Traffic Data Quality Workshop

Work Order Number BAT-02-006

STATE OF THE PRACTICE FOR TRAFFIC DATA QUALITY

White Paper

Prepared for

Office of Policy
Federal Highway Administration
Washington, DC

Prepared by

Battelle
. . . Putting Technology To Work

Texas Transportation Institute

Cambridge Systematics, Inc.

December 31, 2002

"State of the Practice for Traffic Data Quality"

By Rich Margiotta

Introduction

Purpose of Report

This White Paper documents the current state of the practice in the quality of traffic data generated by Intelligent Transportation Systems (ITS). The current state of the practice is viewed from the perspectives of both Operations and Planning personnel; the distinction between these two groups is that Operations personnel use the data primarily for real-time or near real-time applications (e.g., incident management, ramp metering) while Planning personnel use the data for applications that are not nearly as time sensitive (e.g., monitoring trends in travel monitoring). The paper considers:

- What Operations and Planning applications use traffic data and what are the quality requirements for these applications.
- Causes of poor quality in traffic data
- Quality issues specific to ITS-generated traffic data
- Possible solutions to quality problems

For the purpose of this paper, when "Operations" or "ITS" is used, it is meant to refer to the activities of ***Traffic Management Centers (TMCs) in urban areas***. Rural ITS applications are emerging, but the current state of the practice in ITS-generated traffic data is clearly focused on urban TMC deployments.

Methodology

This report draws heavily on past work conducted for FHWA under the Archived Data User Service (ADUS) program. Additional information was gathered from phone interviews with state transportation agency personnel from traffic monitoring programs (usually within Planning divisions) as well as ITS groups. (ITS personnel were usually those directly involved in traffic management center (TMC) operation.)

Types and Applications for Traffic Data

Data Types

Several types of traffic data are collected by both "traditional" and ITS means. Table 1 displays these types of data. Where there is overlap between the two realms, the basic nature and definitions of the data collected are the same. However, there are subtle differences in data collection methodologies that may lead to problems with data sharing and quality. Among these are the polling rate and vehicle classification "bins". (Section 4 discusses these discrepancies in more depth.)

Table 1. Types of Traffic Data Used by Transportation Agencies

Data Type	Description	Collection Details
Volume	Total number of vehicles passing a point on the highway over a given time interval	**Planning:** Collected continuously at a limited number of sites statewide; 24-48 hour counts cover most highway segments (but counts may be up to 3 years old on major highways, more on lower classes); data usually aggregated to hours for reporting from field. **ITS:** Collected continuously on every segment (1/2 mile spacing is typical on urban freeways); data reported at 20-30 second intervals from field; data aggregated for later use anywhere from 20-30 seconds up to 15 minutes.
Vehicle Classification	Same as volume except counts are made by individual vehicle classification	**Planning:** Collected continuously at a limited number of sites statewide; 24-48 hour counts taken at selected locations; FHWA 13-bin scheme based on number axles, type of power unit, and trailering is the most common. **ITS:** For urban TMCs, it is uncommon that vehicle classification is collected – where it is, 3-4 length-based bins are typically used. (CVO deployments used primarily to capture intercity truck movements do collect vehicle classification.)
Truck Weight	Total weight and individual axle weights and spacings of trucks	**Planning:** Same as vehicle classification except that short-counts are less frequent. **ITS:** For Urban TMCs, neither collected by ITS deployments nor used in ITS applications. (CVO deployments used primarily to capture intercity truck movements do collect vehicle weights.)
Occupancy	The percent of time that a roadway detection zone is "occupied" with vehicles	**Planning:** Not collected. **ITS:** Collected continuously on every segment (1/2 mile spacing is typical on urban freeways); data reported at 20-30 second intervals from field; data aggregated for later use anywhere from 20-30 seconds up to 15 minutes. (The same equipment is used for both volume and occupancy measurements.) ***Roadway density*** and ***average headways*** can be calculated from occupancy if length of the detection zone and average vehicle length are known.
Speed	Speed of vehicles passing a point on the highway over a given time interval (also known as "time-mean speed")	**Planning:** Newer equipment used to measure volumes, vehicle classifications, and truck weights are capable of collecting speeds, but the data are rarely used. **ITS:** Either collected directly (same characteristics as for volume and occupancy) or estimated from volume and occupancy measurements (older "single roadway loop" systems).
Travel Time	The measured time a vehicle takes to traverse a highway	**Planning:** Rare for state agencies to collect; local agencies collect using "floating car" method (drivers specifically tasked to collect travel times). License plate matching using

