Traffic Operation

1. Introduction

Traffic operations is the analysis of volume and capacity relationships of transportation facilities, with the goal of predicting operational performance in the form of travel time, speeds, delay, queues, and various other measures. As traffic volumes approach, and in some cases exceed, capacity, these performance measures deteriorate (higher delay, lower speed), up to potentially reaching a failure criterion, such as a queue spilling back beyond the physical storage space available for the movement at an intersection. So, as traffic volume, or the load on a transportation segment or approach to an intersection, increases, it is expected that the performance at some point starts to degrade, and eventually break down. Analysts also refer to the performance of a transportation facility as its level of service (LOS) and sometimes its quality of service (QOS), which are both important concepts.

The field of traffic operations relies on a combination of field data and established analysis methods to provide analysts with methodologies for predicting the performance of an intersection or segment. The methodologies used to conduct operational analyses are typically derived from either broad field observations at representative sites (e.g., through regression), or from traffic flow theory relationships. Those traffic flow relationships in turn can often be calibrated to better fit local operating conditions. The methods then employ a series of analytical steps, each with equations, tables, or charts, to allow for the estimation of performance from available inputs. The analyst needs to have a good understanding of these methodologies and their empirical or theoretical basis to make sure that they are applied correctly and within their intended scope.
This part provides an overview of the basics of transportation operations, including fundamental relationships of traffic flow, and established analysis methods for various transportation facility types.

2. Purpose of Traffic Operations

Transportation operational analysis methods are applied to estimate the current or, more often, a future state of the transportation system. The methods provide a structured process for estimating performance from a set of traffic inputs. These inputs most fundamentally include the traffic volume level and the roadway geometry, with the first expressing the demand side of how many cars, trucks, pedestrians, or bicycles aim to travel through a point, and the latter the supply side, which describes how many cars, trucks, pedestrians, or bicycles can travel through that point. Between the supply and demand side lie empirically derived relationships in the form of equations, tables, and charts that estimate performance based on those inputs. At the most disaggregated level, these methods exist for points (e.g.,

how much traffic can pass through a gate or a bridge), but are more practically useful once aggregated to the intersection, segment, or facility level.
The demand side of a traffic operation problem is quantified through the number of objects (cars, trucks, pedestrians, bicycles, etc.), the mix of the traffic stream (e.g., percent trucks), and the temporal distribution of traffic over say the course of a day. The supply side is expressed through the type of intersection, the number of lanes, or the radius of curves for turning movements. Researchers and transportation analysts have then developed relationships to estimate the capacity of the network element under study, as well as methods to predict performance of those elements through measures that typically include the average speed, travel time, density, queue length, or number of stops.
Traffic operational analyses are performed by transportation professionals with training on these operational methods and an understanding of the relationships between volume and capacity (demand and supply) for various traffic network elements. Commonly, the actual analysis is facilitated by software, as the equations can be complex, and because their application can be repetitive, as an analyst evaluates multiple approaches, to multiple time intersections, over multiple time periods, and for multiple scenarios.
One final key element in traffic operations is data. Data are needed to quantify both demand and supply levels, before being able to estimate the operational performance. In their most basic form, data take the form of traffic counts—demand—and a count of the number of lanes—supply. But as we'll see in this part, the nuances of demand and how traffic is measured, what the vehicle mix is, how traffic is distributed temporally, and so on, tend to be more complicated. Similarly, the supply characteristics are much more complex to quantify than just counting the number of lanes, as the capacity of each lane is impacted by physical characteristics (lane width, shoulder clearance, radius, etc.) and, more importantly, attributes of traffic control devices that include traffic signals, yield lines, and stop lines. The capacity of a transportation system element is also different for different road users, as one can, for example, fit more cars per hour than trucks per hour, or more pedestrians than bicycles through the same point. Put differently, the capacity or supply, in turn, can be a function of the demand on the system.

3. Highway Capacity Manual

The U.S. Highway Capacity Manual, or HCM, is the primary reference for traffic operational analysis, methodologies, and level of service (LOS) concepts in the United States, as well as many other countries. The HCM is a collection of concepts and methods that guide analysts on how to evaluate a particular type of intersection or roadway segment, based on what can be extensive national or international datasets of operational performance. The HCM is also the primary source for defining the capacity of different roadway elements that are used in many applications beyond traffic operations, including transportation planning and even safety analyses. The objectives of the HCM are to:

- Define performance measures and describe survey methods for key traffic characteristics.
- Provide methodologies for estimating and predicting performance measures.

- Explain methodologies at a level of detail such that readers can understand the factors that have an effect on multimodal operation.

