costs due to speed changes and curvature are calculated. The excess VOC is added to the basic costs to yield the overall VOC for the section.

9.3.5 Delay

Nodes and links in the networks of various transportation modes may often experience delay, which translates into higher vehicle operating costs. In evaluating transportation improvements at such facilities, VOC costs, particularly for fuel and inventory, can be expressed as a function of time delay. On highway links, for instance, delay can involve decelerating to a stop, idling, and accelerating from a stopped position. Such stop-and-go traffic leads to additional strain on a vehicle, which is translated into higher use of fuel and oil. All three phases involve fuel consumption rates that generally exceed that of constant-speed travel. The primary share of overall delay costs is attributed to acceleration of vehicles after being slowed or stopped rather than fuel consumed in decelerating or idling during delay periods (AASHTO, 2003). The impact of travel delay on VOC (fuel and inventory shipping cost components) can be estimated using a methodology provided by AASHTO (2003). In the methodology, the analyst estimates the delay with and without improvement using field measurements (applicable only to the existing situation), simulation, or analytical travel delay models. Using the estimated change in delay, fuel consumption rates per minute of delay and fuel price, the total cost of delay can be calculated.

Fuel Consumption (Gallons) per Minute of Delay by Vehicle Type

Free-Flow Speed (mph)	Small Automobile	Large Automobile	SUV	Two-Axle Single-Unit Truck	Three-Axle Single-Unit Truck	Multiple-Unit Truck
20	0.011	0.022	0.023	0.074	0.102	0.198
25	0.013	0.026	0.027	0.097	0.133	0.242
30	0.015	0.030	0.032	0.122	0.167	0.284
35	0.018	0.034	0.037	0.149	0.203	0.327
40	0.021	0.038	0.043	0.177	0.241	0.369
45	0.025	0.043	0.049	0.206	0.280	0.411
50	0.028	0.048	0.057	0.235	0.321	0.453
55	0.032	0.054	0.065	0.266	0.362	0.495
60	0.037	0.060	0.073	0.297	0.404	0.537
65	0.042	0.066	0.083	0.328	0.447	0.578
70	0.047	0.073	0.094	0.360	0.490	0.620
75	0.053	0.080	0.105	0.392	0.534	0.661

Source: Adapted from AASHTO (2003).

a) Change in Fuel Costs due to Delay Change

For a given vehicle class, the change in fuel costs due to a change in travel delay is found as follows (AASHTO, 2003):

$$change\ in\ fuel\ VOC = g \times (D_0 - D_1 \times p) \tag{9.4}$$

where: g is the fuel consumption in gallons per minute of delay (from the above table), $D_0 - D_1$ = change in delay (minutes) due to transportation improvement, and p is the price of the fuel. The parameters g and p are specific to vehicle class.

Example:
Modernization and optimization of the traffic signal system at a busy urban arterial yielded, on average, a 9-minute reduction in delay per trip for users of the arterial. The traffic volume is 4300 vph and is composed of 25% small autos, 30% large autos, 25% SUVs, 10% two-axle single-unit trucks, 5% three axle single-unit trucks, and 5% multiple-unit trucks. After improvement, average free-flow speed increases from 45 mph to 50 mph, and traffic volume and composition remain unchanged. Determine the reduction in fuel costs during peak hours due to the decrease in delay. Assume that fuel cost is $2.20 per gallon. Use the fuel consumption rates provided in the table above, and assume simple averages across vehicle classes.

Solution:
The traffic volume for each vehicle class is determined by multiplying the percentage composition by the total traffic volume. Using the table above, the fuel consumption rates are determined, and the change in fuel consumption cost is presented the table below.

