2 DEVELOPING AND IMPLEMENTING A DATA COLLECTION PLAN

This chapter contains information and guidance on developing and implementing a data collection plan for travel time studies. Adequate planning, training, and preparation are vital to successful data collection activities. Figure 2-1 illustrates a generic travel time data collection process that can be used to plan and execute a travel time study.

The first several sections of this chapter describe the process of establishing study objectives and understanding the uses and users of the data being collected. Guidance is provided on setting the study scope, in terms of the geographic scale and inclusion of different time periods and facility types. The travel time data collection techniques are summarized and compared to assist in selecting the collection technique that is most appropriate. Data collection scheduling and data sampling are also discussed. The use of training and pilot studies are introduced as ways to improve the effectiveness and accuracy of data collection. The chapter concludes with general information about progress tracking, data reduction, and quality control.

2.1 Establish Study Purpose and Objectives

The study purpose and objectives establish the need for data and information in a transportation analysis and should be defined as the first step in any data collection activity. Once established, the study's purpose and objectives will help to guide the data collection to successful completion. Not only will the study's purpose and objectives be used to develop a data collection plan, they may also be used throughout the study process for clarification of tasks or resolution of ambiguous issues.

It is not uncommon for travel time data to be collected for several purposes with the main objective to establish a database of current roadway operating conditions. Similar steps should be taken to identify all required uses and ensure that the travel time data meet the minimum requirements (i.e., "smallest common denominator") for all applications. If the different studies or uses have competing needs, agency personnel may simply have to prioritize their data needs.

Examples of travel time study purpose or objective statements include the following:

> "The purpose of this study is to determine travel time information on major (Kansas City Metropolitan Region) streets and highways . . . The study will be used to . . . identify the extent and location of traffic congestion and specific problem areas . . ., serve as a data base to check speeds in the current computer networks. . ., allow comparisons with the 1987 and 1977 Travel Time and Delay studies. . ., and provide information . . . to determine areas where future studies are warranted. . ." (1);

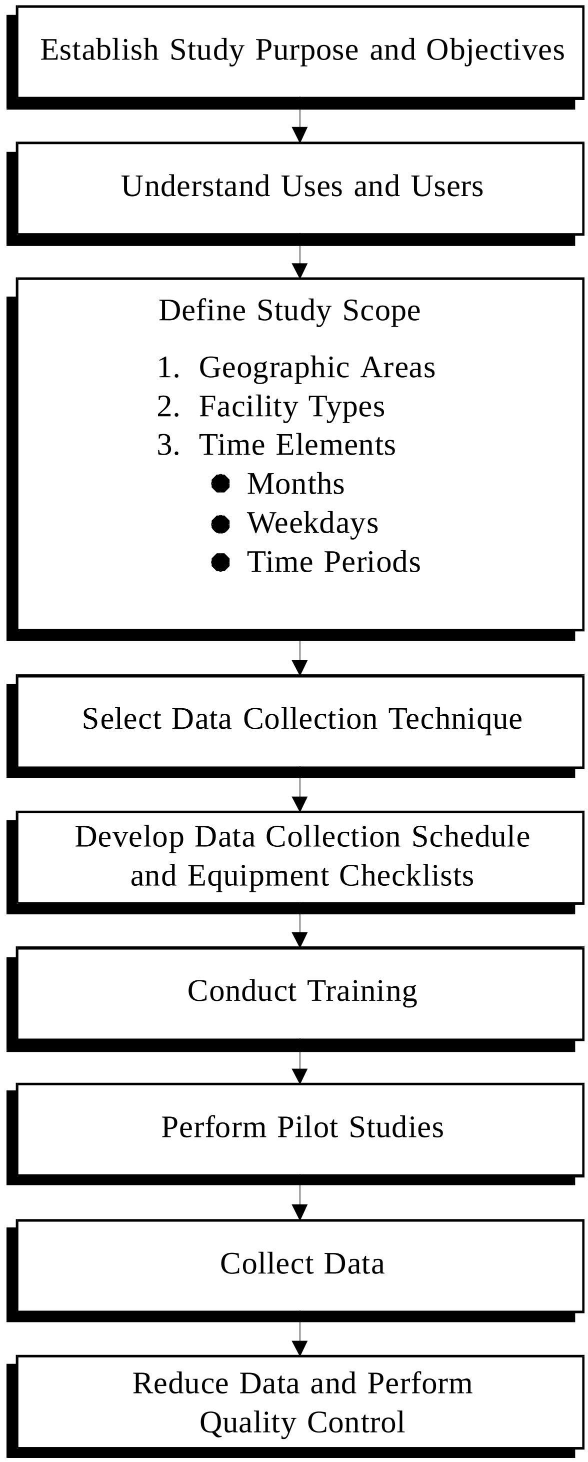


Figure 2-1. Travel Time Data Collection Process

> "The purpose of this study was to obtain effective travel times for representative links in the road network. These are used in computer models. . . In conjunction with the travel time runs, a vehicle delay study was also conducted to identify specific congested locations and also to determine the types and causes of delay . . ." (2)

2.2 Understand Uses and Users

The uses and users of the travel time data to be collected are as important as the study purpose and objectives. In many cases, the uses of the data are the motivation behind the study and should have been considered in establishing the study objectives. The users of the travel time data are an important consideration, as they affect several variables in the collection and presentation of data. Table 2-1 provides a perspective on the uses and users of travel time data. The table matrix illustrates the wide number of uses, and also the different uses for technical and non-technical audiences.

Travel time data often are collected for several purposes or potential uses. For example, travel time data might be collected for the congestion management process and also be used to calibrate travel demand forecasting models or as input to mobile source emissions models. For situations in which the travel time data must be used for several purposes, the data should be collected for the use that requires the finest level of detail. The travel time data can then be aggregated or analyzed to meet other study needs. In the earlier example, the mobile source emissions model may require second-by-second speeds to capture the acceleration and deceleration patterns in congestion. Once the second-by-second speed data has been collected, it can be aggregated for less data-intensive uses such as calibrating a travel demand forecasting model or monitoring area-wide congestion trends.

The emerging practice of using data collected by intelligent transportation system (ITS) applications for planning and evaluation purposes illustrates an important point about understanding uses and users of data. Until recently, ITS components were seen as providing valuable data only for operating transportation facilities. Several transportation agencies have recognized the many uses of ITS data for planning and evaluation applications and are beginning to share data resources where ITS applications have been deployed.