While the HCM is the primary traffic operations resource developed for the United States, many other countries have adopted the HCM for their own use, often with some country-specific modification and calibration of methods to better suit local conditions. Some countries have developed their own traffic operations manuals, such as, for example, the German HBS (manual for measuring street systems, FGSV, 2001).

The HCM, first produced in 1950, is updated regularly based on new research supported by the Transportation Research Board (TRB). The manual is updated through the Highway Capacity and Quality of Service Committee of TRB. HCM 2010, the most recent edition, was distributed in early 2011; a major update of the 2010 HCM is expected for publication in late 2015.

The HCM is principally organized into four volumes, covering:

- general concepts,
- methods for uninterrupted flow (freeways),
- methods for interrupted flow (arterial streets), and
- Supplemental information to further document the methods in the second and third volumes.

Each volume is organized into chapters that describe a particular element of the transportation system, ranging from basic freeway segments to signalized intersections, to modern roundabouts, to shared-used pedestrian and bicycle paths. Table below shows the high-level organization of the 2010 Highway Capacity Manual in the three primary printed volumes. Volume 4 is an online-only volume with additional reference materials in support of the printed chapters.

Table: Overview of HCM 2010 organization.

Volume 1: Concepts	Volume 2: Uninterrupted Flow	Volume 3: Interrupted Flow	Volume 4: Supplemental Material and Applications
			
Ch. 1—*HCM* User's Guide Ch. 2—Applications Ch. 3—Modal Characteristics Ch. 4—Traffic Flow and Capacity Concepts Ch. 5—Quality and Level of Service Concepts Ch. 6—*HCM* and Alternative Analysis Tools Ch. 7—Interpreting *HCM* and Alternative Tool Results Ch. 8—*HCM* Primer Ch. 9—Glossary and Symbols	Ch. 10—Freeway Facilities Ch. 11—Basic Freeway Segments Ch. 12—Freeway Weaving Segments Ch. 13—Freeway Merge and Diverge Segments Ch. 14—Multilane Highways Ch. 15—Two-Lane Highways	Ch. 16—Urban Street Facilities Ch. 17—Urban Street Segments Ch. 18—Signalized Intersections Ch. 19—Two-Way Stop Controlled Intersections Ch. 20—All-Way Stop Controlled Intersections Ch. 21—Roundabouts Ch. 22—Interchange Ramp Terminals Ch. 23—Off-Street Pedestrian and Bicycle Facilities	Ch. 24—Concepts: Supplemental Ch. 25—Freeway Facilities: Supplemental Ch. 26—Freeway and Highway Segments: Supplemental Ch. 27—Freeway Weaving: Supplemental Ch. 28—Freeway Merges and Diverges: Supplemental Ch. 29—Urban Street Facilities: Supplemental Ch. 30—Urban Street Segments: Supplemental Ch. 31—Signalized Intersections: Supplemental Ch. 32—Stop-Controlled Intersections: Supplemental Ch. 33—Roundabouts: Supplemental Ch. 34—Interchange Ramp Terminals: Supplemental Ch. 35—Active Traffic Management Technical Reference Library HCMAG—6 case studies

4. Operational Analysis

Operational analysis is one aspect of the overall transportation system's management process. This process starts with designing the facilities. The design process lays out the physical transportation network and describes the types of intersections or interchanges used to control the interaction of multiple links in the system. The process continues with the forecasting of travel demand in a transportation planning context, where analysts estimate how much traffic is expected on any given link and intersection.

This demand is then influenced by a host of factors that can promote or depress demand, or shift it over time and/or space. Traffic operations involves measuring or predicting the performance of this demand for the system, individual links and nodes, and ultimately the user of the system in the form of quantifiable measures of effectiveness (MOEs). This general process is illustrated in Figure below.

5. Capacity

Capacity is a reflection of the ability of a transportation facility to accommodate a moving stream of people or vehicles. Capacity of any highway system element is defined as the maximum number of vehicles that have a reasonable expectation of passing over a section (in either one or both directions) during a given time period under prevailing roadway and traffic conditions. In other words, capacity is the "supply" measure of transportation facilities. Capacity analysis provides tools for the analysis of existing facilities and for the planning and design of improved or future facilities.

