TRANSPORTATION RESEARCH

CIRCULAR

Number E-C292 June 2024

Advancing Highway Traffic Monitoring Through Strategic Research

2024 Update

NATIONAL ACADEMIES Sciences Engineering Medicine

TRB TRANSPORTATION RESEARCH BOARD

TRANSPORTATION RESEARCH CIRCULAR E-C292

Advancing Highway Traffic Monitoring Through Strategic Research

2024 Update

Submitted
February 2024

Transportation Research Board
500 Fifth Street, NW
Washington, DC
www.trb.org

TRANSPORTATION RESEARCH CIRCULAR E-C292
ISSN 0097-8515

The **Transportation Research Board** is one of seven major program divisions of the National Academies of Sciences, Engineering, and Medicine. The mission of the Transportation Research Board is to mobilize expertise, experience, and knowledge to anticipate and solve complex transportation-related challenges.

The **Transportation Research Board** is distributing this E-Circular to make the information contained herein available for use by individual practitioners in state and local transportation agencies, researchers in academic institutions, and other members of the transportation research community. The information in this E-Circular was taken directly from the submission of the authors. This document is not a report of the National Academies of Sciences, Engineering, and Medicine.

Foreword

Monitoring travel on the nation's roadways is crucial for understanding the dynamics of the U.S. transportation system. This involves both broad assessments, such as calculating annual average daily traffic to report on vehicle miles traveled, and detailed analyses, like examining freeway operations on a minute-by-minute basis. Achieving high-quality, actionable data requires considerable investment in time and intellectual effort by a qualified workforce. It is through the relentless dedication of practitioners that we can generate and sustain accurate measures of roadway usage. High-quality monitoring is vital for addressing transportation challenges, including reducing traffic injuries, minimizing greenhouse gas emissions, and alleviating social inequities, thereby aiding practitioners and policymakers in making informed decisions.

Every five years, TRB's Committee on Highway Traffic Monitoring compiles insights from field experts to evaluate current practices and innovations in travel monitoring, setting the groundwork for future research directions. This document, a collaborative effort of over 60 specialists across different sectors of travel monitoring, encapsulates the collective wisdom in the field within the United States. This document can serve as a valuable resource for practitioners seeking to navigate the complexities of travel monitoring, highlighting both the well-understood areas and the knowledge gaps that need addressing to enhance comprehension of roadway travel.

Acknowledgments

This E-Circular was developed over the past year with the leadership and assistance of the Highway Traffic Monitoring Committee members and its friends. Over 60 subject matter experts from public agencies, academia, and private industry contributed their time and experience to this document. The dedicated efforts, unselfish time commitments, insights, and inspiration of these volunteers made this E-Circular possible.

The Transportation Research Board's Highway Traffic Monitoring Committee (ACP70) member Josh Roll led the development of this E-Circular with support and encouragement from chair Ioannis Tsapakis and research working group coordinators Andrew Nichols and Jaqueline Masaki. Every chapter was steered by a dedicated chapter leader who orchestrated a team of experts to engage in discussions, write content, and collaboratively produce each section, encompassing the compilation of research ideas. The chapter leaders include Russell Lewis, Steven Jessberger, Aaron Moss, Jijo Kulathintekizhakethil, Barbara Ostrom, Weimin Huang, Olga Selezneva, Xu Zhang, and Sirisha Kothuri. Xu Zhang, Debbie Walker, Larry Klein, and Elizabeth Stolz reviewed the full E-Circular, providing important suggestions and changes that were incorporated into the finished product. Without the leadership of these individuals, this E-Circular would not have been possible.

Contents

Continuous Count Programs

Russell Lewis
Wisconsin Department of Transportation

Sean Diehl
Drakewell Inc.

PURPOSE

A robust continuous count program is the strong foundation of every traffic monitoring program. The data from these continuous count stations (CCSs) provide important trends and reports that can be used to make essential transportation decisions daily (*1*). The types of data needed for submittal through the Highway Performance Monitoring System (HPMS) and Travel Monitoring Analysis System (TMAS) are volume, speed, classification, and weight (*2*). The legal obligations for a state to provide this data are explained in 23 Code of Federal Regulations, Part 500 Subpart B (*3*). The primary use of this reporting by the Federal Highway Administration (FHWA) is for allocating road improvement funds throughout the country, which requires states' data submittals to be as accurate as possible. Along with this required federal reporting, each CCS program has different internal stakeholders whose individual data needs should be addressed when being the program is being built. Different types of data can be collected with different technologies which have varying costs associated with them. A state should consider a CCS hierarchy, as indicated in Figure 1, when choosing count locations using input from its internal data users. This figure highlights that all sites should collect traffic volume with fewer sites collecting length classification and speed and fewer still axle classification and weigh-in-motion (WIM) capability.

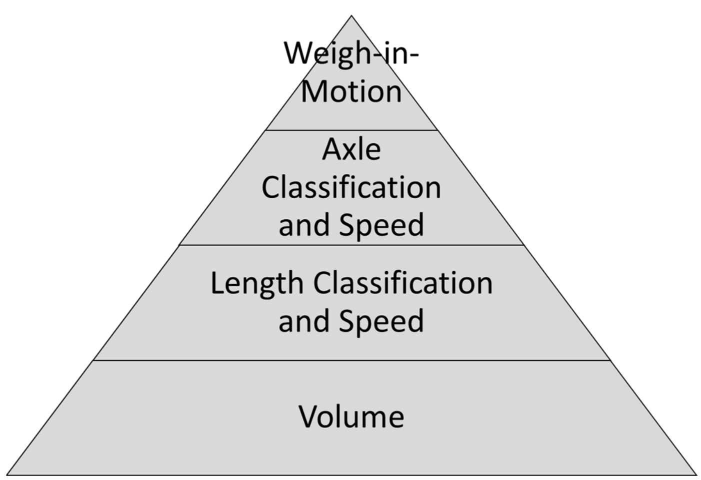

FIGURE 1 Continuous count hierarchy.

CCS Hierarchy

Due to budgetary constraints, not every segment collected can have a CCS, and not every CCS can collect every data type.