IMPORTANT ☞	Clear identification of study objectives, uses and users, and audience are a critical, yet often overlooked, step in the study design process.

Table 2-1. Uses and Users of Travel Time Data

Uses of Travel Time Data	Primary Users	
	Technical	Non-Technical
Planning and Design		
Develop transportation policies and programs		✔
Perform needs studies/assessments	✔	✔
Rank and prioritize transportation improvement projects for funding	✔	✔
Evaluate project-specific transportation improvement strategies	✔	
Input/calibration for air quality/mobile source emission models	✔	
Input/calibration for travel demand forecasting models	✔	
Calculate road user costs for economic analyses	✔	
Operations		
Develop historical travel time data base	✔	
Input/calibration for traffic models (traffic, emissions, fuel consumption)	✔	
Real-time freeway and arterial street traffic control	✔	
Route guidance and navigation	✔	✔
Traveler information		✔
Incident detection	✔	
Evaluation		
Congestion management system/performance measurement	✔	
Establish/monitor congestion trends (extent, intensity, duration, reliability)	✔	
Identify congested locations and bottlenecks	✔	✔
Measure effectiveness and benefits of improvements	✔	✔
Communicate information about transportation problems and solutions		✔
Research and development	✔	

2.3 Define Study Scope

A well-defined study scope that is clearly linked to the study objectives ensures that the travel time study will produce the necessary data. The study scope should answer three important questions:

1. Where do we collect travel time data? (Geographic Areas)
2. On what facilities do we collect travel time data? (Facility Types)
3. When do we collect travel time data? (Time Elements)

The study scope not only defines the ranges of effort during the travel time study, but also delineates the applicability of the results from data collection and analysis. Although this may appear to be a foregone conclusion, inadequate samples of data are often extended or extrapolated to make inaccurate conclusions about an entire population. Sampling procedures, however, can be used in travel time studies to collect statistically significant samples of data. Sampling procedures over both time and space are discussed later in this chapter and in subsequent chapters (for each technique).

2.3.1 Geographic Areas

The geographic scope defines the boundaries of the study. Examples of geographic scope include:

- A short section of roadway in the vicinity of a planned or implemented transportation improvement (e.g., before-and-after study);

- A transportation corridor between defined points, perhaps including a freeway, frontage roads, and parallel arterial street(s) (e.g., major investment study);

- Several transportation corridors that service a central business district or an activity center; and

- All major transportation corridors within a defined zone, sub-area, or region (e.g., congestion management system).

If a study's geographic scope only includes a selected number of corridors or roadways, travel time data would most likely be collected on each facility (i.e., no sampling). However, if the geographic scope encompasses an entire urban area or region, sampling procedures may be applied to achieve cost-effective data collection. Sampling procedures consist of collecting data for a statistically significant percentage of the entire roadway system being considered, then drawing conclusions about the entire roadway system from the sampled percentage. Sampling is most applicable for planning applications in which the required accuracy is typically less than that required for design or operational analyses.

2.3.2 Facility Types

The next step in defining the study scope is specifying the transportation facility types or functional classes of roadways. Like the geographic scope, the facility types considered in a travel time study should be based upon the study objectives. Facility types or classifications can be based upon different schemes, like those used in travel demand forecasting models, traffic operations models, or roadway inventory data bases. Table 2-2 contains examples of common roadway classifications from a variety of different sources. Collector and local streets are typically not considered in travel time studies because of their decreased functional role in providing mobility or throughput (3).

Table 2-2. Urban Roadway Functional Classification Categories

Travel Demand Forecasting Model (varies)	**Highway Performance Monitoring System (HPMS), Urban (4)**	**AASHTO 1994 "Green Book," Urban (3)**	**1994 Highway Capacity Manual (HCM), Urban and Suburban (5)**
Radial freeways Circumferential freeways Principal arterials (divided/undivided) Minor arterials (divided/undivided)	Interstate highways Other freeways and expressways Other principal arterials Minor arterials	Interstate highway Other freeways Other principal arterials Minor arterials	Freeways Multilane suburban highways Class I Arterial Class II Arterial Class III Arterial

The primary or ultimate use(s) of the travel time data will dictate the specific functional classification scheme to be used. For example, travel time data primarily collected for validating planning models will likely use classification categories corresponding to the specific travel demand forecasting model (i.e., similar to first column of Table 2-2). If the data will be used for other purposes, such as congestion management, it can be reclassified into categories that may be more appropriate for operational purposes, such as the HPMS or HCM classifications. Classification schemes used by other state, regional, or local agencies may also influence the choice of classification categories. In addition, the Federal Highway Administration (FHWA) has published additional guidance on highway functional classification (6).

The functional classification scheme should group roadways so that all roadways within a given group have similar traffic and operating characteristics. Grouping roadways with similar operating characteristics into a single classification strata or group permits the use of stratified sampling (if so desired). With stratified sampling, the number of roadway samples collected within the classification groups can be varied based upon the variability of data for roadways within a particular group.

A stratified facility sampling plan (i.e., 100 percent sampling of all facilities versus statistical sampling from each functional class) may also be considered in this step. Sampling of travel times on a regional network of freeways and arterial streets may be desirable where funds are not available for the desired data collection frequency or where you wish to concentrate data collection resources on the most critical routes of the network. The use of a sampling plan may also depend upon the application(s) of the travel time data. For example, regional system performance monitoring efforts may only require travel time data on a sample of freeways and arterial streets in the region. A corridor or before-and-after study, however, may require more detailed travel time data and necessitate data collection on all facilities under study (100 percent sample).

CAUTION !! Stratified sampling of data is typically considered when the desired precision can be achieved through sampling or when funds are not available for the complete roadway network. Proceed with caution and the assistance of a statistician.

The steps for establishing a stratified sampling plan are as follows:

1. **Establish the functional classification groups** to be used in the travel time study (Table 2-2). This step was discussed on the previous page.

2. **Designate the routes** that are located within the geographic scope and a functional classification group. This step consists simply of designating corridors that are within the study boundaries.