Data Type	Description	Collection Details
	segment	imaging technology becoming more prevalent. **ITS:** Collected with vehicle-based technologies: (1) GPS transmission of location and time, or (2) roadway-based "readers" of vehicle tags. (Most of the vehicle "tags" in current use are from automated toll collection systems. Readers may also be installed off of toll highways to detect the passage of "tagged" vehicles.)
Queues	Stopped or slow moving vehicles impeded by a bottleneck	**Planning:** Not usually collected. **ITS:** Where collected, restricted to queues at ramp meters.

Applications: Planning-Related Traffic Monitoring

Planning-related traffic monitoring activities are usually conducted as a service to support a variety of other functions with transportation agencies. Brief examinations of the Planning applications that use traffic data are presented in Table 2. Also included in Table 2 is an assessment of the advantages of using ITS-generated traffic data for these applications. It is clear that ITS-generated data ***potentially*** offers many advantages over general use traffic data:

- **The continuous nature and detailed geographic coverage of traffic data generated by ITS removes temporal sampling bias from traffic measurements.** The vast majority of traffic data currently collected for planning, administration, and research applications are based on short-duration traffic counts. Although attempts are made to adjust or expand the sample, the procedures are imperfect. With continuous data, there is no need to perform adjustments to control sample bias. (Equipment-based errors are still present, though).

- **Continuous data from ITS sources allows the direct study of variability in travel times.** This variability is often termed the ***reliability of travel times*** and it is becoming an important factor in both the operations and planning communities. Continuous data also capture the full range of factors influencing reliability, most notably incidents and weather – short duration counts either completely miss these events or are unduly biased by them. (Many agencies will discard short counts and floating car runs taken during "unusual" events.)

- **ITS-generated traffic data can supplement – and in some cases supplant – traffic data collected for Planning and general use.** Traffic monitoring on heavily traveled urban highways has become extremely difficult for field personnel. Installing portable devices on the mainlines of these highways has become practically impossible for safety reasons, and the reliance on ramp-based methods requires that multiple devices be installed and that all devices be operating properly during the data collection. By accessing data that already exist through ITS sources, these problems are avoided.

Recent work indicates that ITS data can be used as volume resource in these circumstances.

- **Data to meet emerging requirements and for input to new modeling procedures will have to be more detailed than what is now collected.** The next generation of Travel Demand Forecasting (TDF) models (e.g., TRANSIMS) and air quality models (modal emission models) will operate at a much higher level of granularity than existing models. Traditional data sources are barely adequate for existing models and there is little doubt that they will be incapable of supporting the next generation of models. ITS can provide many of the data types to support these models, especially at the detailed geographic and temporal resolutions that are required. For example, roadway surveillance data (volumes, speeds, and occupancies) are typically reported every 20 seconds and GPS-instrumented vehicles can report positions and activity at time intervals as short as one second. Also, GPS-derived locations can pinpoint incident locations to within a few meters. This level of detail will be required for the input and calibration data used by the new models. Finally, as data generated by ITS are used more frequently for non real-time purposes, it is likely that additional uses not currently foreseen will emerge. In addition, data on activity patterns and how travelers respond to system conditions will be important for the next generation of models.

- **As the focus of transportation policy shifts away from large-scale, long-range capital improvements and toward better management of existing facilities, the creation and use of system performance measures is taking on greater significance.** Measures of mobility have been used for many purposes, ranging from site-specific operations analysis to corridor-level alternative investments analysis to area-wide planning and public information studies. Transportation agencies have adapted a wide range of mobility performance measures and these have been reviewed to develop the performance measures most appropriate for national mobility monitoring. In the past few years, the issue of performance monitoring has been elevated by transportation agencies to be responsive to the demands of the public and state legislatures, and TEA-21's emphasis on system operations and management have extended this trend. The demands of performance monitoring are more rigorous than traditional planning applications, which are geared to estimating investment requirements to the "nearest extra lane of capacity." In other words, data with the gross resolution to meet traditional transportation planning applications will be incapable of detecting more subtle changes in system performance.