Cost is the largest limiting factor in the number of CCSs in a program. A 2-lane WIM site can cost upward of $100,000 for materials and installation, whereas a 2-lane volume-only site can cost less than $10,000. Typically, the higher the cost of a site, the more information the user can obtain from it. Volume data is very important, but other vehicle characteristics are needed for areas like pavement design (*2*), making it important to have some sites that collect traffic data at a higher level of detail.

Another important use for the CCS data is creating factors for the short-term traffic count program. These counts make up most of the statewide coverage for traffic counts and it is important that each factor group and functional class has enough CCSs to represent them (*2*). More information on the factoring process can be found in the section on Short-Term Traffic Count Programs.

STATE OF THE PRACTICE

Site Selection

The *Traffic Monitoring Guide* provides more guidance on calculating the number of sites needed for each factor group (2). As stated before, CCS locations should be determined by a centralized decision process based on internal stakeholder needs. For instance, safety analysis relies on speed data, while pavement design is tied to axle class and weigh-in-motion data. Working together with the stakeholders is key to making sure the funds to install these sites are being used correctly. There should be an agreement on where and how many of each type of site there are for each factor group.

Traffic monitoring programs should also use funding from road construction projects to install CCSs. The program should be able to coordinate with project managers and explain to them the importance and potential uses of these sites. The overall cost of a CCS is minuscule in comparison to the total cost of a construction project and data provided from it will likely be useful for future projects as well as for the entire CCS program. For more information about on-site selections for pedestrian and bike counts, see Bicycle and Pedestrian Traffic Monitoring.

Calibration and Quality Control

All CCSs should be calibrated for each datatype they collect at least once a year. Data from a site that is out of calibration can cause significant repercussions, including skewing short-duration factoring, impacting safety analyses and project prioritization outcomes, and potentially affecting project designs. Testing methods to validate and monitor the quality of data at CCS sites therefore must be considered an essential aspect of the budget for a CCS program.

The accuracy of these sites can be checked and monitored by processing the files every day and running parameter checks to verify that the data are in line with the previous calibration. Speed data should be validated using a radar gun and adjusting loop spacings or sensor sensitivities. Axle spacings (specifically the 2nd–3rd axle space on a Class 9 truck) should be monitored to make sure the calibration is holding. Weight data should be calibrated using any of the standards (test truck, continuous calibration) to ensure that accurate axle weight data is being used to create the load spectra in the Mechanistic-Empirical Pavement Design Guide (MEPDG) (*3*).

For more information about calibration and quality control for pedestrian and bike counts, see Bicycle and Pedestrian Traffic Monitoring.

Length Classification Data

Being able to use every datatype from a CCS is important as it provides the reasoning behind installing the site in the first place. CCSs are expensive to install and maintain, so not using all the data provided would waste funds. Length classification sites are relatively cheap compared to traditional vehicle classification collection methods but can provide valuable data for areas such as pavement design and freight operations.

Axle factors and truck estimates can be obtained from these sites with the correct bin schemes and calibration methods (*4*). Being able to use accurate length class data gives better coverage for axle factor groups and allows for accurate truck data collection in areas where road tubes will not last or could not be safely deployed.

Accurate and Affordable WIM Data Collection

WIM data is one of the most important outputs a CCS can produce. This data is not only used for FHWA submittals but is a key input in the MEPDG (*3*). Programs should not just look at the accuracy of gross vehicle weight (GVW), but the individual axles as well (single, tandem, tridem) since these axle groups drive the MEPDG (*4*). A CCS can produce accurate GVW but have skewed individual axle weights, which will change MEPDG results (*3*).
Traffic monitoring programs should test new sensors that can collect these accurate individual axle weights. Sensors that can provide this information are becoming more affordable. More information on this topic can be found in the Weigh-in-Motion section.

STATE OF THE ART

Calibration and Quality Control

Some counters can passively test and alert to faults in equipment, including sensors and batteries, to help ensure timely maintenance and maximize quality uptime of CCS sites. Additionally, software can store and present this type of data in reports, dashboards, and automated notification systems, all of which help ensure the best use of this information for the traffic monitoring program's benefit.

Side-Fire Axle Classification Technology

Keeping traffic counting equipment out of the roadway saves money, creates a safer working environment, and results in less impact to the traveling public during installation and maintenance. CCS programs should research and learn how to install sites that can collect 13-bin axle classification data from the side of the road. Several options are being used to collect side-fire axle classification data.

Infrared: Uses infrared beams to classify vehicles. These are as accurate at providing the FHWA 13-bin axle classification data as traditional in-pavement methods, but data accuracy levels are more susceptible to weather events (e.g., snow, rain, fog). These sites cannot be installed just anywhere; they need enough shoulder drop-off as well as the right crown in the road to collect accurately.

Camera: These sites collect video recordings and then use artificial intelligence (AI) to automatically classify passing vehicles. This new technology bypasses having to pay for video counts to process and can provide next-day data rather than having to wait for a company to process the data. These cameras do have issues with occlusion and certain weather events, but these issues can be solved with more height and with machine learning.

Installing these types of sites instead of the traditional axle classification (loop/piezo) will save money upfront with the initial installation cost. They will also save money in the future through not having to repair sensors in the roadway (lane closures are one of the most expensive costs in a repair budget).

Individual Vehicle Record Collection and Use

For years, CCS programs had issues with costs associated with storing individual vehicle record (IVR) data. Vendors also had maximum memory limits that did not allow for IVR data storage and collection (*5*). However, the cost of storage has become reasonable, and most counters now allow for IVR data collection. CCS programs should attempt to convert as many sites to IVR data collection as possible, assuming the data is accurate enough. One way to check the accuracy is through a “quality control and process method (*5*).” Not all IVR data is accurate, so it is important to run quality control on it when necessary. IVR data can also be reprocessed if changes in binning schemes or vehicle adaptations occur so the data can capture these changes.

There are a vast number of uses for IVR data, from calculating vehicle headways/gaps and individual classification vehicles’ speeds to calculating the exact average or 85th percentile speed. Users of the data can comb through and find new information about a site with IVR data, which traditional binned data does not allow.