Estimation of Change in Fuel Consumption Costs due to Delay

	Small Auto	Large Auto	SUV	Two-Axle Single-Unit Truck	Three-Axle Single-Unit Truck	Multiple-Unit Truck	Total
Traffic volume (vph)	1075	1290	1075	430	215	215	4300
Fuel consumption rate (gals/min)	0.025	0.043	0.049	0.206	0.280	0.411	
Fuel price ($/gal)	$2.2 (average for all vehicle classes)						
Change in delay due to the improvement, $D_0 - D_1$	9 min (average for all vehicle classes) per peak hour						
Change in fuel consumption costs	$532	$1098	$1043	$1754	$1192	$1750	$7369

Therefore, the total reduction in fuel costs during the peak hours due to the decrease in delay is $7,369/hr.

b) Change in Shipping Inventory Costs due to Delay

AASHTO (2003) provides a methodology for estimating the impact of time delay on shipping operating costs, as follows: The change in inventory cost per shipping vehicle due to a change in delay is given by $\Delta I(D) I(D) \times \Delta D$, where ΔD is the change in delay (minutes) and inventory costs (cents per vehicle-minute) is given by,

$$I(D) = 100 \times \frac{r}{365 \times 24 \times 60} \times P \tag{9.5}$$

where r is the interest rate (per annum) and P is the dollar value of the cargo being transported by the shipping vehicle.

In some cases, the analyst is provided with an estimate of the expected change in delay, but in other cases, change in delay will need to be estimated (by calculating the delay before and after the improvement). Delay can be estimated on the basis of prevailing traffic conditions and road inventory.

Numerical Example:
A freeway was constructed in 2005 to bypass a city center. This improvement led to a 10-minute reduction in travel delay per trip for shippers who transport goods across the city. If the average value of cargo is $265,000 per truck and the interest rate is 6%, determine (a) the shipping inventory costs per vehicle before the construction, (b) the reduction in shipping inventory costs due to the construction in 2005, and (c) the change in user benefits accrued to shippers in 2005 compared to pre-construction conditions. The pre-construction period daily truck traffic (ADTT) was 33,000, and the trip time was 1.5 hours. Assume a 5% ADTT increase due to induced demand.
Solution:
a) Shipping inventory cost per truck
The unit inventory cost of the shipment before improvement can be calculated as follows:
$I(D) = P \times [r/(365 \times 25 \times 60)] = 265000 \times [0.06/(365 \times 25 \times 60)] = \$0.03025/truck - minute$

The unit inventory cost after the improvement is the same as that before the improvement because there is no change in the total cargo value and the annual interest rate. Since the travel time reduces by 10 minutes after the improvement, the total inventory cost saved due to the improvement is
Change in unit inventory cost $= \$0.03025/truck - minute \times 10\, minutes =$ $0.3025/truck

b) Reduction in Shipping Inventory Cost
The unit shipping inventory cost in dollars/truck-mile, U, can be calculated as follows:
Before Improvement:

U_{before} = \$0.03025/$truck - minute \times (60/S_{before})$ = \$1.815T/L per truck-mile

where S_{before} = average speed before improvement (mph) = L/T, L = average truck trip length (miles) and T = average truck trip time (hours).

$\therefore$ Total yearly inventory cost = (\$1.815T/L)(ADTT)(L)(365)

After Improvement:

U_{after} = \$0.03025/$truck - minute \times (60/S_{after})$ = \$1.815[T - (10/60)]/L per truck-mile

Therefore, the reduction in total shipping inventory cost = Total yearly inventory cost before the improvement - Total yearly inventory cost after the improvement = 1.815 ADTT (365) [T- 1.05T- (10/60)] = 1.815 ADTT (365) [0.175 - 0.05T] = \$2,186,168 in the first year.

c) Change in user benefit (Change in consumer surplus)

The change in user benefits can be calculated based on the change in consumer surplus, which is given by the following formula:

User Benefits = $\frac{1}{2} \times (U_{before} - U_{after}) \times (VMT_{before} + VMT_{after})$
= 0.5 ([(1.815 (10/60)/L)][2.05 ADTT (L)(365)] = = \$3,734,703.