3. **Sub-divide routes into "segments,"** which are shorter sections of roadway (lengths vary by functional classification) with similar operating characteristics and geometric cross sections. The segment lengths may vary depending upon the data collection technique, but should be no longer than the following general ranges:

Freeways/Expressways:	1.6 to 4.8 km (1 to 3 mi)
Principal Arterials:	0.6 to 3.2 km (1 to 2 mi)
Minor Arterials:	0.8 to 3.2 km (½ to 2 mi)

Shorter segment lengths than these maximum lengths can be used for specific operational analyses with the caveat that segments lengths less than 0.4 to 0.8 km (¼ to ½-mile) may produce travel times with greater variability. Segment breakpoints, or route checkpoints, may be located at major interchanges, major signalized intersections, jurisdictional boundaries, and transition points between different roadway cross sections or land uses. For freeways and expressways, on-ramp merge points and lane drop locations are the best breakpoints for matching the cause of the traffic speed to the effects. For arterial streets, segment breakpoints are best located at major intersections or where changes in roadside activity occur. Professional judgment and local knowledge of traffic conditions should be used in defining segments. Site surveys or corridor reconnaissance during peak periods can also help in defining segment termini.

4. **Use sample size and finite population correction equations** (Table 2-3) to determine the number of roadway segments to sample within each functional classification group. The sample size calculations rely on three variables:

    - **Coefficient of variation (c.v.)** - a relative measure of variability, defined as the standard deviation divided by the mean. The c.v. can be estimated from existing data or the default values in Table 2-3 can be used for estimates of c.v. (7).

    - **Z-statistic** (or t-statistic for samples less than 30) - a function of the desired confidence level (e.g., 95 percent confidence level) for the sample mean (Table 2-3). The most commonly used confidence levels for stratified segment sampling are typically in the 80 to 95 percent range, but may also depend upon budget constraints.

    - **Relative permitted error** - expressed as a percentage, which is one-half of the desired confidence interval for the sample mean (e.g., ± 5 percent). The most commonly used error levels are between 5 and 10 percent, but vary depending upon the use of the travel time data (Table 2-3).

5. **Select the roadway segments to sample** within each functional classification (stratification) group. Theoretically, stratified random sampling techniques are used to randomly select the necessary sample size of roadway segments. The random selection of roadway segments scatters data collection sites around the geographic area, significantly increasing the costs of data collection for methods such as test vehicle. An alternative to random sampling is presented in the following paragraphs.

    Prioritized sampling, in which 10 to 20 percent of the critical or most congested segments are fully sampled, while the remaining 80 to 90 percent of the roadway segments are randomly sampled from different routes. Although prioritized sampling may not conform to the thorough statistical methods that exist for stratified random sampling, it concentrates data collection efforts on the most critical or congested locations. One or more of the following factors can be used for prioritizing data collection:

- perceived bottlenecks or congested locations;
- percent change in congestion level (if available);
- average daily traffic volume per lane; or
- average daily traffic volume.

These factors should rank the roadway segments with the highest congestion or the fastest growing congestion as “high priority.” Depending upon the number of roadway segments, the top 10 to 20 percent of segments could be designated as “high priority,” thereby collecting data on these segments on an annual or frequent basis.

Once the top 10 to 20 percent of roadway segments have been designated as “high priority” for data collection, the remaining roadway segments should be randomly chosen from the routes to accomplish the sample sizes for each strata group as outlined earlier. With this technique, data will be collected on the “high priority” segments and some randomly selected segments on a frequent basis (e.g., annual).

The prioritized sampling technique ensures that reliable, timely data exists for severely congested segments, and that the remaining, less critical segments are sampled on a less frequent basis. It will ensure that travel times at major bottlenecks such as lane drops, bridge/tunnel approaches, and freeway entrance locations are measured.

Table 2-3. Sample Size Estimation Equations

Coefficient of Variation, c.v.

$$c.v. = \frac{\sigma}{\mu} \approx \frac{s}{\bar{x}} \qquad (2\text{-}1)$$

where: σ = population standard deviation
μ = population mean
s = sample standard deviation
$\bar{x}$ = sample mean

Uncorrected Sample Size, n′

$$n' = \frac{c.v.^2 \times z^2}{e^2} \qquad (2\text{-}2)$$

where: z = z-statistic based on confidence level
e = relative permitted error level (%)

Sample Size, n (corrected for finite population):

$$n = \frac{n'}{1 + \frac{n'}{N}} \qquad (2\text{-}3)$$

where: n′ = uncorrected sample size
N = population size

Coefficient of Variation can be estimated from existing data or the following default values can be used for estimates of c.v. (7):

Freeways/Expressways: c.v.'s range from 15 to 25 percent (depending upon traffic volume)

Principal/Minor Arterials: c.v.'s range from 20 to 25 percent (depending upon traffic volume)

The **Z-statistic** is based on the desired confidence level. Z-statistics are provided below for commonly used confidence levels (8):

Desired Confidence Level	Z-statistic
99 percent	2.575
95 percent	1.960
90 percent	1.845
85 percent	1.440
80 percent	1.282

The **relative permitted error, e,** is expressed as a percentage of the sample mean, and is typically based upon the use of the data. Commonly used error levels are (7):

- ± 5 percent for design and operational analyses; and
- ± 10 percent for planning and programming studies.

2.3.3 Time Elements

There are several time elements that must be considered in establishing the scope for travel time data collection activities:

- months of the year;
- days of the week; and
- time periods, or time of day.

These three time elements are discussed in the following sections.

Months of the Year

Travel time data are commonly used to represent "typical" or "average" annual conditions, and should be collected during months that have typical or average traffic volume patterns. As with defining other time elements in the scope, traffic volume patterns from automatic traffic recorder (ATR) stations can be used to determine typical or average months for data collection. Table 2-4 contains data from an ATR station in Houston, Texas, and illustrates how this data can be used to define typical or average months. Those months with traffic volumes within approximately 2 percent of the annual average daily traffic (AADT) volumes are candidate months for data collection. As a general rule of thumb, the spring (i.e., March, April and May) and fall months (i.e., September, October, and November) are commonly considered average conditions if no ATR traffic volume data is available for specific areas or corridors.

For special studies that seek to examine congestion associated with non-work trips, one may wish to look at specific times of the year in which traffic patterns differ from typical or average months. Examples include (but are not limited to): summer or winter months near high use recreational areas; the holiday shopping season (late November and December) near large retail shopping centers; months when large universities or schools are not in session; or months coinciding with regional festivals or special events. However, if travel time data are desired for typical daily traffic conditions, these times of the year should be avoided.