EMERGING TRENDS AND DRIVERS OF CHANGE

Privatization of Traffic Monitoring Programs Several states have instituted a “pay-for-data” arrangement for collecting traffic data. Managing operational requirements remains a formidable challenge for CCS programs due, in part for many, to persistently constrained funding levels and loss of institutional knowledge and experience. Knowledge loss due to retirements, workforce reductions, attrition, and other operational or workforce realities are well known obstacles to successful CCS programs. Additionally, CCS programs that contract construction, maintenance, and software (polling and telemetry, data processing) often have more than one contract to receive all goods and services they require. Managing multiple contracts for one program's goal means managing the space between the contracts for the program. For example, when data is not appearing for a CCS site in the state's traffic monitoring system, are the sensors, the recorder, the modem, the cellular network, or the polling software at issue? Troubleshooting becomes the job of the CCS program, and the answer is not always clear. This leaves the state to rely largely on disparate contracted vendors to communicate and cooperate with one another. Maintaining staff levels with enough knowledge and expertise to conduct fieldwork or to manage several contracts for many CCS and traffic monitoring programs may be increasingly difficult.

Some state DOTs (such as Georgia and South Carolina) have awarded "pay-for-data" contracts of varying scopes to overcome some of these challenges. Some include CCS, short-term counts, polling, and traffic data management software. Others are more limited, bundling CCS with polling. The consistent theme with CCS is that the data (of specified quantity, quality, types, durations, locations) are the deliverable, not the specific construction or maintenance activities of data collection sites themselves. In this model, the vendor is paid for quality data of a specified type and quantity. The contracts often include penalties for missing delivery targets, which incentivizes uptime and therefore vendor responsiveness and proactiveness to site functionality.

Research still needs to be done on whether this is a cost-effective approach based on internal users' data needs. Additionally, the approach taken must be carefully considered given the unique needs of the CCS program, the broader traffic data monitoring program, and the stakeholders. However, some states are seeing success in initial implementations of this contracting and delivery model. These successful state approaches could be a worthwhile consideration for any potential CCS program.

Data Sharing

- Traffic Monitoring Center (TMC): CCS programs are not the only areas that collect continuous traffic data. Most agencies have a TMC that may collect similar data but with different needs. These business units are more interested in real-time data and may not put as much importance on complete data sets or have the same accuracy level requirements for specific vehicle types. However, there are such large overlaps in technologies used and data collected that, with agreed-upon specifications, coordination and planning for site selections, and planning for construction and maintenance, agencies could vastly expand their CCS and even ITS programs. Several programs have implemented this strategy successfully. However, there is a need for research into understanding how these programs have been successful and gathering their lessons learned. There is also a need for research to understand from those agencies not engaging in such data sharing why they are not, what barriers to success they face, and what tools or funding they might use for meaningful engagement.

- DOT State Patrol (DSP): DSP monitor roadways for overweight vehicles using a combination of safety and weight enforcement facilities (SWEFs) and virtual WIM stations (VWIMs). Most of these sites have mainline WIM data that are being used to screen vehicles. These data are available and can be used by CCS programs for their WIM data needs. Since these are such expensive sites, sharing this data is a large cost savings for the program. Identifying the available data and its owners is the first step. The second step is identifying the barriers to data sharing. These are often structural barriers, such as IT infrastructure and security between agencies, or data fidelity concerns that can be overcome through communication between entities and technology.
- Metropolitan Planning Organization (MPOs)/Local: Some MPOs, cities, and counties collect their own CCS data that can be shared with the state CCS program. The first step in data sharing is engaging with the traffic data collection community in one's area to learn what is being collected and where. Writing clear and reasonable specifications for data quality can help potential data sharing partners understand an organization's data needs, possibly making the partners willing to adjust their collection programs to meet these needs. Like the barriers discussed above, communication and technology can overcome barriers to enable the sharing of data.

GAPS IN PRACTICE AND KNOWLEDGE

CCS Program Maturity Model

It would be beneficial for CCS programs to have a way of determining how well their program is doing with data collection. Several ways would be helpful in determining this.

- Benchmarking CCS Outputs: Researching common ways that CCS programs can evaluate themselves on staffing levels, number of CCSs, road miles evaluated, and associated costs.
- Performance Metrics for CCS Programs: Discovering what metrics should be used to evaluate a traffic monitoring program.
- Evaluating Network Data Completeness: Quantifying data completeness across the agency network where there are no counters for validation.

Traffic Data as a Source of Public Agency Revenue

As stated before, the costs to install and maintain a CCS program are high. Most data are given to consultants for free for the consultants to analyze or use to create products that are then sold. Researching whether a program could sell this data would provide a potential new funding source that would allow a program to not only maintain its current program but grow it as well.

CURRENT AND PROPOSED NATIONAL RESEARCH INITIATIVES

This section describes existing and proposed research to addresses gaps in practice and knowledge for pavement engineering applications related to traffic monitoring.

Current Research

National Cooperative Highway Research Program (NCHRP) Project 03-144, "Leveraging Existing Traffic Signal Assets to Obtain Quality Traffic Counts and Enhance Transportation Monitoring Programs" https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=5126

State departments of transportation (DOTs), MPOs, counties, and other local agencies manage extensive traffic counting programs and have a need to ensure that traffic count data covers a variety of modes of travel (e.g., cycling and walking). These counts support decision-making with the aim of enhancing safety and mobility for the traveling public. There are thousands of existing traffic detection assets throughout the nation that serve traffic management operations. Moreover, other customers of traffic count data such as traffic engineers, traffic monitoring staff, transportation and active transportation planners, and data scientists, as well as non-transportation stakeholders (e.g., those responsible for realty, billboards, economic development), need to combine traffic count data sets in new ways to support various business processes.

As sensor detection technologies mature in assisting traffic operations and intelligent traffic system (ITS) programs, the providers of traffic count programs recognize the potential benefits of using existing infrastructure and data to supplement their counts. However, the diverse efforts underway are generally not summarized, publicized, or leveraged.

Key issues associated with using the data from traffic signal equipment for traditional traffic volume measurement include:

- Inconsistency in data quality and format that varies across vendors and technologies;
- Inconsistency in availability of sensors at all intersections as well as approaches to individual intersections; and
- Variable configuration of sensor equipment causing possible gaps in data availability, quality, and storage even though the equipment itself may be capable of counting vehicles, bikes, and pedestrians.