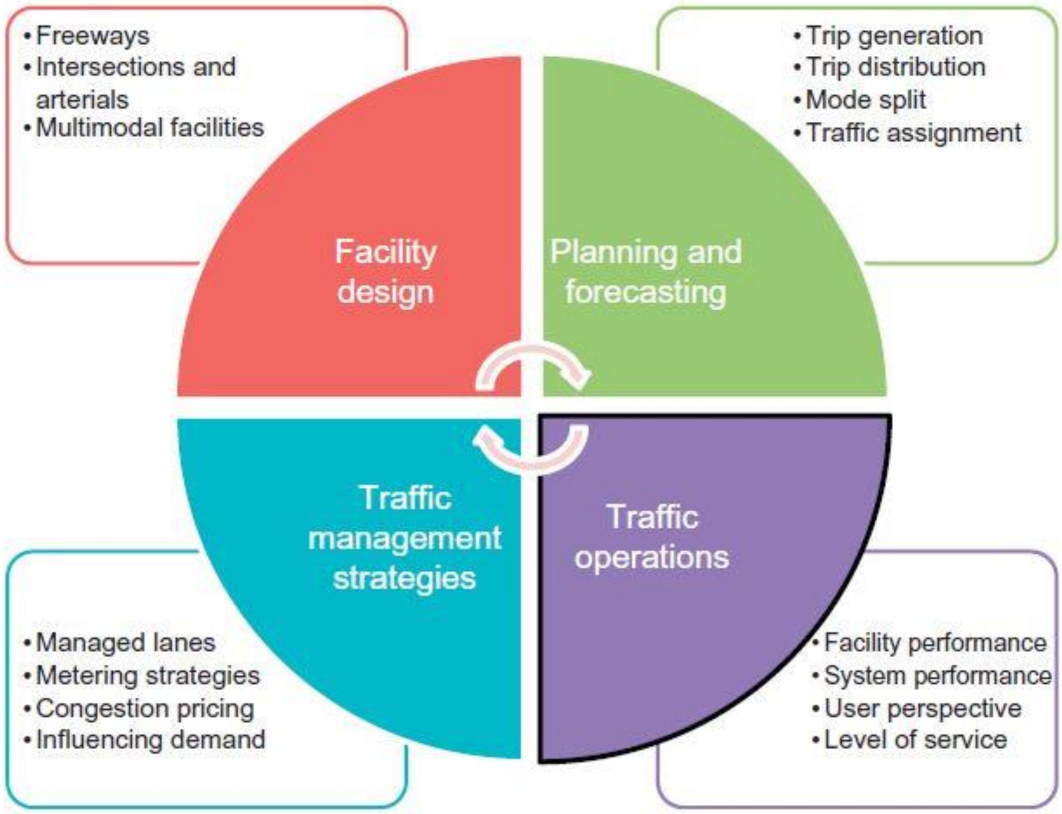


Figure: Traffic operations in the transportation systems management process.

In the Highway Capacity Manual (TRB), capacity is defined as follows:
Capacity is the maximum sustainable flow rate at which vehicles or persons reasonably can be expected to traverse a point or uniform segment of a lane or roadway during a specified time period under given roadway, geometric, traffic, environmental, and control conditions; usually expressed as vehicles per hour, passenger cars per hour, or persons per hour.
The capacity determines a key operating threshold for two traffic flow regimes, which are concepts used throughout traffic operational analysis. They are described as follows and illustrated in previous Table:

Undersaturation or uncongested flow is a traffic condition in which the arrival flow rate is lower than the capacity or the service flow rate at a point or uniform segment of a lane or roadway.
Oversaturation or congested flow is a traffic flow condition in which the arrival flow rate is greater than the capacity. Congested flow is often caused by a downstream bottleneck, which limits throughput, and results in queuing upstream of the bottleneck or choke point.
Typical capacities for highway systems are given in Table below.

<u>Table:</u> Typical capacities of highway systems.

Facility	Definition	Capacity in passenger cars
Freeways and expressways away from ramps and weaving sections	Per lane of freeway per hour	2400
Two-lane highway	Total in both directions, per hour	2800
Urban signalized intersection	Total per lane for through movement per hour of continuous green (urban areas with population over 250,000)	1900
Small town or rural signalized intersection	Total per lane for through movement per hour of continuous green	1750
Modern roundabout	Total per approach lane without any conflicting traffic in circle, depending on roundabout geometry and configuration	1400–1600

The capacities in Table above should be treated with care, as these represent the base or ideal conditions of the particular transportation system element. As such, the listed capacity values are rarely observed, as they are typically reduced by a range of factors that affect the capacity of roadways and intersections.

<u>These factors are broken down into three main categories</u>:

- Roadway conditions: These refer to the geometric characteristics of the street or highway including the type of facility, the surrounding development, the number of lanes, lane and shoulder widths, lateral clearances, design speed, and horizontal and vertical alignment.
- Traffic conditions: These refer to the characteristics of the vehicles using the facility. This includes the distribution of vehicle types, the amount of traffic in the available lanes, and the directional distribution of the traffic.
- Control conditions: These refer to the types and specific design of control devices and traffic regulations in use on a roadway or intersection. The location, type, and timing of traffic signals have a significant impact on capacity. Other important controls include stop and yield signs, lane restrictions, and turning restrictions.