- **ITS technologies have the potential to capture urban vehicle classifications, a large gap in the current traffic data programs.** Nearly all of the equipment used by Planning-oriented traffic monitoring units to perform automatic vehicle classification is based on devices placed on or in the roadway surface. The current state of this equipment does not allow for accurate vehicle classification where vehicle speeds are variable, as in congested urban areas. Emerging technologies used for ITS-related traffic monitoring have demonstrated potential for collecting vehicle classification in addition to the typical "suite" of volumes, speeds, and occupancies. Although the classifications from this equipment are length-based (3-4 bins are common) and therefore not as detailed as data from Planning-oriented monitoring activities, they nonetheless can fill a large void.

Applications: Operations

In urban areas, Operational responses originate at TMCs whose primary focus is freeway performance. Roadway surveillance is a typical feature of TMCs, both in terms of visual coverage (e.g., CCTV) and electronic traffic data. Electronic traffic data always include volumes and detector zone occupancies and most TMCs also include measured traffic speeds. (The same equipment is used to measure all three data types.) Current TMC applications ***that potentially can use traffic data include:***

- Ramp meter control – most algorithms for dynamically adjusting ramp metering rates are based on occupancies.
- Lane control – speeds caused by bottlenecks are used to provide lane control guidance.
- Traffic signal control – real-time traffic adaptive control strategies (e.g., SCOOT, SCATS) rely on detailed information about signal performance and mid-block speeds.
- Incident detection – incident detection algorithms use speeds, occupancies, or some combination.[1]
- Variable speed limits – adjusting speed limits based on current environmental and traffic conditions.
- Evacuation, special event, and military deployment – these functions usually have special traffic control needs.
- General bottleneck performance – speeds are used by TMC personnel to gain a general understanding of real-time system performance.
- Traveler information – maps showing current speeds by link are a typical form of information disseminated by TMCs. Also, messages of general congestion (based on speeds) and specific incidents are often posted on dynamic message signs and broadcast over highway advisory radio.
- Evaluations and Performance Monitoring – where these are conducted, volumes and speeds are used.

[1] Experience with incident detection algorithms has been mixed. Many areas have found that algorithms produce too many “false alarms” and no longer rely on them. Other areas still use them as a screening mechanism. In general, incident detection can be efficiently performed by fielding cell phone calls from motorists, especially if a dedicated number for reporting incidents exist.

Table 2. Traditional Applications for Traffic Data

Category	Specific Application	Current Traffic Data Used	Advantages of Using ITS-Generated Data
Travel Demand Forecasting Models	Validation of predicted link volumes	AADTs for 24-hour forecasts (generally used in smaller areas); peak hour volumes in larger areas	Continuous data removes sampling and adjustment bias present in short counts and in developing peak hour volumes from K- and D-factors.
	Validation of predicted link speeds	None available for this purpose	Can be derived directly from measured data for either daily or peak hour.
	Free flow speeds	None available for this purpose; based on speed limit or judgment	Can be derived directly from measured data.
	Link capacities	None available for this purpose; based on judgment and (rarely) *HCM* analysis	Direct measurement of highest flow rates based on actual link conditions.
	Link truck percentages	Based on limited amount of urban vehicle classification	New technologies can provide much better estimates of urban vehicle classification (length-based, continuous, greater coverage).
Congestion Management Systems	Performance measures (mobility-based)	Limited floating car data; synthetic methods based on volume estimates	Direct measurement of long-term performance and speeds, including the effects of incidents, weather, work zones, and other sources of non-recurring congestion missed with synthetic methods.
Emissions Models (MOBILE6)	Hourly speed estimates by functional class	Synthetic methods based on volume estimates	
	VMT by 28 vehicle classes	Based on limited amount of urban vehicle classification and vehicle registrations	Length-based classifications can be a basis for developing these.
Highway Design	Design volumes	Estimated using forecasted AADTs with areawide K-, and D-factors	Facility-specific K- and D-factors can be derived.
Safety Analysis	Crash rates for performance monitoring and specific studies	Exposure (typically VMT) derived from short-duration traffic and vehicle classification counts; traffic conditions under which crashes occurred must be inferred.	Continuous volume counts, truck percents, and speeds, leading to improved exposure estimation and measurement of the actual traffic conditions for crash studies.
Freight Analysis	Truck travel patterns	Data collected through rare special surveys or implied from available vehicle classification	Electronic credentialing, AVI, and new roadway technologies for vehicle classification allows tracking. Improved understanding of truck patterns and can lead to improved assessments of inter-modal