Research is needed to examine whether the data provided by traffic signal assets can provide accurate traffic counts. The objectives of this research are to:

- Determine the feasibility of using existing or enhanced traffic equipment to collect, store, and disseminate data for purposes other than traffic operations, particularly for traffic monitoring programs;
- Determine the suitability of traffic count data from already installed and existing traffic assets for this purpose; and
- Develop effective practices for obtaining and integrating traffic counts from existing traffic assets.

The research will evaluate types of currently installed traffic monitoring assets and assess the suitability of traffic count data for non-operational traffic data usage. In this research:

- The term "suitability" includes the quality, applicability, and type of the data obtained from the traffic equipment.
- The term "traffic" includes motorized vehicles, micromobility devices, and nonmotorized modes including bicycles and pedestrians.
- The term "traffic assets" includes, but is not limited to, signalized intersections, crosswalk signals, video, loops, magnetometers, radar, and traffic detection cameras.

National Traffic Sensor System Evaluation Program
https://rip.trb.org/View/1957079

Traffic sensors are essential components of all highway traffic monitoring and traffic management systems and reporting. Traffic monitoring depends upon reliable detection and accurate measurement of flow rate, speed, classification, and other parameters for various modes of transportation. Active traffic management systems and other intelligent transportation systems applications require these parameters and more for varied uses like traffic control systems, wrong-way driving detection, near-miss crash analysis, and predictive analysis. Sensor systems based on new and emerging technologies—such as optics, electronics, communications, artificial intelligence, and connected and autonomous vehicle (CAV) applications—are rapidly supplanting traditional traffic sensor systems, but they typically lack independent evaluation of their accuracy and performance. State and local agencies must often rely on informal, inconclusive evaluations and pilot deployments to assess sensor systems' suitability for highway applications. The burden to test every sensor type and revision that comes to market creates massive duplication of effort and wastes time, effort, and funding.

Although millions of traffic sensors are in use, manufacturers and distributers can rarely provide independent third-party test results demonstrating their real-world performance. Sensor errors can seriously affect safety, mobility, and public trust. The lack of information regarding system performance and reliability in different operational domains can lead to misapplication of sensor systems, unacceptable performance, or short service life. An authoritative method and a national testing program are needed to characterize the performance and identify the operational domains of current and emerging traffic sensor systems. The objective of this research is to develop evaluation criteria and testing methods for traffic monitoring sensors and systems, which could serve as the foundation for a national sensor system evaluation program.

Improving Traffic Detection Through New Innovative i-LST Technology Demonstration Pilot
https://rip.trb.org/View/2189923

Current traffic monitoring practices primarily focus on counting the number of vehicles, classifying vehicles by length or axle arrangement, and weighing vehicles. Additional critical information such as body type (tractor or trailer) is not readily captured due to technology limitations. However, body type data and information are vital for goods movement and freight analysis as different commodities are transported by different vehicles. For example, perishables and other temperature-sensitive goods are carried by the so-called reefer, a sealed trailer with a refrigerated unit to keep the truck contents at a cooled temperature. Liquids and gases are typically transported by tanker trucks. Flatbed trailers have open decks with no roofs or sides, offer the greatest flexibility to carry not only oversized goods but also a wide range of other freight. Intermodal trucks carry standardized ocean containers. Dry freight trailers provide significant protection to the freight being moved from both inclement weather and other detrimental effects. In addition to the body type identification challenge from current traffic monitoring technologies, current practices also miss the highly desired data regarding the travel time and on/off points where vehicles enter or exit a roadway network, such as where and when a given truck enters or exits a particular highway. Information like this is vital to modeling and projecting vehicle routing associated with demand analysis.

The proposed pool fund study will deploy and demonstrate a set of effective technologies previously developed through the U.S. Department of Transportation Small Business Innovation Research (SBIR) program (SBIR: Measuring Traffic Performance with the Inductive Loop Detector Signature Technologies | SBIR.gov) to capture not only the legacy traffic monitoring data items but also additional body type and system usage information. The new-to-be-deployed technology requires no new-on-the-roadway physical activities. The new technology relies on utilizing existing roadway embedded loop sensors to gain all needed data.

Proposed

Ground Truth Method and Tools for Evaluating Accuracy, Precision, and Bias of Counting Equipment

This research would document acceptable methods for agencies to employ for consistent evaluation of traffic counting devices so results (accuracy, precision, and bias) can be shared between agencies. Doing this will help vendors—once accepted by one agency the results can be shared so multiple tests will not need to be performed. Having standard practices and methods will help agencies be able to quickly determine if equipment meets or does not meet given specifications and methods. Industrywide this will improve data because only those methods and equipment that pass minimum specs will be used for data collection. Methods and tools for various counting technologies and arrays such as length/axle class, WIM, and differences between intrusive and nonintrusive will need to be considered.

REFERENCES AND OTHER RESOURCES

1. Office of Highway Policy Information - Policy | Federal Highway Administration (dot.gov).
2. *Traffic Monitoring Guide*, Federal Highway Administration, U.S. Department of Transportation, Washington, DC, Jan. 2023. https://www.fhwa.dot.gov/policyinformation/tmguide/.
3. Pierce, L. M., and G. McGovern. *NCHRP Synthesis 457: Implementation of the AASHTO Mechanistic-Empirical Pavement Design Guide and Software*. Transportation Research Board of the National Academies, Washington, DC, 2014. https://doi.org/10.17226/22406.
4. Scott Peterson. Viability of Vehicle Length in Estimating Vehicle Classification and Axle Factors. Madison, WI, 2017. https://www.pooledfund.org/Document/Download/7526.
5. Herbert Southgate and Steven Jessberger. Innovative Individual Vehicle Record Traffic Data Quality Analysis Methods. Washington, DC, 2021. https://doi.org/10.1177/03611981221113566.

Short-Term Traffic Count Programs

Steven Jessberger

U.S. Department of Transportation, Federal Highway Administration

Sean Diehl

Drakewell Inc.

Susie Ford

High Desert Inc.