Traffic engineers calculate the capacities of roadways and intersections by using the procedures presented in the Highway Capacity Manual (TRB), which are often implemented in software to facilitate what can be rather complex methodologies.
The volume-to-capacity (v/c) ratio identifies how close traffic volumes are to the calculated capacity. This unit less ratio is calculated by dividing the existing traffic volume by the calculated capacity. The value of the v/c ratio will never exceed 1. As it approaches 1, traffic congestion and delays are expected.

6. Level of Service

Level of service (LOS) is a measure of the quality of flow along a highway. In the HCM, there are six defined levels of service, designated LOS A (best operating conditions) through LOS F (worst operating conditions). Each of these levels of service represents a range of operational conditions within a traffic stream. These operational conditions are characterized by such factors as speed and travel time, freedom to maneuver, traffic interruptions, and comfort and convenience. For example, during the evening rush hour period for a particular stretch of highway, the users of that highway may experience LOS D; but a user traveling along that same exact roadway at 3 am would probably experience LOS A.

In applying and interpreting LOS, it is important to consider that it is a step function with thresholds determining the boundaries between different letter grades. These boundaries and thresholds have been set by the Transportation Research Board’s (TRB) Committee on Highway Capacity and Quality of Service, based on expert judgment and longstanding expertise in the application of the HCM methods. The thresholds, however, are fixed limits to describe a phenomenon that is highly variable due to day-to-day traffic fluctuations and estimation errors.
Analysts should take great care when applying LOS, as its step-function nature may push the transportation element under analysis into the next LOS category, despite an only marginal change in the underlying service measure. To illustrate this point, Figure below shows the LOS thresholds for signalized intersections, which are based on the average control delay per vehicle in seconds.
The figure illustrates that a 12-s change in the average control delay can lead to an LOS changing by two letter grades from A to C, by one letter grade from B to C, or can leave the LOS entirely unchanged for LOS C and above. As such, an analyst should always consider the net effect in the service measure (and other performance measures, as discussed in the next section) and be cautious whenever the numerical value of a service measure is very close to an LOS boundary.

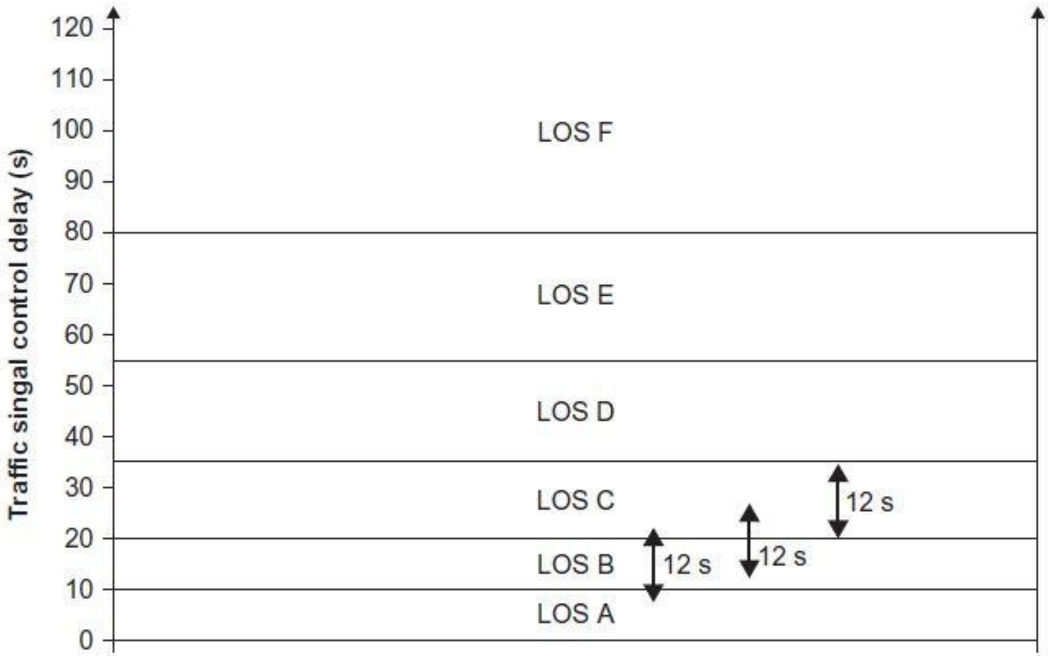


Figure: Illustration of LOS as a step function for traffic signals. Source: TRB, 2015.