Table 2-4. Using ATR Station Traffic Volume Data to Select Typical Months

ATR Station S-139, US 59, Houston, Texas				
Month and Season	**Average Day**		**Average Weekday (Mon - Fri)**	
	Volume	**Percent AADT**	**Volume**	**Percent AADT**
January	173,684	93.4	194,036	104.3
February	188,691	101.4	208,530	112.1
March	187,877	101.0	205,970	110.7
April	189,651	102.0	209,040	112.4
May	183,365	98.6	201,914	108.6
June	185,515	99.7	204,611	110.0
July	180,276	96.9	198,917	106.9
August	189,668	102.0	209,063	112.4
September	183,898	98.9	205,073	110.2
October	196,253	105.5	216,036	116.1
November	188,704	101.4	207,470	111.5
December	184,524	99.2	202,960	109.1
Annual Average	**186,009**	**100.0**	**205,302**	**110.4**

Source: adapted from reference (9).

Note: Shaded months ± 2 percent of annual average (e.g., candidates for data collection).

Days of the Week

Traditionally, data collection efforts for many transportation agencies have been focused on the middle weekdays (i.e., Tuesday, Wednesday, and Thursday). Monday and Friday are typically excluded from data collection because a small number of weekdays are sampled (typically less than 20 for most study budgets), and these days' high variation from conditions during the middle of the week would necessitate a much larger sample of weekdays.

ATR station data or other 24-hour traffic volume counts can also help to establish the day-to-day variation between weekdays. As a general rule, if study budgets only permit data collection to occur on a given facility less than five separate weekdays, then sampling should be concentrated between Tuesday and Thursday. If budgets permit sampling of five or more weekdays for a given facility, then data collection should be evenly distributed over all five weekdays (e.g., Monday through Friday). Weekend days (e.g., Saturday or Sunday) should be sampled if the study focus relates to recreational or weekend-based trips.

For example, Figure 2-2 shows morning peak hour data from a Houston freeway (I-10, Katy Freeway). The average speeds were collected throughout the peak hour by about 20 passive probe vehicles equipped with AVI transponders. The figure illustrates that Friday speeds are consistently higher than other weekdays and the average weekday speed. If a study budget only permitted the sampling of speed data from, say two to three weekdays per facility, then those weekdays that exhibit conditions closest to average "typical" conditions would be chosen. For this example, the study should sample weekdays between Monday and Thursday. If, however, the study budget permitted the sampling of speed data from five or more weekdays, the data collection should be distributed over all weekdays.

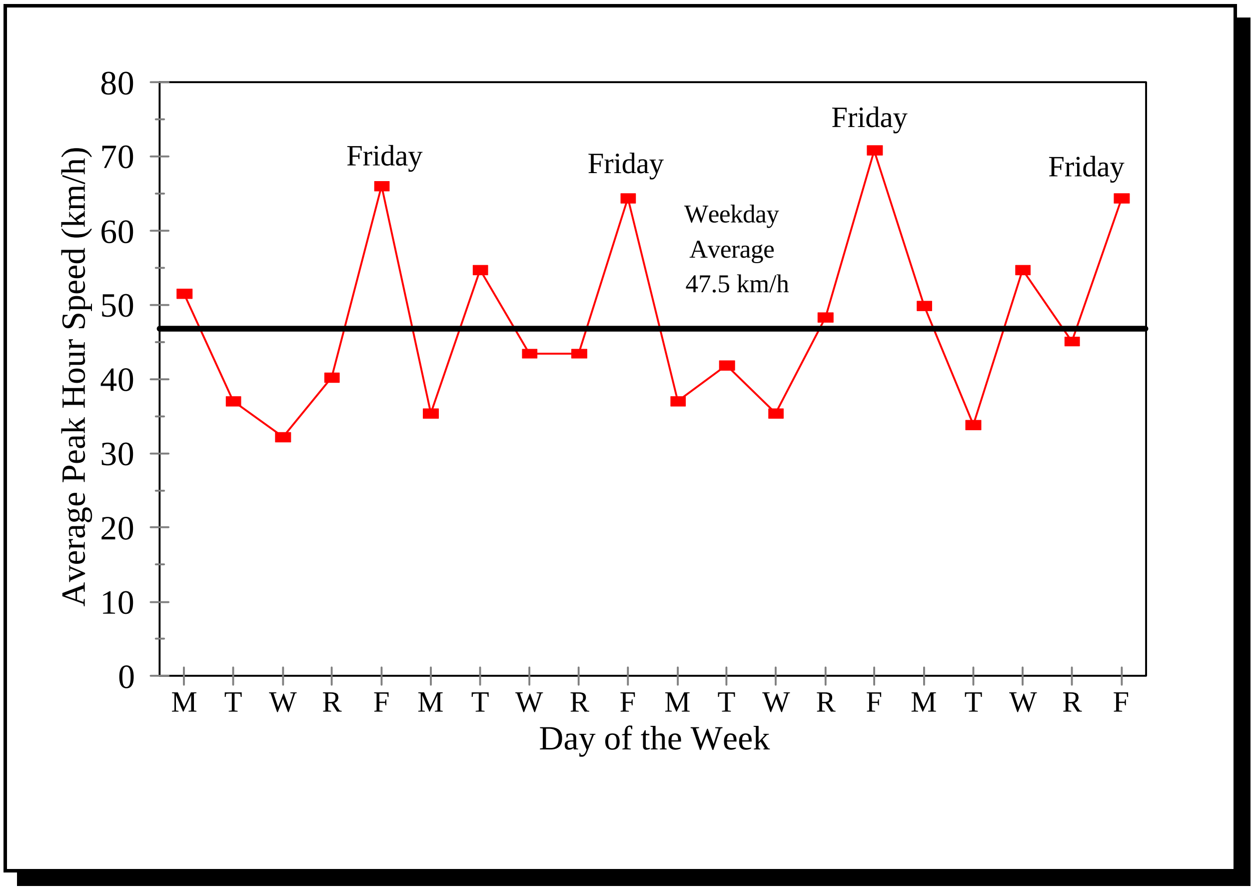
EXPERT TIP ☺

When budgets only permit data collection on less than five weekdays, concentrate on Tuesday through Thursday. If data can be collected for five or more weekdays, use all five weekdays to obtain average travel times.