Kent Taylor

North Carolina Department of Transportation

Gordon (Stuart) Thompson

New Hampshire Department of Transportation

Scott Gordon Vockeroth

Alaska Department of Transportation

Jacqueline Masaki

Florida Department of Transportation

INTRODUCTION

Short-duration traffic count programs have emerged as an essential tool in modern transportation planning and management, providing cost-effective means of collecting vital data on traffic volume, vehicle classification, and weight distribution. These programs can be divided into two subsets: coverage counts and special needs counts, each addressing specific data requirements and objectives. Agencies throughout the nation employ a range of strategies in terms of count duration, location, and frequency to obtain accurate representations of traffic volume for highway segments. Data collected through portable traffic recorders (PTRs) and automatic traffic recorders (ATRs) may range from axle counts to vehicle speeds, depending on the sensor technology used. This section delves into the current state of the practice for short-duration count programs, exploring the sensor technologies, methodologies, and data collection approaches that have contributed to the widespread adoption and success of these programs in meeting diverse traffic monitoring needs.

Advancements in short-duration traffic monitoring technology have been rapid over the past two decades, spurred by factors such as the growing demand for accurate, real-time data; developments in computing and communication technology; limited resources; safety and performance concerns; and increasing traffic volumes. Nonintrusive sensors have gained popularity due to their ease of installation, access, and maintenance, all without disrupting traffic flow. Additionally, automated data processing and quality assurance have streamlined traffic analysis and software compatibility and geographic information systems (GIS) integration have improved reporting flexibility and data management. Progress in data storage, battery life, and solar panel use have also extended the maximum duration of short-duration counts. Consequently, the efficiency of these programs has been enhanced through coordination and collaboration among various data collection efforts, with several state DOTs taking the lead in sharing data across state and local agencies.

STATE OF THE PRACTICE

Short-term counts have had different names over time, including short-duration count, portable count, short count, coverage count, special count, manual count, and temporary count. Many agencies and professionals still use these names, and these terms are all typically synonymous with one another. The 2022 *Traffic Monitoring Guide* consistently uses the term short-term count. Short-term counts can be conducted to collect motorized (vehicle class, speed, volume, weight) and micromobility data.

The primary objective of a short-term program is to conduct counts at enough locations (coverage) on the roadway system so that the traffic data and statistics for a given road segment accurately portray the actual traffic on that segment. Short-term count programs are a core component of a traffic monitoring program. They provide the geographic and system diversity of count coverage needed for broad representation of the transportation system. Highway agencies use short-term programs to help meet federal reporting requirements (e.g., the Highway Performance Monitoring System, or HPMS).

A secondary objective of short-term count programs is to conduct a subset of special counts for the various needs of the agency. This may include supplying information for individual projects (e.g., corridor studies, pre-construction studies, pavement design, traffic control studies), developing policies and specifications (e.g., lane closure policies), and providing broad knowledge of roadway use. Sometimes the data collected for these two objectives can be aligned or shared. For example, a classification count for pre-construction for a bridge replacement project in early design phase could be used as a coverage count. Other times the data is not appropriate to align or share. An example is a small collection of 24-hour volume counts near a stadium during a major event.

The location, duration, and frequency of short-term counts are a function of each agency's policies, funding levels, geographic areas of responsibility, and special needs. Short-duration count data is collected by vendors and agency staff. The amount of collection by vendors versus agency staff varies from agency to agency and even business unit to business unit within agencies tasked with short-term count data collection. However, the trend does seem to be increasingly more vendor-based collection.

The duration of a short-term count can be a period of a few hours up to several days or even a few weeks. Unlike continuous count stations that can have high installation, operating, and maintenance costs, short-term counts are significantly less expensive to conduct per location, providing the time and budget opportunity to collect traffic data at many more locations. Advances in data collection, management, and reporting technologies in recent years have greatly increased efficiencies in collecting, processing, and sharing data. This in turn has helped short-term programs stretch their budgets further to collect the data that is in ever-growing demand and value.

The various ways in which agencies balance the benefits and costs of addressing their objectives against their limited traffic counting budgets have led to different data collection programs nationwide. FHWA's *Traffic Monitoring Guide* (TMG) recommends that short-term count data collection consist of a periodic comprehensive coverage program over the entire system on a maximum 6-year cycle. The coverage plan includes counting the HPMS sample and full-extent sections on a shorter (maximum) 3-year cycle to meet the national HPMS requirement (*1*).

Additionally, a 2015 multistate cooperative pooled fund study led by FHWA reviewed the accuracy and precision of annual average daily traffic (AADT) estimations for various short-term count durations. It resulted in several important conclusions for short-term count programs, including restricting short-term counts to weekday results in more precise AADT estimation and determining that Friday is comparable to other weekdays (it is not necessary to exclude Friday from volume counting from either an accuracy or precision perspective). It also found that a 24-hour count results in statistically similar outcomes as a 48-hour count, allowing for 24-hour counts to be used for AADT calculations for HPMS. This finding agreed with a 2015 study in Illinois (*3*) that 24-hour and 48-hour counts were statistically comparable, given appropriate adjustment factors.

Short-term counts are collected with equipment that typically includes PTRs, commonly referred to as counters. Most short-term counts are conducted by field personnel who bring all necessary items to the location. Typical short-term counts conducted with PTRs use pneumatic tubes stretched across the road or, increasingly, nonintrusive technologies (e.g., video detection and radar). Table 1 summarizes existing short-duration counting equipment along with a short summary of known issues related to those technologies. Some short-term count locations have permanent installation attributes, such as loops in the road with leads to an empty cabinet where a field technician can bring a counter and battery to conduct periodic extended short-term counts. Recent advances in edge-process camera AI are leading to some short count options becoming available to agencies. More advances are seemingly on the horizon and present new forms of nonintrusive technologies that some agencies are beginning to adopt and use (more on this subject in the next section, State of the Art). The more recent emergence of probe-based data (from connected vehicles and cellular phones) presents a new and different technology and data source for short-term count programs that is in early stages of adoption and use by some agency programs.