7. Measure of Effectiveness and Service Measure

A measure of effectiveness (MOE), sometimes referred to as performance measure (PM), is a metric designed to describe and quantify the operations of a traffic systems element. Common MOEs include delay, travel time, speed, and number of stops, queue lengths, and density. Different MOEs are useful to quantify different aspects of the traffic operations of a system element, and often analysts look at more than one to get the full picture of how something operates. For example, the delay for vehicles at a signalized intersection is interesting and important, but a full picture of the intersection's operations is not possible without also looking at the number of stops (do drivers make it through with one or fewer stops, or are multiple stops—also known as cycle failures, as described later—observed in an approach?) or the queue lengths (are queues contained in the provided storage space, like a turn pocket, or do they spill onto other travel lanes?). Similarly, density is a useful measure for analysts to quantify a freeway system, but one could argue that other measures such as the average travel time or queue lengths are more directly tied to the user perception and "bottom line" for drivers traveling along a freeway facility.

The service measure is the property one looks at to determine what LOS a user is experiencing. The service measure varies depending on the type of facility that is being analyzed. For example, for freeway segments, the service measure is density. For signalized intersections and roundabouts, the service measure for vehicles is control delay, but for pedestrians it is an index that combines not only delay, but also aspects of comfort, safety, and convenience. Once again, the Highway Capacity Manual is a valuable reference that explains the methods for estimating capacity and determining LOS. The service measures for the 2010 HCM are summarized in Table below.

Table: Service measures by system element in the HCM.

Transportation element	Service measure by mode				System measure
	Auto	Pedestrian	Bicycle	Transit	
Freeway facility	Density	—	—	—	Speed
Basic freeway segment	Density	—	—	—	Speed
Multilane highway	Density	—	LOS score	—	Speed
Freeway weaving segment	Density	—	—	—	Speed
Merges and diverges	Density	—	—	—	Speed
Two-lane highway	PTSF[1], speed	—	LOS score	—	Speed
Urban street facility	Speed	LOS score	LOS score	LOS score	Speed
Urban street segment	Speed	LOS score	LOS score	LOS score	Speed
Signalized intersection	Delay	LOS score	LOS score	—	Delay
Two-way stop intersection	Delay	Delay	—	—	Delay
All-way stop intersection	Delay	—	—	—	Delay
Roundabout	Delay	—	—	—	Delay
Interchange ramp terminal	ETT[2]	—	—	—	Travel time
Alternative intersection	ETT[2]	—	—	—	Travel time
Off-street pedestrian/bike facility	—	Space	LOS score	—	Speed

PTSF = percent time spent following; ETT =experienced travel time (sum of delay and extra distance travel time)