Category	Specific Application	Current Traffic Data Used	Advantages of Using ITS-Generated Data
			access and highway design for heavily used truck highways.
Pavement and Bridge Management	Historical and forecasted loadings	Volumes, vehicle classifications, and vehicle weights derived from short-duration counts (limited number of continuously operating sites)	Continuous volume counts and vehicle classifications taken over a larger area.

- Weather Management – includes detecting and forecasting weather-related hazards such as snowy/icy road conditions, dense fog, high winds, and approaching severe weather fronts. This knowledge can be used to more effectively deploy road maintenance resources. It can also be used in conjunction with other core functions such as traffic control (e.g., variable speed limits, signal coordination timings), incident management (e.g., routing response vehicles), and traveler information (e.g., general advisories, location specific warnings).

Traffic Data Quality: Characteristics

What Causes "Bad" Traffic Data

Several sources contribute to inaccuracies in traffic data. These relate to the nuances of specific equipment and how data are collected and transmitted from the field. A more thorough discussion of data quality issues associated with particular technologies is covered in the white paper, *Innovative Approaches to Traffic Data Quality*. A few generalizations can be made about the sources of data quality problems:

- **Type of equipment.** Roadway-based devices (inductive loops are the most common) are placed in each lane of traffic. Non-intrusive devices (such as radar, acoustic, and video imaging) are usually configured as one device per direction of travel. That is, a single device measures all lanes of traffic in a direction. All devices establish a detection zone within which measurements are taken, but the methods of how conditions are determined are each different from the others. Recent tests by the Minnesota Department of Transportation reveal that volume performance at the freeway test site revealed that most non-intrusive sensors had an absolute error of between 2 percent and 10 percent when mounted within vendor-recommended ranges. Also, all of the sensors were within 8 percent of the baseline speed data.

- **Interference from environmental conditions.** Roadway surface conditions can affect the performance of equipment installed in the pavement. Precipitation and light conditions can affect the sensing abilities of non-intrusive devices.

- **Installation.** Roadway-based equipment is sensitive to how it is placed in the pavement. Non-intrusive devices must be placed in such a manner that detection zones in all lanes can be established. Further, installation of non-intrusive devices on the roadside creates an "occlusion" problem – vehicles (especially trucks) can block the detection zones of some lanes. The problem increases with the number of lanes that must be monitored by a single device. Overhead mounting of non-intrusive devices greatly diminishes (if not eliminates) the occlusion problem, but increases maintenance requirements. For example, optimal performance of video sensors is attained when the cameras are located closest to the freeway and as high as feasible.[1]

- **Calibration.** All equipment must be calibrated to local conditions to some degree. Often this relies on judgment by field personnel because "ground truth" data on which to perform the calibration do not exist. For roadway-based loop detectors, the loops must be "tuned" correctly.

- **Inadequate Maintenance.** Poorly maintained field equipment can lead to both subtle errors creeping into the data as well as catastrophic failures.

- **Communication failures.** Transmission problems – both intermittent and long-term – can lead to gaps in the data (i.e., missing data) even though data may be correctly collected in the field.

- **Equipment breakdowns.** Physical or software-related failures of the equipment are a major source of traffic data quality problems.

Detection of "Bad" Data

The white paper, *Defining and Measuring Traffic Data Quality*, presents a full discussion of how questionable/inaccurate data are identified after they are collected from the field. A variety of methods are used including: internal range checks, cross-checks, time series patterns, comparison to theory, and historical patterns are used.

Correction of "Bad" Data

Once suspect data are identified, the question then is what to do about them. Most applications flag the records failing quality control or set the measurement values to missing or other special codes. Editing the measurement values is far less common, although some experimentation with "imputing" values has taken place. Imputation appears to be most applicable where intermittent gaps appear in the data rather than large portions of time with missing or suspect data. A variety of techniques have been explored including time series smoothing and historical growth rates by location and day and week. However, there is little consensus in the profession on what techniques to be used, or if imputation should be done at all.