TABLE 1 Sensors for Collection of Short-Duration Counts of Motorized Vehicles

Technology	Number of Sensors for Speed Data Collection	Vehicle Class Data Collected	Number of Lanes Collected per Sensor	Issues
Pneumatic tubes	2 per lane	Axle-based (FHWA 13+)	1 per pair of sensors	Accuracy limitations under very heavy traffic volumes or stop-and-go conditions; not suited to snowy conditions.
Tape switches	2 per lane	Axle-based (FHWA 13+)	1 per pair of sensors	Need protection of lead wires if placed on lanes not adjacent to shoulders; not suited to snowy or wet conditions.
Magnetometers	2 per lane	Length-based or obtained from vehicle undercarriage profiles	1 per pair of sensors	Some sensors are placed in the pavement, others on the pavement, or even both alongside the road with others under the pavement; requires a short lane closure for sensor placement.
Video detection systems	1 camera for one or multi-lanes; multiple cameras are used for many lanes	Length- or axle-based (FHWA 13+)	Multiple	Mounted to roadside infrastructure such as a pole or sign bridge or with a portable camera mounted on a trailer; possible performance degradation with reduced visibility; susceptible to vehicle occlusion in distant lanes when camera is side mounted.
Piezo	1 or 2 per lane depending on the array used	Axle-based (FHWA 13+)	1 or 2 per array of sensors	Need protection of lead wires if placed on lanes not adjacent to shoulders; very cold weather may affect performance for WIM applications; difficulty taping down with moisture and prevailing traffic.
LiDAR	1 per location	Axle-based (FHWA 13+)	Usually 1 (multiple lanes with some models)	No in-road installation required; possible performance degradation with reduced visibility, including vehicle occlusion in distant lanes when device is side mounted.
Microwave radar	1 per direction (side- mounted), 1 per lane (overhead), some overhead models monitor multiple lanes and multiple directions	Length-based	Multiple	Sensor is mounted on an extensible pole on a trailer pulled to the site or mounted to a roadside infrastructure/pole device; side-mounted radars may have occlusion issues in lanes furthest from sensor with heavy or stop-and-go traffic in a multiple lane scenario.
Acoustic (passive and active)	1 sensor	Length-based	Multiple	Sensor is mounted on an extensible pole on a trailer; background sounds may interfere.
Infrared beam	1 pair of sensors on each side of the roadway	Axle-based (FHWA 13+)	Up to 9 lanes	Aligning the sensor beams can be challenging; large crown in road and snow or debris on the roadside can block beams.

To develop AADT estimates, many agencies factor (i.e., multiply) the Average Daily Traffic (ADT) of a short-term count using one or more temporal and other adjustment factors such as month of year, day of week, time of day, axle, and change rate factors. Short-term counts may be conducted to record different types of data depending on the needs of the program and technology employed (e.g., counter and sensing technologies). For motorized vehicle collection, these can include number of axles, axle spacing, axle weight, traffic volume, volume by vehicle classification, speed, length, gap, and headway. For micromobility, this can include pedestrians, bicycles, e-bikes, scooters, speeds, volumes, and safety adherence (e.g., helmet wearing). The data may either be over a certain time period (binned) or on an individual vehicle basis. FHWA recommends collecting data in IVR format whenever possible. Data storage requirements must be considered when purchasing equipment or evaluating older equipment because IVR data reporting requires large storage capacity (*1*).

Many agencies use vendor-provided technologies such as traffic data management software, mobile field apps, and GIS to plan, collect and conduct, quality assurance (QA) and quality control (QC), process, calculate statistics, and share short-term data and statistics with their data consumers. Short-term count data is now more accessible and widely available than ever before thanks to advances in internet technologies. Internet of Things (IoT) technology improves flow of data from sensors to data systems; cloud computing provides scalable platforms for storing and processing large amounts of data; application programming interfaces (APIs) allow different applications to communicate and share data; big data analytic tools allow traffic data to be processed and analyzed in real-time; mobile and web applications provide field tools to improve data collection safety and accuracy; geolocation services (such as global positioning systems [GPS]) improve count quality assurance; and improved web data visualization tools provide capabilities such as dashboards and maps.

STATE OF THE ART

Short-term traffic monitoring technology has evolved over the past 30 years due to a combination of factors:

- Need for more timely and accurate information
- Advancements in portable collection technology (e.g., magnetometer, radar, and others)
- Need for counting without being in the roadway (using, for example, video or side-fire radar) and the high cost associated with providing traffic control or closing lanes during the installation and maintenance of intrusive sensors
- Advancements in sharing information (traffic counts, project needs) across agencies
- Advancements in low-cost computing and communications technology
- Limited resources devoted to the collection of short-term data
- Office of Safety Model Inventory of Roadway Elements (MIRE) and performance issues caused during the installation and maintenance of intrusive sensors
- Improved portable traffic equipment lifetimes and reduced maintenance cost
- New data reporting requirements (IVR and speed by class)
- Increased traffic volumes

Many agencies have started to test and use nonintrusive sensors (i.e., those located above or to the side of the roadway) to obtain short-term counts. These can often be installed and maintained without personnel having to enter the travel lane. However, although side-mounted sensors have the advantages of easy installation, access, and maintenance, vehicles in lanes farthest from the sensor can be obscured by long and tall vehicles traveling in the lanes closer to the sensor (*5*).

In addition, data processing and QA/QC procedures have largely been automated. In the past, QA/QC was primarily done manually by engineers. However, the majority of the most recent data processing advancements is being provided by software processing vendors or the equipment vendors. Only a few agencies still use in-house software for QC. The equipment or other software typically provides:

- Several tools that allow users to create graphs and traffic reports,
- Calibration for equipment and traffic parameters,
- The ability to edit information about the traffic collection sites,
- Study details,
- Analysis of data,
- An option to set up factoring groups,
- Automated tests to check analysis results for certain conditions,
- The ability to export and email results in commonly used formats (e.g., Excel),
- Customized settings for time formatting, and
- Units of measurement (speed, temperature, spacing, and weights).

Software compatibility across manufacturers has started to expand, thus increasing reporting flexibility. Manufacturers have also started to integrate GIS into their software and hardware by improving data management practices. Some equipment can be accessed remotely (e.g., over the Internet), allowing users to monitor and transfer real-time data and minimizing the need to manually extract the data from the equipment. Some equipment automatically, or with minimal effort, allows for the collection of the GIS positioning when placing the equipment for a traffic count study. Other advancements in traffic count technology, including lower power consumption and addition of on-device solar panels, support an overall increase in the maximum duration of a short-term count to a period of a week or more.