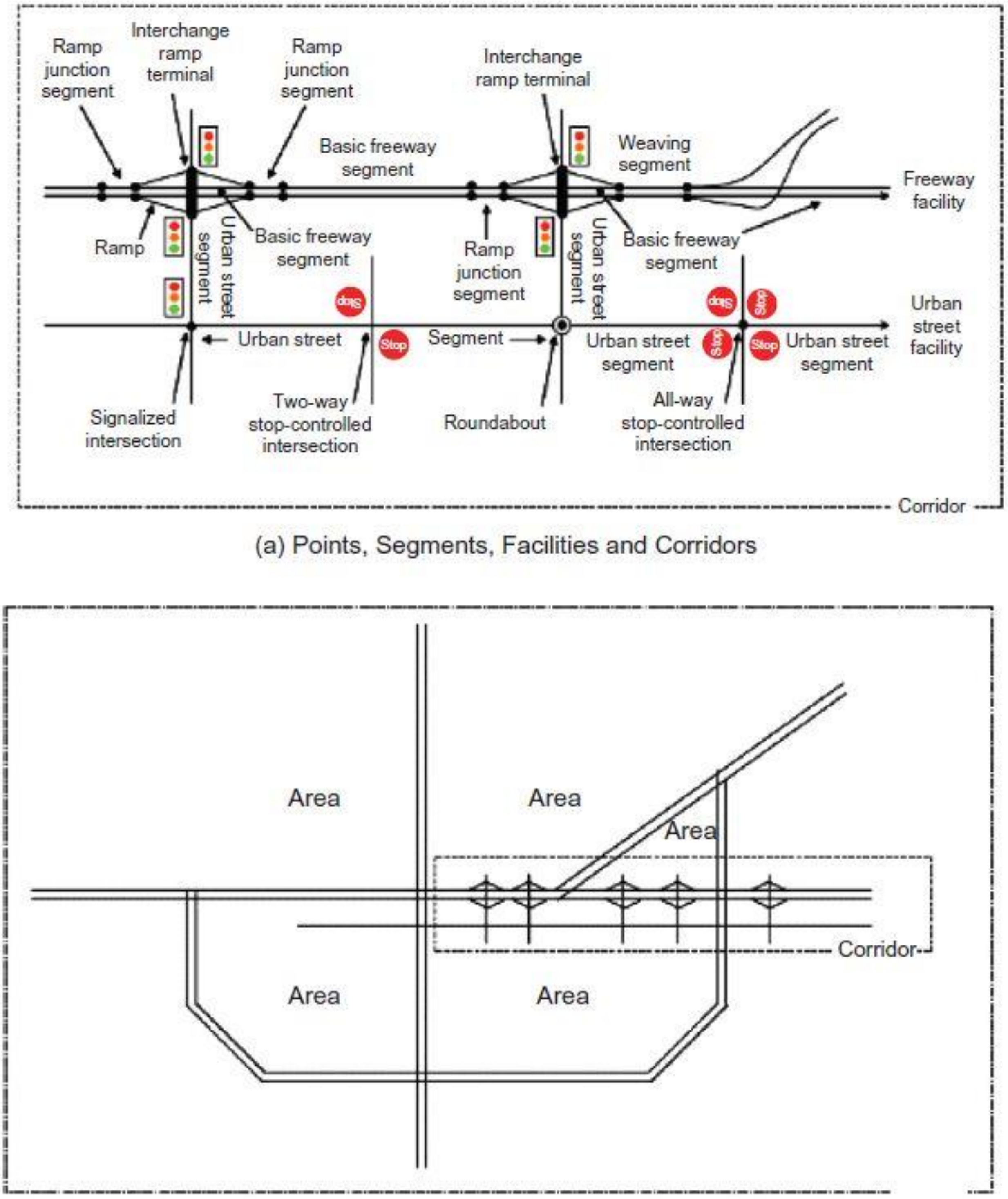
8. Spatial Analysis Scope

The HCM contains methods for varying spatial scopes. At the most disaggregated level, analyses are performed for individual point or nodes, which represent a single intersection. In fact, the HCM typically evaluates each approach to these intersection points separately, as traffic demands and geometric characteristics can easily differ from one leg of an intersection to the next. Points also include interchanges, which are the junctions between the surface streets and freeway networks.

The next aggregation level is the HCM segment, which for arterial streets represents the combination of an intersection and the link immediately upstream of it (in the direction of travel). Segments are also the lowest aggregation interval for freeways, where freeway, merge, diverge, and weaving segments are the basic building blocks. For both arterials and freeways, multiple segments can be aggregated to the level of facilities, which represent extended urban streets or sections of freeway that are of interest for analysis.

The facility is the highest level of analysis that can be readily performed in the HCM. However, other, more aggregated levels of spatial scope exist. For example, a system of two parallel facilities, an arterial and a freeway, are referred to as a corridor. A corridor analysis scope assumes that there is some interaction between the two facilities, for example, in the form of traffic diversion when an incident or general congestion impacts the operations on one or the other. From the corridor level, an area analysis may refer to the level of an entire neighborhood or suburb that combines multiple links, nodes, and segments. Finally, multiple areas make up the overall transportation system of, say, a metropolitan area.

The analysis of corridors, areas, and system typically requires the use of modern software tools that either integrate the methods from the HCM or use traffic simulation principles to estimate network performance, as discussed later in this part of the book. The various system elements are illustrated in Figure below.



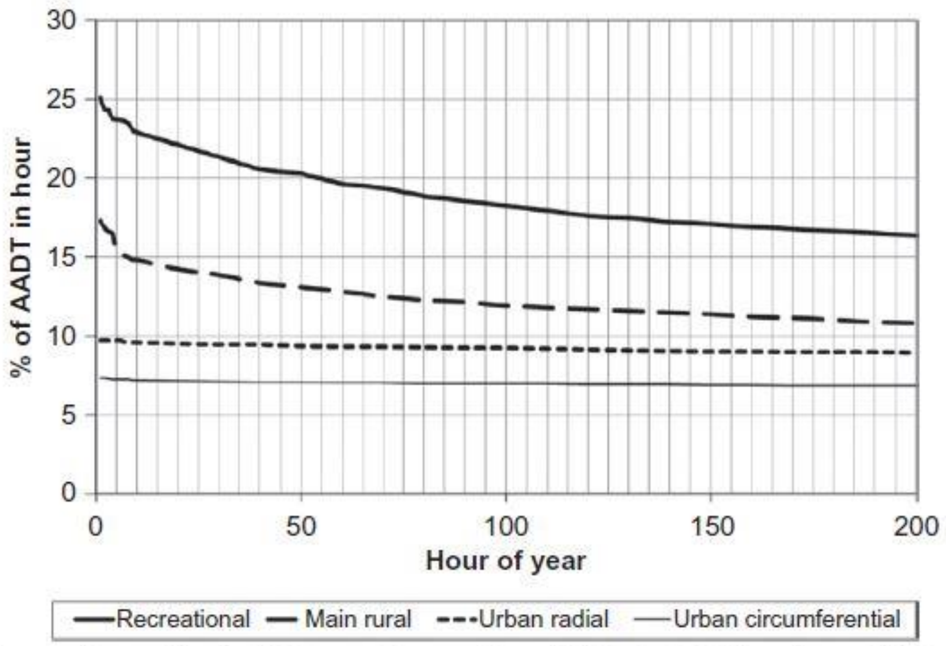
(a) Points, Segments, Facilities and Corridors