Quality Issues for Using ITS-Generated Data for Traditional Uses

Operational vs. Traditional Uses of ITS-Generated Traffic Data

The applications that traffic data support in each of the realms – as well as the nuances of data collection in both cases – can have an impact on data quality. Several differences exist based on these points, as discussed below.

Volumes vs. Speeds. A review of operational and traditional applications was presented in Section 2. Based on these applications, the most notable difference between operational and traditional use of traffic data is the emphasis on speeds and occupancies in the former and on volumes in the latter. Traditional applications use volumes are their basis – speeds are often modeled after the fact in specific applications. Yet, most ***current*** operational uses do not use volume very much, if at all. This lack of focus on volumes may lead to ignoring data quality problems related to volumes. This situation is highlighted by the case of Houston's Transtar system. Originally, roadway-based traffic detection was installed on many of Houston's freeways. Later, as electronic toll tags were implemented, Transtar instrumented both toll and non-toll roads to monitor travel times of tag-equipped vehicles. For their applications up to this point, Transtar has found the tag-based travel times to be sufficient and use the roadway-based traffic data as a supplement.

Data Quality Control Methods. The interviews with Operations and Planning personnel revealed that while Planning personnel are used to performing in-depth reviews of traffic data, including the use of QC software, Operations personnel rarely examine the data at this level of detail. Data review from an Operations perspective review is typically limited to whether the detector is reporting any data at all and identifying obvious outliers. Planning review of data is more likely to include more sophisticated range checks, cross-checks, checks against theory, checks against history profiles, and equipment quirks (e.g., consecutive values).

Level of Accuracy. Data quality requirements (i.e., level of accuracy) also vary between the two realms. In terms of volume, a review of the INFOstructure effort reveals that for advanced traffic management purposes, volumes with a +/-10% accuracy would suffice. (Presumably these are applications behind the current state-of-the-practice in traffic management.) This level of accuracy corresponds roughly to those of Planning-oriented traffic monitoring for short-duration counts, considering the inherent problems in the adjustment process. For continuous count data, however, +/-10% accuracy may be too lenient a threshold – most traffic monitoring units would like a much tighter error bound on these data. Therefore, ITS-generate data with +/-10% error tolerance are probably adequate for estimating AADTs on roadway segments, but other applications of continuous count data (factor and temporal distribution development) are questionable.

The INFOstructure's estimates of speed accuracy requirements are 5-10% for traffic management and 20% for traveler information applications. For performance monitoring purposes, an error tolerance of 5-10% is probably adequate. However, the degree to which this tolerance is currently achieved is largely unknown and likely varies significantly from area to area.

Recent work by Mitretek Systems on data accuracy requirements for advanced traveler information systems (ATISs) indicates that familiar commuters benefit from knowing point-to-point travel times within 10-20 percent of their true values. Travel time estimates beyond 20 percent accuracy range still benefit certain subsets of commuters, but most commuters would be better off just relying on their own experience and sticking to a habitual route. In the Mitretek study, squeezing error below 5 percent doesn't seem to have a great deal of benefit. The Mitretek results correspond to the estimates subjectively developed in an earlier ATIS effort that found the desired error rate of travel times developed by aggregating point speeds should be "less than 15 percent". However, these results need to be tempered by the method used to estimate travel times. Direct measurement systems – those that measure the passage of vehicles over extended highway segments (such as probes) – provide the most accurate estimates. If point-to-point travel times are synthesized using a series of roadway-based detectors (spot speeds), then the accuracy of the individual measurements becomes more critical. If the individual measurements are independent (unbiased), then errors will tend to cancel out so that the accuracy of any given detector can be in the 10-20 percent range. If, however, the measurements are biased in one direction, then the errors will be additive, and the accuracy of individual detectors will have to be more stringent.