9.3.6 Speed Changes

Vehicles travel at different speeds due to geometric and/or traffic conditions. It has been shown that the more frequent the speed change of a vehicle, the higher the associated operating cost, particularly its fuel component. When vehicles slow down or pick up speed, they experience additional strain that is translated into a higher use of fuel and oil. As such, highway projects that smoothen traffic flow by reducing the frequency and intensity of speed changes ultimately reduce the costs of vehicle operation.

An extreme case of speed change is stop–start conditions, which are usually typical of city driving. Barnes and Langworthy (2003) showed that for maintenance, repair, and depreciation, worsening stop–start conditions will increase costs of fuel consumption and to a smaller extent, the costs of maintenance, repair, and depreciation.

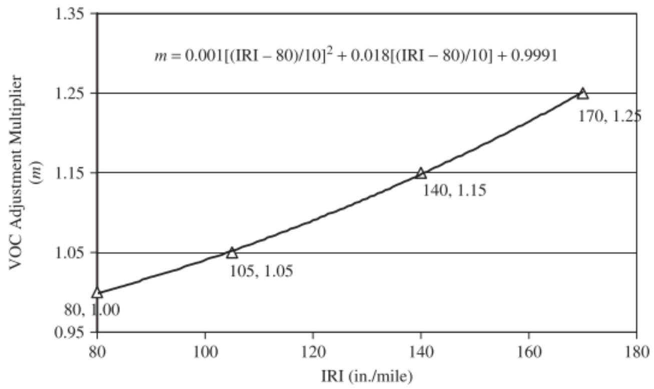
9.3.7 Horizontal Curvature

A vehicle negotiating a horizontal curve requires extra energy to counter centrifugal forces in order to stay in a radial rather than a tangential path. Furthermore, the side friction increases tire wear and tear, and the frequency and cost of maintenance and replacement. The VOC due to curves involves fuel, tire, and maintenance and repair, and is typically expressed as a function of the rate of consumption and unit prices of these VOC components, vehicle type, and average speed. In the HERS methodology, VOC for curve negotiation speed is estimated separately for sections with low speeds ($<$ 55 mph) and those with high speeds ($>$55 mph).

9.3.8 Road Surface Condition

To some extent, pavement roughness, often measured in terms of the present serviceability index (PSI) or international roughness index (IRI), can affect the maintenance, tire, repair, and depreciation cost components of VOC. This is because the motion of vehicle tires on a rough pavement surface is associated with greater resistance to movement, which can lead to higher levels of fuel consumption compared to traveling at a similar speed on a smooth surface; and a bumpy ride, which leads to increased vibration and wear and tear of vehicle parts. Also, an indirect effect of poor pavement conditions is that road users may be forced to drive at lower speeds, leading to higher fuel consumption. Transportation projects such as resurfacing that improve pavement surfaces can therefore lead to reductions in VOCs.

Zaniewski (1982) suggested that there can be significant impacts of pavement roughness on nonfuel vehicle operating cost components, particularly for rough pavements. Most other research on the relationship between pavement condition and VOC has been conducted outside the United States by the World Bank and other international agencies. Examples include a New Zealand study (Opus Central Laboratories, 1999) which suggests that at superior levels of pavement condition (low roughness), increments in condition have relatively little incremental effect on vehicle operating cost (Figure 7.4), and that additional costs of vehicle operation start to accrue only when the IRI exceeds approximately 100 in/mi (3.33 m/km). For paved roads in poor condition and for gravel roads, changes in road surface condition can lead to significant reductions in VOC. Barnes and Langworthy (2003) reported on a previous study that suggested that a unit increase in IRI (in m/km) can generally lead to an increase of $200 (1.67 cents/vehicle-mile, assuming 12,000 annual mileage) in vehicle maintenance and repair costs alone. Also, Barnes and Langworthy (2003) developed adjustment factors for all VOC components combined, as a function of pavement condition (in the figure below). They assumed a baseline PSI of 3.5 or better (an IRI of about 85 in/mi or 1.35 m/km), at which further increases in pavement condition would have no impact on operating costs, and then adjusted for three levels of rougher pavement as shown in the figure. The figure can be used to estimate the VOC corresponding to a given pavement condition. For the depreciation component, there seem to be relatively few studies that have explicitly shown a relationship with pavement roughness. However, it seems obvious that in the long term, a vehicle which is operated on a rough pavement surface is likely to lose its value faster than one that is operated on a smooth-surfaced pavement.