Recurring holidays or events should be considered when scheduling the specific days for data collection. These days are avoided when sampling a small number of weekdays because of their variance from "typical" day-to-day operating conditions. Unless data is desired specifically for these events, the following times should be avoided when sampling weekdays :

- established holidays (e.g., Memorial Day, Independence Day, Veteran's Day);
- other celebrated days (e.g., St. Patrick's Day, Valentine's Day);
- changes in local school schedules (e.g., spring break, summer recess);
- day after time changes (e.g., Daylight Savings, Standard Time changes); and
- special events (e.g., professional sports games, regional festivals).

If large samples of weekdays (i.e., 75 to 100 percent of all possible days) are available, as is the case with some ITS data collection technologies, data from all days should be included to provide a truly representative value for the "average" weekday.



Source: data from reference (<u>10</u>), Appendix B

Figure 2-2. Illustration of Weekday Speed Variation (Morning Peak Hour Speeds, I-10 Freeway in Houston, Texas)

Time Periods

The time periods define the ranges in the time of day that travel time data will be collected. Like other elements of the study scope, the time periods will likely be determined by the study objectives. For travel time studies that are focused on identifying congestion trends and problems, three time periods are commonly considered:

1. **Morning Peak Period** - encompasses all congestion during the peak morning commute, typically sometime between the hours of 6 a.m. and 9 a.m.;

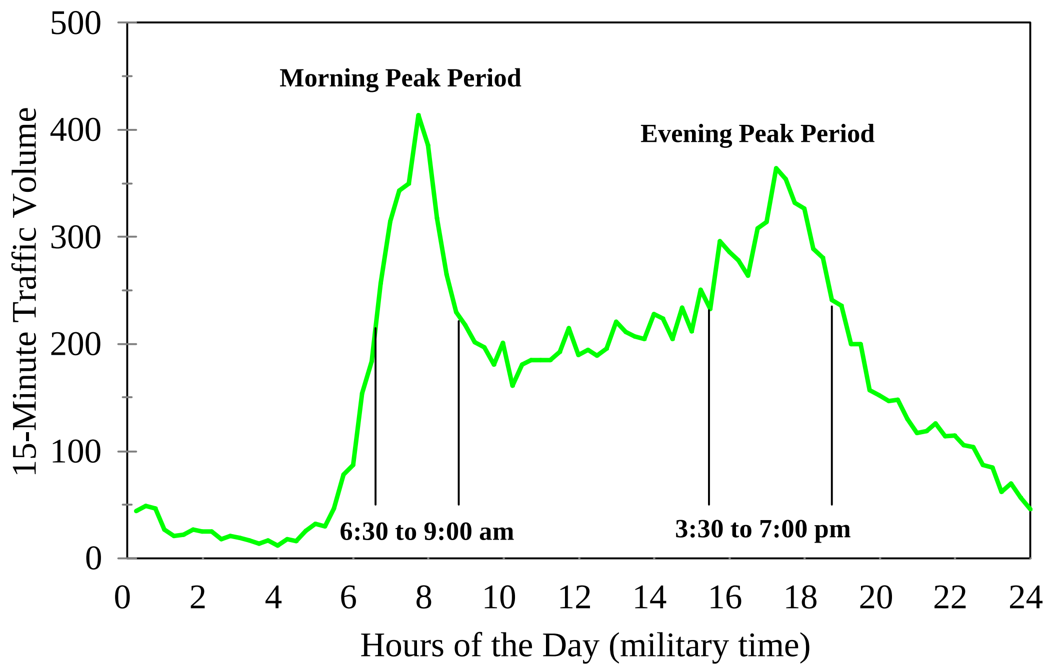
2. **Off-Peak Period** - includes periods of free-flow traffic during the middle of the day or late in the evening, typically between 10 a.m. and 11 a.m., 1 p.m. to 3 p.m., or after 7 p.m. The hours before and after 12 noon (11 a.m. to 1 p.m.) should be avoided if the "lunch hour" traffic is significant. Off-peak travel times are used to establish free-flow conditions for calculating congestion measures; and

3. **Evening Peak Period** - encompasses all congestion during the peak evening commute, typically some time interval between the hours of 4 p.m. and 7 p.m.

For studies relating to weekend or recreational travel, these typical "commuter" time periods should be adjusted to coincide with the times of congestion or peak traffic conditions.

The time periods for data collection should be matched to local traffic conditions and congestion patterns for the geographic area under consideration. The time periods can be defined by examining travel time data from previous studies or traffic volumes from inductance loop detectors, ATR stations, or 24-hour counts. The traffic volumes should come from a representative sample of facilities on which data is to be collected. On single corridors, traffic volumes taken at both end points and the middle of the corridor can better establish predominant congestion and traffic patterns throughout the corridor.

Figure 2-3 contains an illustration of defining time periods based upon traffic volume data. Note in the figure that the peak period includes both the build-up and dissipation of congestion, as evidenced by the peak volumes. The duration of the time period(s) depends upon the duration of congestion, which commonly varies by the population size of communities. In large metropolitan areas like Los Angeles, Houston, or New York, the peak periods may last two to three hours or more. In smaller towns and cities, congestion and the resultant peak period may last less than a single hour. The minimum duration of time for a peak period definition should be one hour (peak or "rush" hour).



Source: 15-minute loop detector data from TransGuide (San Antonio) web site "http://www.transguide.dot.state.tx.us/statistics.html"

Figure 2-3. Defining Peak and Off-Peak Time Periods Using Traffic Volumes

2.4 Select Data Collection Technique

Several data collection techniques are available to measure travel times. A specific technique, or combination of techniques, should only be selected after considering the study and data needs and the advantages and disadvantages of each technique. The travel time data collection techniques in this handbook are grouped into four general categories:

- "active" test vehicle techniques;
- license plate matching techniques;
- "passive" ITS probe vehicle techniques; and
- emerging and non-traditional techniques.

The first step in selecting a data collection technique should be to **investigate any existing sources of travel time or speed data**. For example, travel time or speed data may be available from other agencies or through transportation management centers. The deployment of ITS in major urban areas is a potentially rich source of travel time and speed data for a number of operational and planning studies. Because data from ITS components may only be available for a limited geographic area or corridor, additional data collection may be necessary to supplement existing data for area-wide or regional studies.