In general, short-term count programs become more efficient as various data collection efforts are coordinated so that one counting study meets multiple needs. Examples of coordination include:

- Sharing data collection activities, equipment, and schedules with local agencies;
- Using technologies that include access to software that makes possible integration, dissemination, and conversion of schedules and data collected from state and local agencies;
- Establishing data governance committees with members from national, state, and local agencies. State DOT leaders sharing data across state and local agencies can be found across the country, including in New York, Colorado, and Ohio.

Local government organizations are most familiar with roadway data and location in their jurisdiction. As state and national purposes expand past the interstate and state roadway system, local data can be an untapped opportunity to validate and fill in gaps for state and federal data needs.

Quality local traffic count data can supplement a state DOT's traffic count program, reduce staffing needs for counting lower functional class roadways, provide data where the state typically does not count, and provide additional quality checks on the state's traffic monitoring program to improve service delivery and residents' quality of life.

Assignment of Short-Term Counts to Temporal (Monthly) Adjustment Factor Groupings

The TMG (*1*) recommends the use of cluster analysis in conjunction with traditional methods for the creation of factor groups. However, a known limitation of the cluster analysis method is that it can produce clusters that may not follow any clear stratifications or boundaries (a lack of specified and definable characteristics). This can limit the clusters' applicability in assigning short-term counts to temporal (monthly) adjustment factor (MAF) groups. The majority of research studies on improving AADT accuracy from short-term counts focused mainly on lowering the errors associated with the creation of adjustment factor groupings. Prior research has shown that the assignment step is the most critical element in the AADT estimation process. Potentially ineffective allocation of short-term counts to MAF groups may triple the prediction error (*6*). In the absence of relevant guidelines and recommendations, further research is needed to fill this gap (*7*).

EMERGING TRENDS AND DRIVERS OF CHANGE

Short-term count programs are trending in different directions and are mainly being driven by economies of scale and improved efficiency. One trend at many agencies is hiring out short-term counting practices to one or more companies who provide some or all short-term counting for that agency. Another trend involves coordinated efforts to buy the short-term counting devices centrally and have statewide memoranda of understanding in place for local agencies to perform counts for designated locations and additional locations. This is what the New York DOT does for its short-term counting program. More recently, some agencies have begun purchasing data to augment and, in some cases, replace short-term counts; Minnesota DOT and others are taking this approach. Some of these practices are occurring due to reductions in the number of staff available to undertake the short-term counting programs.

Over the years, some DOTs have successfully established cooperative collaboration with their local government entities. This relationship has provided contacts at the local level to navigate and garner support for initiatives that include traffic data collection. These initiatives can support national goals (e.g., Safe Routes to School, National Bridge Inventory, Highway-Rail Crossing Inventory Data, use of passive data (probe or others), All Roads Network of Linear Referenced Data (ARNOLD), requirements for HPMS, Office of Safety Model Inventory of Roadway Elements (MIRE) critical to Safety Management, and more.

These relationships require time and staffing but can help improve data sharing practices; lead to all-inclusive policies that speak to federal, state, and local government; and facilitate communication that will better serve the public.

Connected vehicles and probe data more broadly (datasets sourced from connected vehicles, mobile data such as from cell phones, and private fleet GPS) present additional sources of vehicle data analysis for decision-making processes. As with any technology, there are advantages and disadvantages to consider. Improvements in the quantity and quality of data feeding probe data models as well as the models themselves have greatly increased their relative accuracy to field-collected traffic counts.

The advantage of this is the ability to easily obtain a geographically and temporally scalable traffic dataset. However, since the accuracy of these models is tied in part to sample size, low-volume roads and intersections may be prone to errors or variation from field-collected data that are unacceptably large for decision-making. Since the saturation of field counts tends to be primarily on higher-volume roads, the need for greater geographic and temporal coverage may not yet be resolved between field and probe data. A further consideration of probe data versus field-collected data is probe data's limitations on vehicle classification (3 types vs. 13 vehicle types) and micromobility data (bicycles and other nonmotorized vehicles on the road, pedestrians on sidewalks or in crosswalks).

Continued interest, research, and use of this data, particularly in conjunction with field-collected data by public agencies, the research community, and industry, could improve the accuracy and trust in these data sources in ways that make them an integral and integrated data source for measuring traffic for analyses and decision-making processes.

GAPS IN PRACTICE AND KNOWLEDGE

This section presents issues, challenges, undiscovered areas, and other gaps pertaining to the collection and analysis of short-term traffic data.

Data Collection on Non-Federal Aid System Roads

A major aspect of the Highway Safety Improvement Program rulemaking is the requirement that states must collect and use a subset of MIRE fundamental data elements (FDEs) for all public roadways, including Non-Federal Aid System (NFAS) roads. NFAS roads are typically rural minor collectors and both rural and urban local roads. States were supposed to define anticipated improvements to collect MIRE FDEs in their traffic records strategic plan by July 1, 2017. By September 30, 2026, data must be accessible for all public paved roads. Traffic volumes (AADT) are already collected under the HPMS for federal aid roadway segments and ramps. States are concerned about their ability to collect and maintain data on local roads. Though new data models and products are becoming available to agencies, more research is needed to help agencies set standards and specifications for the data to both meet the MIRE requirement and properly communicate critical metadata regarding such data to their data consumers.

Lack of Intra-Agency and Interagency Coordination

Many state agencies collect similar but redundant traffic data for various operational reasons such as traffic operations, planning and pre-construction, construction, safety, and traffic monitoring. Additionally, many state agencies lack agreements with local agencies to coordinate data collection activities. Lack of coordination among departments and agencies often leads to duplication of efforts and an inability to share resources toward making traffic counting programs more efficient and cost effective. Lack of coordination also affects the consistency and quality of the data if it is shared. Differences in data collection, such as varying quality standards or classification schemes, can lead to challenges with quality control, data management, or simply overall usability. The concept of "collect it once for many purposes" can pay dividends for public agencies. Research and guidance are needed to help develop more universal data collection and content standards, collection program coordination models, and workable and permissible contracting ideas to share costs or pool funding.