(b) Corridors, Areas and Systems

<u>Figure:</u> Illustration of spatial analysis scope and system elements. Source: TRB, 2015.

9. <u>Temporal Analysis Scope</u>

Equally important to the spatial scope is the temporal analysis scope. Traditional HCM analyses have focused on the analysis of the peak 15 min of the 30th highest hour of the year in terms of traffic demands. This 30th highest hour is also referred to as the design hour. The design hour approach acknowledges that it is likely inefficient and cost- prohibitive to base transportation analyses (and the resulting investment decisions) on the single highest hour of the year. In other words, the maximum hour of the year is by definition a very rare event, and has been found to not be reflective of “typical” congestion levels (Figure below).



Notes: Recreational, US-2 near Stevens Pass (AADT = 3,862); main rural, I-90 near Moses Lake (AADT = 10,533); urban radial, I-90 in Seattle (AADT = 120,173); urban circumferential, I-405 in Bellevue (AADT = 141,550).

Figure: Illustration of traffic volume by hour of year. Source: TRB, 2015.

In practice, analysts typically perform a traffic count on a typical weekday, which often avoids Monday and Friday, as well as any holidays. It is often cost-prohibitive to count traffic for multiple days or weeks at the same intersection, resulting in the general assumption that the measured peak period volume is a reasonable approximation of the design hourly volume (DHV).

However, the analyst should take great caution to select the observation period carefully, especially when evaluating a site with seasonal variability. For the purpose of analysis, the peak 15-min flow rate of the DHV is used traditionally as the volume input for operational analyses. The peak 15-min flow rate can be measured directly in the field (preferred), or can be estimated from the DHV using a peak hour factor (PHF). The PHF was described in detail in Part 2: Transportation Planning, and is defined as the peak hour volume divided by the product of four times the peak 15-min volume. The resulting factor is less than or equal to 1.0, and is used to convert the DHV to a peak 15-min flow rate, by diving the hourly volume by the PHF. This yields a flow rate that is likely greater than the hourly volume, and represents the 15-min peak.

10. Reliability Analysis

The traditional operational analysis scope in the HCM focuses on the operations of the peak 15-min flow rate of the 30th highest hour of the year, as described. This convention is often needed to manage limited resources for traffic counts in support of projects, and to obtain analysis data in a timely fashion. But by definition, this (simplified) approach of single-day counting ignores day-to-day variability of traffic. More importantly, it does not consider effects of nonrecurring sources of congestion that are known to impact the reliability of the system. The Federal Highway Administration (FHWA) defines seven common sources of unreliable travel as:

- Physical bottlenecks or capacity limitations.
- Weather impacts.
- Work zones.
- Traffic incidents.
- Traffic control devices, including signal timing.
- Special events.
- Day-to-day variability and fluctuations in normal traffic.

A sample of peak travel times for a route in Seattle, Washington is shown in Figure below. The figure illustrates that using any single estimate for the purpose of analysis fails to capture the true day-to-day variability in travel.