Once all existing sources of data have been identified, the second step is to **consider all needs and potential uses** for the travel time data. Some studies or analyses may require detailed travel time and delay information for specific corridors. Test vehicle techniques are most appropriate when analyses require detailed data about intermediate travel times and delay. In these cases, detailed information can be obtained by active test vehicles that traverse the routes or corridors of interest. License plate matching techniques are not well-suited for gathering intermediate travel time and delay, but do gather large sample sizes and provide more insight into travel time variability among drivers and throughout the time period. For example, license plate matching techniques have been used in several instances to compare the travel time reliability of general-purpose freeway lanes and high-occupancy vehicle (HOV) lanes.

The final consideration is the **budget and equipment resources allocated to data collection or available to the agency**. Available equipment (e.g., portable computers, video cameras) or study budgets may limit the data collection to one of several techniques. Some agencies may have analysis tools that are capable of exploiting certain data collection techniques. For example, agencies with geographic information systems (GIS) capabilities should consider the many advantages of GPS data collection. The chosen data collection technique should also match personnel capabilities and experience. Some test vehicle and license plate matching techniques are technology-intensive and require adaptable, experienced personnel that have available time and/or resources.

Table 2-5 contains a qualitative comparison of the travel time data collection techniques (Liu provides a complementary comparison in 11), and Table 2-6 summarizes the major advantages and disadvantages of each technique. These tables can be used to determine which technique best fits your data needs.

EXPERT TIP ☺

Steps in selecting a data collection technique: investigate existing potential data sources, assess potential applications and data needs, and determine budget and equipment allocation. Then refer to Tables 2-5 and 2-6.

Survey or interview methods are also used to obtain estimates of travel times for various purposes, including development and calibration of planning models. Most travel survey methods require drivers to record or recall travel times for trips during a given time period. This handbook does not include descriptions for survey recall methods, as the accuracy level is not consistent with other techniques in this handbook (12). However, useful travel time data may be extracted if travel survey methods use vehicles instrumented with GPS devices or similar instrumentation. In these cases, the appropriate sections in this handbook may be used based upon the vehicle instrumentation. Readers should refer to FHWA's *Travel Survey Manual* for detailed documentation of survey methods (13).

The following criteria are used in Table 2-5 to compare travel time data collection techniques:

- **Initial or capital costs** - typical costs of equipment necessary to perform data collection. For all ITS probe vehicle techniques except GPS, it is assumed that the vehicle-to-roadside communication infrastructure does not exist;

- **Operating efficiency** - relative costs of data collection per unit of data;

- **Required skill or knowledge level** - typical skill or knowledge level required for data collection personnel;

- **Data reduction and/or processing** - typical time and cost associated with reducing and/or processing field data;

- **Route flexibility** - transportability of data collection equipment to different routes in short periods of time;

- **Accuracy** - typical accuracy of the technique relative to the true average travel time (assumes adequate quality control procedures);

- **Sampling rate over time** - ability to collect travel time data at frequent time intervals (for a given facility) without excessive equipment;

- **Sampling rate over space** - ability to collect travel time data at closely-spaced distance intervals (for a given facility) without excessive equipment; and

- **Sampling rate of vehicles** - ability to collect travel time data that is representative of the numerous vehicle types and driving behaviors in the traffic stream.

Table 2-5. Qualitative Comparison of Travel Time Data Collection Techniques

Technique	Initial or Capital Costs	Operating Costs (per unit of data collected)	Required Skill or Knowledge Level	Data Reduction and/or Processing	Route Flexibility	Accuracy and Representa-tiveness[a]	Sampling Rate		
							Time	Space	Vehicles
Test Vehicle									
Manual	low	high	low	poor	excellent	fair	low	moderate	low
DMI	moderate	moderate	moderate	good	excellent	good	low	high	low
GPS	moderate	moderate	moderate	good	excellent	good	low	high	low
License Plate Matching									
Manual	low	high	low	poor	good	fair	low	low	moderate
Portable Computer	moderate	moderate	moderate	good	good	good	moderate	low	high
Video with Manual Transcription	low	moderate	moderate	fair	fair	excellent	high	low	high
Video with Character Recognition[b]	high	low	high	good	fair	excellent	high	low	high
ITS Probe Vehicle[c]									
Signpost-Based	high	moderate	high	good	poor	good	moderate	low	low
AVI	high	low	high	good	poor	excellent	high	low	moderate
Ground-based Radio Navigation	high	low	high	fair	good	good	moderate	moderate	moderate
GPS	moderate	low	high	fair	good	good	moderate	high	moderate
Cellular Phone Tracking	high	low	high	fair	good	good	high	moderate	moderate

Rating scales are relative among the techniques: [high, moderate, low] or [excellent, good, fair, poor].

Notes:
[a] Assumes that adequate quality control procedures are used.
[b] Assumes that necessary equipment is purchased (as opposed to contracting data collection services).
[c] Assumes that vehicle-to-roadside communication infrastructure does not exist.

Table 2-6. Advantages and Disadvantages of Travel Time Data Collection Techniques

Data Collection Technique	Advantages	Disadvantages
Test Vehicle		
Manual	• low initial cost • low required skill level	• high operating cost per unit of data • limited travel time/delay information available • limited sample of motorists
Electronic DMI	• moderate initial cost • very detailed speed/delay data available	• lacks geographical referencing (e.g., GIS) • limited sample of motorists
GPS	• moderate initial cost • data easily integrated into GIS • detailed speed/delay data available • can provide useful data for travel surveys	• reception problems in urban "canyons", trees • limited sample of motorists
License Plate Matching		
Manual	• low initial cost	• high operating cost per unit of data • accuracy may be questionable • data reduction time-consuming
Portable Computer	• low operating cost per unit of data • travel times from large sample of motorists • continuum of travel times during data collection	• accuracy problems with data collection, spurious matches • limited geographic coverage on single day
Video with Manual Transcription	• travel times from large sample of motorists • continuum of travel times during data collection	• data reduction time-consuming • limited geographic coverage on single day
Video with Character Recognition	• low operating cost per unit of data • travel times from large sample of motorists • continuum of travel times during data collection	• high initial costs (if equipment purchased) • limited geographic coverage on single day
ITS Probe Vehicle		
Signpost-based Transponders	• low operating cost per unit of data	• typically used for transit vehicles (includes loading/unloading times) • sample dependent on equipped vehicles
AVI Transponders	• low operating cost per unit of data • very accurate	• very high initial cost for AVI infrastructure • limited to instrumented locations • sample dependent on equipped vehicles
Ground-based Radio Navigation	• available consumer product	• typically used for transit vehicles • sample dependent on equipped vehicles
GPS	• increasingly available consumer product • low operating cost per unit of data	• sample dependent on equipped vehicles • privacy issues
Cellular Phone Tracking	• widely available consumer product	• accuracy questionable for detailed applications • privacy issues