Data Collection Nuances. Differences in data collection methodology can also lead to quality problems. One of the most significant is the polling rate and how communication failures interact with it. In traffic monitoring programs, continuous traffic volumes are usually accumulated to hour summaries by the field equipment and then transmitted to a central location every 24 hours. If the communications link for this transmission fails, it is simply re-established. ITS traffic data are typically accumulated to 20- or 30-second intervals by the field equipment and then transmitted immediately. However, if the transmission fails, the field equipment is not likely to be re-polled since it's well into its next reporting cycle. This potentially leads to intermittent gaps in ITS-generated traffic data.

Data Management. An issue related to the aggregation and polling issue is that of data management. Because of the lower level of aggregation and the multitude of sensor locations in an urban area, the sheer volume of ITS-generated data can easily overwhelm Planning-oriented traffic monitoring programs. While this is largely an issue that can be dealt with by increasing computer resources and developing software, it is still a barrier to the sharing of data between the two realms.

Level of Coverage. Another problem raised by the differences in data collection methodology is that of coverage. Detailed traffic data collection for operations only currently cover a portion of urban freeways (22% of urban freeway miles in the 76 largest metropolitan areas had electronic surveillance in 2000) and a smaller portion for signalized arterials. (Generally only advanced control systems like real-time traffic adaptive control collect the type of traffic data useful for traditional applications.) While ITS deployments will continue to grow, they will still tend to be ***concentrated on congested freeway corridors*** because these are the ones in need of operational control strategies. Thus, the data needs of Planning-oriented traffic monitoring programs can never be fully replaced by ITS sources, but ITS can supply information in areas that are historically difficult to place portable equipment.

Vehicle Classification Definitions. It is possible that length-based vehicle classifications will become more prominent in ITS installations. While the length-based bins are useful on their own for a variety of purposes, locally-developed procedures for translating length-based classes and both axle/power unit/trailering (FHWA) and weight class/fuel type (EPA) classification schemes may be possible.

Institutional and Data Sharing Issues. As ITS deployments advance throughout the country, traffic management centers and traditional traffic departments are pursuing innovative approaches to collect, share and disseminate data that is better in quality, more reliable, and easily available. Quality of data is critical, especially when sharing data between regions or jurisdictions, and when this data is made available to the public to make better informed decisions (mostly applicable to ITS generated data). A recent report, addressed specific issues on data sharing techniques, mechanisms, and policies that public agencies use to share data among other public agencies or private agencies. The report collected information from a literature search and enhanced it by conducting a total of 34 telephone interviews with the public sector. Some of the salient features regarding data sharing and its applicability to data quality include:

- Most of the agencies that were interviewed are concerned with collecting traffic data and in some cases multi-modal data.

- When asked what was the main reason for sharing data, most agencies responded that they were motivated to share public travel data to enhance coordination among the region's transportation agencies and to improve overall travel conditions.

- Highway-related data and real-time highway data are the most common type of information that were shared between agencies. Types of information included electronic/digital form (the most popular (24-25 of the 34 agencies), verbal and video.

- Agencies did not distinguish what types and form of data was shared based on who was receiving it. Public agencies shared similar types of data with other public agencies and private enterprises.

- But, when the public agencies were asked whom they share the data with the most, of the 33 agencies that answered this question, 31 share data with other public agencies. The category "other public agencies" is followed by, in order of frequency mentioned, local TV, traffic reporting organizations, local radio, Internet service providers, other organizations, and local newspapers. About a third of the data providers supply local newspapers with information.

- In terms of the types of public sector organizations data was shared with, the most frequently cited were other local jurisdictions such as counties and cities and more specific departments such as the department of public works. Other organizations frequently mentioned include the state police, 911 systems, the State DOT, and transit agencies. Mentioned less frequently were emergency management departments, an airport, a university, and a state parks agency.

- Addressing the need for data quality while data sharing, one public agency respondent mentioned that having a common format and protocol along with data consistency and reliability is necessary.

Recommendations: Possible Solutions

Sampling of ITS Locations and Data Streams

Planning-oriented traffic monitoring programs have begun to recognize the value of ITS-generated traffic data. However, the number of locations where ITS data are collected is quite large. States accustomed to roughly 100 continuous count locations statewide can have that number doubled or tripled if they accepted data from all ITS sensor locations in a single urban area. To get around this problem, some states have identified selected ITS sensor locations where they accept continuous data. The feeling is that for the time being, continuous data collected at ½-mile intervals is not necessary for characterizing traffic in a corridor – short counts at other locations can suffice, especially if they can be adjusted with facility-specific factors from the continuous locations. An extension of this strategy would be to take samples from the remaining ITS sensor locations (say, 48-hour counts once a month or season), but this has not been tested to our knowledge.