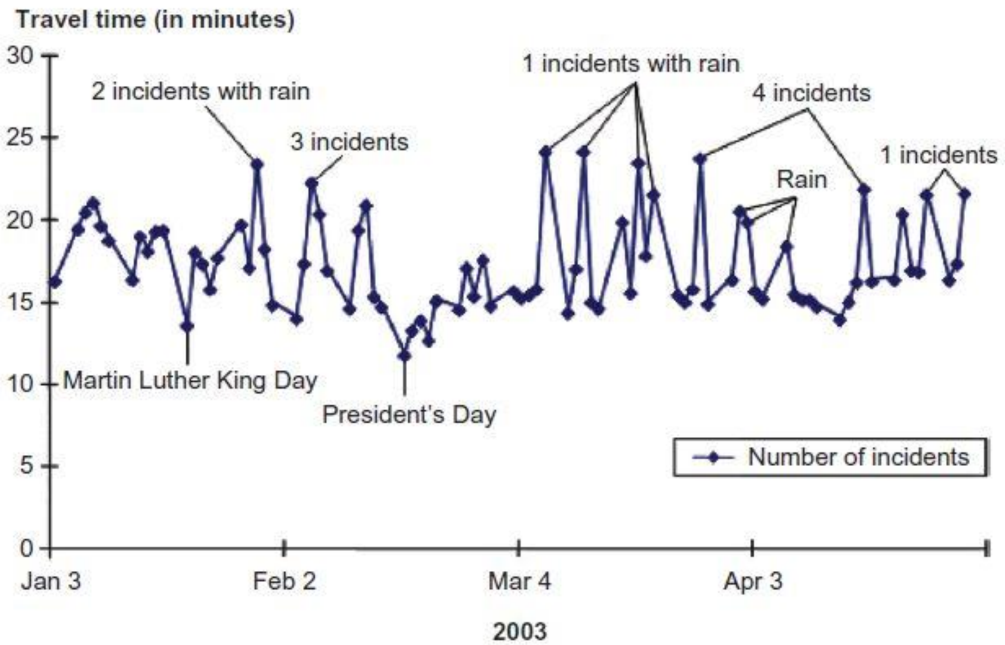
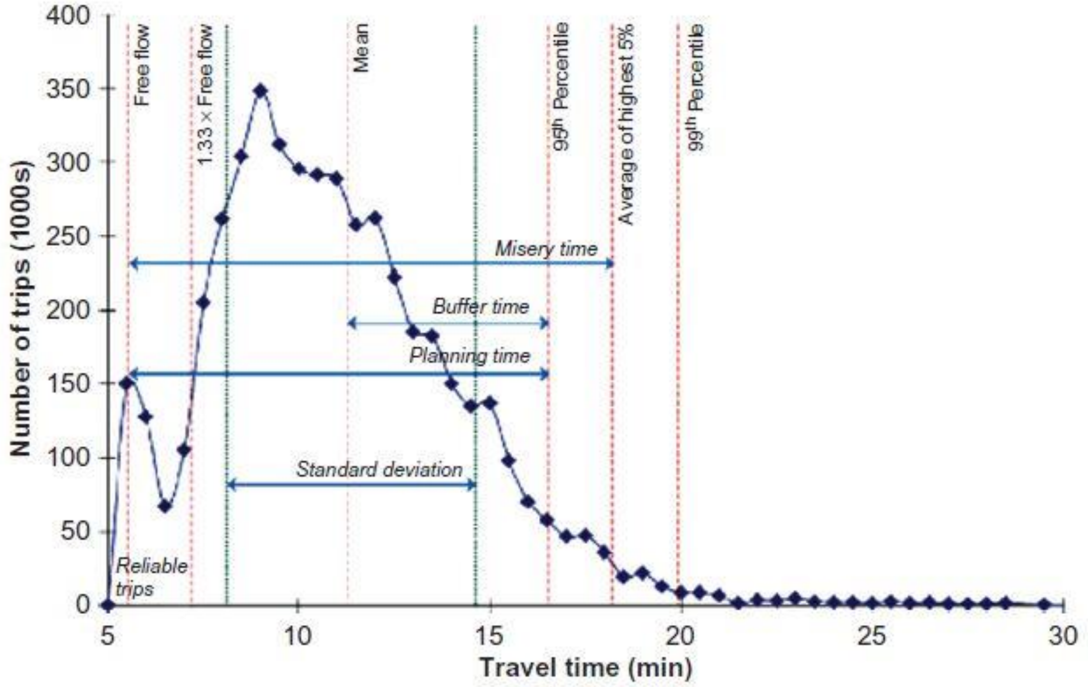


Figure: Example of day-to-day fluctuation in peak travel time. Source: FHWA, 2005a; http://www.ops.fhwa.dot.gov/congestion_report/executive_summary.htm.

Significant research in recent years has focused on defining performance measures and developing methodologies to quantify travel time reliability for traffic operational analysis. Travel time reliability is also integrated in the HCM as a key new concept in the latest edition, with the HCM now offering methodologies to estimate reliability for both surface streets and freeways.

At the heart of travel time reliability research and the HCM methodologies is the travel time distribution, which is the basis of all reliability performance measures. An example travel time distribution is shown in Figure below, which illustrates several key reliability performance measures, including the standard deviation of travel time, the 95th percentile of travel time, planning time (comparing 95th percentile to free-flow travel time), and buffer time (comparing 95th percentile to average travel time).



<u>Figure</u>: Example travel time reliability distribution. Source: Zegeer et al., 2014; http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2_S2-L08-RW-1.pdf.