The comparability of average travel times from different data collection techniques may be of concern if agencies compare data from several different sources. Little research has been performed to quantitatively compare average travel time samples using different techniques to the true average travel time. To some degree, each technique may yield slightly different values for the same conditions. Steps should be taken to avoid biases that affect the representativeness of travel time data from certain techniques. Analysts should also recognize and understand the source of potential biases in the data they are using, especially in cases where biases may be suspected. For example, test vehicle techniques may yield data with less variability than license plate matching data (because the test vehicle driver purposefully minimizing variability by "floating" with traffic). Or, probe vehicle data may be biased toward higher speeds because of the driving characteristics of select groups of motorists.

UNKNOWN **??**	Travel times collected using different data collection techniques may yield slightly different results. Data managers and users should recognize the potential biases inherent in the technique(s) they are using.

2.5 Develop Data Collection Schedule and Equipment Checklists

A schedule of data collection activities should be developed once the study scope, data collection technique, and other major parameters have been determined. An example of a data collection schedule for test vehicle runs is shown in Table 2-7. Similar tables showing date, time, and facilities being surveyed can be developed for other travel time collection techniques. A schedule is particularly helpful with implementing the data collection effort and in informing data collection personnel of their specific responsibilities. The content of the schedule includes the specific days and time of day that data is to be collected. If possible, the schedule should also contain the names of persons assigned to specific duties or stations for each day of data collection.

Equipment checklists should be used to ensure the proper assignment and continued operation of data collection equipment. These checklists are especially important for equipment-intensive travel time collection methods, such as computerized methods of test vehicle or license plate matching. A sample equipment checklist for video license plate matching is shown in Table 2-8. Checklists typically have columns to record the following information:

- time, date, and location of use;
- make, model, or serial identification number;
- names or initials of person(s) using equipment (check-in and check-out); and
- instructions for any necessary field calibration.

Table 2-7. Example of Data Collection Schedule for Test Vehicle Runs

Travel Time Runs--Week 1 (May 26, 1997)				
Day & Date	Time Period	Freeway Facility	VAN 3582 **HALF HR** Start by 3:30 pm Last run by 6:00 am	VAN 3583 **EVEN HR** Start by 4:00 pm Last run by 6:30 am
Monday, 5/27/97	PM	I-45	Bill S.	Jim R.
Tuesday, 5/28/97	AM	I-10	Dale T.	Sam P.
	PM	I-10	Bill S.	Jim R.
Wed., 5/29/97	AM	I-45	Dale T.	Sam P.
	PM	I-45	Bill S.	Jim R.
Thursday, 5/29/97	AM	US 59	Dale T.	Sam P.
	PM	US 59	Bill S.	Jim R.
Friday, 5/30/97	AM	I-45	Bill S.	Jim R.

Table 2-8. Example of Equipment Checklist

Equipment Checklist and Sign-Out

Route Number: Equipment Set Number:
Segment: Survey Team Number:
Location Description: Date: Time Period:

Equipment	**Person**	**Make**	**Model**	**ID/Serial #**
Hi-8 mm camcorder				
camera remote				
camera case				
CL-2x extender				
camera batteries				
camera remote batteries				
video leads (red and black)				
S-video cable				
linear polarizing filters				
UV lens filter				
DC-10 DC adaptor				
marine battery				
5" color monitor				
monitor cable				
monitor case				
Hi-8 mm video tapes				
tape labels				

Other Incidental Supplies:

☐ Watch ☐ Pen/Paper ☐ Safety Cones ☐ Cellular Phone ☐ Chains/Locks
☐ Phone Numbers ☐ Safety Vest ☐ Rain Gear ☐ Clipboard

Source: adapted from reference (14).

2.6 Conduct Training

The attitude and knowledge of data collection personnel play a major role in the quality of collected data. All data collection personnel should be adequately trained on the travel time data collection technique to ensure a consistent level of knowledge. The training or briefing may be best accomplished in small groups in which each person has the ability to ask questions and practice using the data collection equipment. A training session should include the following key points:

- purpose(s) of the data collection, including sponsorship, analysis goals, and end uses of the data;

- step-by-step details of the data collection technique and equipment operation;

- troubleshooting techniques to fix equipment problems in the field; and

- specific procedures or requirements for canceling data collection because of weather, traffic incidents, or equipment problems.

IMPORTANT ☞ Adequate training is necessary for a consistent level of quality data. Ensure that your data collection effort has training built into the budget.

Training sessions should also impart the following attitudes to data collection personnel (13):

- *This job is important* - stress the importance of the study, how it is to be used in solving problems and meeting project needs;

- *I must follow instructions* - teach the importance of following instructions, the necessity of proper field procedures, and the importance of accuracy and consistency;

- *I am a professional* - each person collecting data should believe: “I have a job to do; I am a professional being paid for services rendered”;

- *Research is important* - communicate the value of research, how research information improves our ability to make decisions, solve problems, and save money and resources; and

- *The accuracy and reliability of data is my responsibility* - stress that each person is responsible for collecting data that is accurate, reliable, and has been collected according to instructions provided.

2.7 Perform Pilot Studies or Trial Runs

Pilot travel time studies or trial runs should be conducted before the actual data collection begins. If the data collection personnel are experienced, pilot studies may be considered optional. Pilot studies can be performed over several days on a sample (approximately five to ten percent) of the facilities that will be included in the data collection effort. The purpose of pilot studies or trial runs are the following:

- become intimately familiar with the data collection equipment and process;
- become familiar with data collection corridors and cross streets;
- perform corridor or site surveys and measure exact distances; and
- identify problems or necessary resources as early as possible.