VOC adjustments for pavement roughness levels.

Numerical Example:
A warranty HMA resurfacing project on Interstate 599 yielded a performance jump of 40 IRI (in/mi). If the base vehicle operating cost is \$143 per 1000 vehicle-miles, (a) determine the change in unit VOC due to resurfacing. Use the Barnes and Langworthy relationship. The IRI before improvement was 110 in/mi. (b) If the traffic volume is 67,500 vpd and the section is 6.5 miles in length, determine the overall change in VOC.
Solution:
a) Before Improvement:
IRI = 110 in/mi, and the VOC adjustment multiplier is given by

$$\text{m} = 0.001 \times \left(\frac{110-80}{10}\right)^2 + 0.018 \times \frac{110-80}{10} + 0.9991 = 1.06$$

$$\therefore VOC = (1.06)(143) = \$151.58/1000 \text{ VMT}$$

b) After Improvement:
IRI = 110 - 40 = 70 in/mi, m = 1.00 since 70 is less than 80, and therefore VOC = \$143/1000 VMT. Change in unit VOC = 151.58 - 143 = \$8.58/1000 VMT.

$\therefore$ Overall Change in VOC = (\$8.58)(67,500)(365) (6.5)/1000 = \$1.374 million.

9.3.9 Other VOC Factors

Other factors that can influence the cost of vehicle operation include driver behavior, condition of vehicle, vehicle weight, prices of vehicle maintenance (reflected in costs of labor, vehicle consumables, and spare parts), and weather severity. Operating costs for transit vehicles (such as buses and trolleys) are also affected by other factors, such as transit schedules (which typically depend on passenger demand) and vandalism.

9.4 Procedure for Assesing VOC Impacts

The framework for assessing VOC impacts of transportation interventions revolves around three tasks:

1. Estimating the unit VOC rates (i.e., dollars/vehiclemile) with and without intervention

2. Estimating the amounts of travel (VMT) with and without the intervention

3. Calculating the user VOC benefits of intervention

Bibliography

[1] Insurance regulation. Insurance Act, 2049, 1993.

[2] Tolga Bektas. Freight Transport and Distribution. CRC Press, 2017.

[3] B. Kent Lall C. Jotin Khisty. Transportation Engineering: An Introduction. Pearson Education Inc., 3 edition, 2003.

[4] Milan Janic Dusan Teodorovic. Transportation Engineering: Theory, Practice, and Modelling. Elsevier Inc., 2017.

[5] Michael J. McNamara George E. Rejda. Principles of Risk Management and Insurance. Pearson Education Inc., 13 edition, 2017.

[6] Samuel Labi Kumares C. Sinha. Transportation decision making : principles of project evaluation and programming. John Wiley Sons, Inc., 2007.

[7] N. Gregory Mankiw. Principles of Economics. CENGAGE Learning, 8 edition, 2018.

[8] Dr. Tom V. Mathew. Transportation Systems Engineering. IIT Bombay, 2014.

[9] P. Rama Murthy. Operations Research. New Age International Publishers, 2 edition, 2007.

[10] Lester A. Hoel Nicholas J. Garber. Traffic and Highway Engineering. CENGAGE Learning, 4 edition, 2009.

[11] JK Sharma. Operations Research: Theory and Applications. Trinity Press, 6 edition, 2016.

[12] Hamdy A. Taha. Operations Research: An Introduction. Pearson Education Inc., 10 edition, 2017.

[13] Michael P. Walsh. Vehicle Emission Reductions. European Conference of Ministers of Transport (ECMT), 2000.