Shared Resources

Operations personnel are generally aware of data quality problems but routinely cite the lack of funding for maintenance as a barrier to correcting them, especially in light of the fact that most of their current applications do not require highly accurate data. (As discussed later, this situation may be changing.) Conversely, Planning-oriented traffic monitoring programs generally follow rigorous maintenance schedules when equipment produces data of poor quality. The difference is due to the missions of each group and the level of redundancy in equipment. Traffic monitoring units are in business to collect data while data collection for operations personnel is a tool used to implement operational response strategies. Also, the high density of ITS equipment placement means there is a high degree of redundancy – if a sensor goes down, there are others located close by. This is not a luxury for traffic monitoring activities where permanent equipment is highly isolated.

Given these facts, the potential exists for sharing maintenance resources. Traffic monitoring units have accumulated a long history of maintenance experience that could be tapped by operations personnel if appropriate institutional and funding arrangements can be negotiated. The data quality control methods used by traffic monitoring units is another potential shared resource that can be tapped, although the time scales for ITS-generated data (1- to 15-minute intervals) are typically much smaller than those used for Planning purposes (typically 1 hour).

Maintenance, Calibration, and Performance Standards

Data quality issues are increasingly creeping into the mindset of Operations personnel. Part of the problem is that funding for equipment maintenance was not originally estimated accurately and has not been adequately documented since. In response, some locations are undertaking formal studies of data quality by setting standards and goals for the quality of data they need to support operational strategies and the funding necessary to achieve these standards and goals. This formalization of the process provides a basis for operations personnel to request the additional funding.

Calibration methods and benchmarks are another area worth exploring. Guidance on how to test newly installed equipment – as well as to perform periodic field checking – would be helpful to Operations personnel responsible for detector maintenance.

Contractual Arrangements

A noticeable trend in Planning-oriented traffic monitoring is the outsourcing of data collection activities to private firms. Under such arrangements, contractors are responsible for maintaining equipment and data quality. Some ITS deployments also use contractor personnel as staff extensions for data collection and maintenance. An even more radical model is now being supported by FHWA under the Intelligent Transportation Infrastructure Program (ITIP) where a private firm collects and archives data using their own equipment. They then build traveler information products for sale in the consumer market. Presumably these data can also be made available to public agencies for other types of operational strategies. (ITIP is currently in 2 cities today with 21 more to be added in the next 2 years.) However, the current ITIP effort is subsidized by FHWA – the long-term independent viability of this business model is problematic. When the private sector is involved in data collection, there exists a potential for using formal data quality performance standards as an incentive.

More Sophisticated Operations Applications as a Data Quality Leader

Perhaps the best way to influence the quality of ITS-generated traffic data is to foster the development of more sophisticated operational response strategies that require more accurate and timely data. In truth, the current generation of operational strategies do not require extremely accurate data – operators typically need to know where the big problems are and their responses are geared to this.

However, there are indications that the situation is changing. Information on system performance in real-time is at the core of implementing Operational strategies. As recently noted in an FHWA-sponsored effort: "As more transportation agencies move aggressively toward system operations and performance measurement, the need for comprehensive quality data becomes imperative". In addition to Operations, the same information can also be used in a historical sense to develop performance monitoring statistics. Recent Federal efforts on specifying the so-called INFOstructure and the "data gap" for traveler information systems have taken a big step toward identifying data requirements for Operations. Performance monitoring has also been advanced by efforts such as FHWA's Mobility Monitoring Program. However, it is clear that these efforts are built around the current state of the practice. The Future Strategic Highway Research Program (F-SHRP), a proposed multiyear effort that has improved Operations as one of its four focus areas (under the heading of "travel time reliability") offers the potential for advancing Operations practice significantly. The Reliability portion of F-SHRP includes several proposed projects on performance monitoring, improved data use, and advanced data collection technologies that if implemented, will improve the long-term prospectus for data quality.