Also, travel time variability data obtained during pilot studies can potentially be used to check and/or adjust previously calculated sample sizes. After the pilot studies have been completed, all data collection personnel should provide feedback about the ease and utility of the data collection process. The feedback can then be used to modify the data collection procedures to ensure quality data.

2.8 Collect Data

Depending upon the scope of the study, data collection may extend through several months or even throughout the entire year. A manager of data collection activities should be assigned to track the progress of data collection, troubleshoot equipment and personnel problems, and supervise the data reduction and quality control measures.

The data collection supervisor should establish clear policies and procedures for canceling data collection in the field because of extreme or unusual conditions. Such extreme or unusual conditions that could merit field cancellation of data collection include:

- severe weather (e.g., heavy rain, tornados, ice);
- unusual traffic conditions (e.g., severe accidents, police chases); and
- equipment malfunction (e.g., dead batteries, broken video camera lens).

EXPERT TIP ☺

Define clear protocol for unusual circumstances prior to data collection. The use of cellular phones by data collection personnel in the field can save time and money.

Several other types of qualitative information should be gathered during data collection that could prove useful in the data reduction and analysis stages. Useful qualitative information includes:

- weather conditions (e.g., sunny, rain, foggy);
- pavement conditions (e.g., dry, wet, icy);
- observations about unique traffic conditions or incidents; and
- media reports about construction closures, incidents, or other special events that may affect traffic conditions.

Information that may be roadway or site-specific, such as weather or pavement conditions, should be recorded on data collection sheets or summaries. General area or regional information, such as special events, should be recorded in a common file location.

2.9 Reduce Data and Perform Quality Control

The first several days of travel time data should be reduced and analyzed soon after it has been collected to ensure that field personnel are collecting quality data. This early data reduction and quality control can potentially identify equipment problems or data discrepancies that are not obvious, particularly in electronic data collection systems. Data reduction or quality control records, such as those shown in Table 2-9, can also help to track progress.

Table 2-9. Example of Quality Control and Progress Tracking Forms

Morning Peak Period Travel Time Runs (7 to 9 a.m.)										
Freeway	Run 1		Run 2		Run 3		Run 4		Run 5	
	Driver	Q.C.	Driver	Q.C.	Driver	Q.C.	Driver	Q.C.	Driver	Q.C.
I-10 (Katy)	S.T.	R.B.	S.T.	R.B.	S.T.	R.B.				
I-45N (North)	B.E.	R.B.	B.E.	R.B.						
I-45S (Gulf)										
US 59N (Eastex)										
US 59S (Southwest)				Note: Responsible personnel can either initial or place an “X” in each cell once the run/QC has been completed.						
US 290 (Northwest)										

2.10 References for Chapter 2

1. "1990 Travel Time and Delay Study: Technical Memorandum." Draft, Mid-America Regional Council, Kansas City, Missouri, May 1991.

2. "Travel Time/Vehicle Delay Study, 1987." Preliminary, Association of Central Oklahoma Governments, Tulsa, Oklahoma, 1987.

3. American Association of State Highway and Transportation Officials, *A Policy on the Geometric Design of Highways and Streets*. AASHTO, Washington, DC, 1994.

4. U.S. Department of Transportation. *Highway Performance Monitoring System Field Manual for the Continuing Analytical and Statistical Database*. OMB NO. 2125-0028. Federal Highway Administration, August 1993.

5. *Highway Capacity Manual: Special Report 209*. Transportation Research Board, Washington, DC, 1994.

6. *Highway Functional Classification: Concepts, Criteria and Procedures*. Report No. FHWA-ED-90-006. Federal Highway Administration, U.S. Department of Transportation, Washington, DC, March 1989.

7. Lomax, T., S. Turner, G. Shunk, H.S. Levinson, R.H. Pratt, P.N. Bay, and G.B. Douglas. *Quantifying Congestion: User's Guide*. NCHRP Report 398. Transportation Research Board, Washington, DC, November 1997.

8. Devore, J.L. *Probability and Statistics for Engineering and the Sciences*. Third Edition, Brooks/Cole Publishing Company, Pacific Grove, California, 1991.

9. State Department of Highways and Public Transportation. *1988 Annual Report: Permanent Automatic Traffic Recorders*. Texas Department of Transportation, 1989.

10. Turner, S.M., G.A. Carlin and R.H. Henk. *Quantifying the Benefits of High-Occupancy Vehicle Facilities Using Automatic Vehicle Identification Technology*. Report No. SWUTC/95/465020-1, Southwest Region University Transportation Center, Texas Transportation Institute, November 1995.

11. Liu, T.K. "Travel Time Data Needs, Applications, and Data Collection." In *National Traffic Data Acquisition Conference (NATDAC '96) Proceedings, Volume II*. Report No. NM-NATDAC-96, Alliance for Transportation Research, Albuquerque, New Mexico, May 1996, pp. 402-418.

12. Murakami, E. and D.P. Wagner. "Comparison Between Computer-Assisted Self-Interviewing Using GPS with Retrospective Trip Reporting Using Telephone Interviews." Presented at the 8th Meeting of the International Association of Travel Behavior Research, Austin, Texas, September 1997.

13. Cambridge Systematics. *Travel Survey Manual*. U.S. Department of Transportation, U.S Environmental Protection Agency, Washington, DC, June 1996.

14. Liu, T.K. and M. Haines. *Travel Time Data Collection Field Tests* - *Lessons Learned*. Report No. FHWA-PL-96-010. Federal Highway Administration, U.S. Department of Transportation, Washington, DC, January 1996.

2.11 Additional Resources for Data Collection Planning

AASHTO Guidelines for Traffic Data Programs. American Association of State Highway and Transportation Officials, Washington, DC, 1992.

Ferlis, R.A., L.A. Bowman, and B.T. Cima. *Guide to Urban Traffic Volume Counting*. Report FHWA/PL/81/019. Federal Highway Administration, U.S. Department of Transportation, Washington, DC, September 1981.

"Guide for the Data Quality Objectives Process," Report EPA QA/G-4. United States Environmental Protection Agency, Quality Assurance Management Staff, Washington, DC, September 1994.

Traffic Monitoring Guide: Third Edition. Office of Highway Information Management, Federal Highway Administration, U.S. Department of Transportation, Washington, DC, February 1995.