Even without the benefit of F-SHRP, other Federal and state efforts are considering more advanced forms of Operational control strategies. As Operational strategies become more sophisticated – and performance monitoring becomes more detailed – data requirements are expected to increase. Specifically, several applications on the short-term horizon can be identified as driving the need for more intricate and accurate data:

- Posting estimated travel times to common destinations on dynamic message signs (DMSs).
- Real-time predictive models that forecast short-range traffic conditions rather than just simply providing a snapshot of current conditions (e.g., the expected queue build-up in 15 minutes from an incident that just occurred).
- Customized traveler information, including alternative and dynamic route guidance.
- Decomposition of delay into its component sources for performance monitoring purposes.
- Integrated freeway/arterial traffic control as well as cross-jurisdictional traffic control.
- Advanced forms of evacuation and military deployment routing.

The recent field operational test on TMC use of archived data is seen as a mechanism for highlighting many of these emerging applications. This operational test is an excellent opportunity to promote data quality, especially with regard to TMC applications, and should be monitored closely.

New Technologies

Monitoring of traffic conditions in real-time is a crucial component of Operational response strategies. When ITS deployment originally was initiated, inductive loop detectors imbedded in pavement were the predominant technology used to monitor vehicle speeds, volumes, and (indirectly) roadway density. In the past decade, increasing use has been made of "non-intrusive" technologies such as video image processing, radar, and acoustic devices to collect the same data. These are termed "non-intrusive" because the devices are mounted on the side of the roadway or overhead, thus avoiding the damaging effects of traffic and the maintenance difficulties with loops. Some areas are using data from probe vehicles (usually toll-tag equipped) to generate travel times. Despite these advances, a number of issues still remain that must be addressed if Operational strategies are to reach their full potential:

- Capital, installation, and maintenance costs – there is a need to reduce these costs so that greater deployment can be achieved. A better understanding/documentation of these costs would also lead to better deployments.
- Coverage – instrumentation is usually done on only roadways of great interest. However, knowledge of traffic conditions on alternative routes as well as the entire system is necessary for sophisticated Operational strategies to have an effect.

- Signalized highway conditions – point-based detectors provide adequate data for freeway performance but are not very useful on signalized highways where most delay occurs at the signal itself.
- Data types – point-based detectors provide spot speeds yet travel times over roadway segments are more useful for many Operational strategies (e.g., traveler information)
- Probe vehicle shortcomings – unless a substantial portion of the fleet is equipped as probes, accuracy may be a problem; roadside readers need to be placed at relatively short distances to provide the level of detail required; volumes are not collected (these are expected to be required for advanced short-term predictive algorithms).

REFERENCES

1. Hu, Pat *et al*, *Proof of Concept of ITS as an Alternate Data Source: A Demonstration Project of Florida and New York Data*, prepared for FHWA, September 30, 2001, http://www-cta.ornl.gov/Publications/Proof_of_Concept.pdf

2. MNDOT and SRF Consulting Group, *NIT Phase II: Evaluation of Non-Intrusive Technologies for Traffic Detection, Final Report*, September 2002.

3. **http://www-cta.ornl.gov/Publications/Proof_of_Concept.pdf**

4. Battelle Memorial Institute and Cambridge Systematics, Inc., *Potential Use of Archived Intelligent Transportation Systems Data for Government Reporting,* prepared for FHWA, September 2002.

5. Tarnoff, P.J. *Getting to the INFOstructure*. White Paper prepared for the TRB Roadway INFOstructure Conference, August 2002.

6. Conversation with Karl Wunderlich. Publication of results is forthcoming under the HOWLATE series of documents developed by Mitretek.

7. *Closing the Data Gap: Guidelines for Quality ATIS Data*, Prepared for: ITS America and the United States Department of Transportation, April 2000.

8. Battelle Memorial Institute, *Sharing Data for Traveler Information: Practices and Policies of Public Agencies*, prepared for USDOT, July 2001.

9. Schumann, Rick, *Summary Of Transportation Operations Data Issues*, PBS&J, August 2001.

10. Cambridge Systematics, Inc. *et al.*, *Research Plan for Providing a Highway System With Reliable Travel Times (Draft)*, prepared for NCHRP 20-58(3), December 2002.