Impact of Construction Activity, Incidents, Weather, and Events

Construction and incidents may have a significant impact on alternative routes that carry rerouted traffic, resulting in increased traffic volumes captured by the traffic equipment. Likewise, the route from which traffic is being diverted will experience decreased traffic volumes. Unless identified and clearly specified by data collection personnel, the final data user has no way of knowing the underlying reasons for abnormality in the data. Weather, particularly around recreation locations such as parks, campgrounds, or trails, can strongly influence short-term traffic patterns on a year-to-year basis. Special events also result in deviations from average conditions and need to be handled with care.

Data Quality Assurance

The ability to efficiently process and assess the quality of data from different collection equipment is a challenging task for many agencies. Not all traffic equipment has its own data processing software and not all vendors produce equipment of the same quality. Some equipment has been heavily tested and operates more reliably than others. Different vendors' equipment can produce different results for any given detection technology because the electronics and the vendors' software may perform differently.

Accuracy of Classification Data Collected in Saturated Traffic Conditions

High traffic volumes or congestion may not be accurately captured by vehicle classification equipment. For example, under saturated traffic conditions, traffic detectors may fail to determine whether a count of four axles represents two cars or one truck. Further, traffic detectors that work on vehicle presence detection often produce erroneous data under stop-and-go traffic conditions (*8*). Urban areas also present the challenge of frequent entrances to the roadway; acceleration and deceleration of vehicles in these areas can lead to the misclassification of vehicles. Many agencies use their own classification scheme for data collection rather than vendor-supplied schemes to improve accuracy. The industry would benefit from the collective documentation of the successes and failures of these efforts to help improve vendor-supplied equipment and the use of various equipment by other agencies.

Securing Road Tubes on the Pavement

The inability to properly secure road tubes on the pavement surface for the duration of a count affects the amount and quality of data collected. In addition, attaching road tubes to the pavement can lead to roadway hazardous conditions and damage to vehicles. This issue is more profound on routes with significant truck traffic and high-volume roads.

Equipment Failures

Equipment malfunctions, communication problems, and other technical failures affect the amount and quality of data collected. Some equipment failures are caused by external factors such as inclement weather conditions, vandalism, utility operations, pavement repair, external circumstances (fire or floods), and maintenance.

Seasonality (Monthly)

As road tubes are still a commonly used collection technique, seasonality in some climates can restrict collection to certain months when tubes can be set successfully. This leaves a shorter timeframe in which to collect roadway data. Agencies also have various types of roadways that only operate for a portion of the year; these roads are only accessible or maintained during particular periods of the year. Seasonal roadways are more common in northern environments with icy roads or roadways that are not plowed during winter months. There is a lack of methodology to accurately reflect travel patterns for these types of roads.

Safety of Traffic Personnel

The safety of staff that install or maintain short-term traffic equipment is a major concern, particularly on high-volume routes. The need to safeguard data collection staff without having to apply traffic control is the main reason that many efforts have focused on advancing nonintrusive detection technologies. Some agencies have policies to always run the engine with a seat belt on when a person is working in the vehicle so the airbag will deploy if there is a collision.

In addition, several documents—such as FHWA's Traffic Detector Handbook (3rd Edition) (9), A Summary of Vehicle Detection and Surveillance Technologies Used in Intelligent Transportation Systems (10), Sensor Technologies and Data Requirements for ITS (11), and ITS Sensors and Architectures for Traffic Management and Connected Vehicles (5)—provide strengths and limitations of various sensor technologies and applications.

CURRENT AND PROPOSED RESEARCH INITIATIVES

This section describes existing and proposed research to addresses gaps in practice and knowledge for pavement engineering applications related to traffic monitoring.

Current Research

NCHRP Project 07-30, "Methods for Assigning Short-Term Traffic Volume Counts to Adjustment Factor Groups for Estimating AADT"
https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4945

Annual average daily traffic (AADT) which represents traffic on a typical day of the year is used by transportation agencies for reporting requirements, allocating resources, informing decision-making, and supporting various agency functions. Transportation agencies use different methods to derive AADT estimates from short-term counts of traffic data from permanent and portable traffic counting equipment installed at selected locations.

Commonly used methods for estimating AADT do not adequately address how short-duration counts should be assigned to adjustment factor groups. Also, there are concerns about the inherent errors in these methods, their applicability to roadways with insufficient traffic data, and the accuracy of the derived AADT estimates. There is a need to improve existing methods and develop new methods for functional classes of roadway where insufficient continuous counting exists to improve accuracy of AADT estimates. These methods will help transportation agencies improve the quality of traffic information and support the decisions regarding capital investment programs and budgets as well as design and maintenance programs.

The objective of this research is to develop rational methods for assigning short-duration traffic volume counts to adjustment factor groups for estimating AADT. The research is concerned with all functional classes of roadways and traffic volumes.

Proposed Research

Interagency Coordination to Increase Number of Counts and Share Data

This research would document examples across the United States in which states, counties, MPOs, and cities collaborate well to collect traffic count data. Researchers would find these collaborative agencies and survey or interview them to understand how their cross-agency collaboration occurs, what data sharing platforms are used, and what data governance principles are deployed to ensure consistency across agencies.

Determine Accuracy and Bias of Portable Technologies for Obtaining Short-Term Traffic Volumes

This research would document errors involved with short-term counting. It would include the total error associated with errors in equipment, air switches, arrays used, sensor choice, age of equipment, and expansion factoring.

REFERENCES AND OTHER RESOURCES

1. *Traffic Monitoring Guide*. Office of Highway Policy Information, Federal Highway Administration, U.S. Department of Transportation, 2016.
2. Handbook of Simplified Practice for Traffic Studies, Chapter 3, Iowa Department of Transportation and Iowa Highway Research Board, Center for Transportation Research and Education, Iowa State University, Ames, Nov. 2002. http://www.ctre.iastate.edu/pubs/traffichandbook/3trafficcounts.pdf.
3. Guide for Interpreting Short-Duration Traffic Count Reports, Washington State Department of Transportation, Dec. 2010. http://wsdot.wa.gov/mapsdata/travel/pdf/Guide_for_Interpreting_Short_Duration_Traffic_Count_Reports.pdf.
4. Hallenbeck, M. E., and H. Weinblatt. *NCHRP Report 509: Equipment for Collecting Traffic Load Data*, Transportation Research Board of the National Academies, Washington, DC, 2004.
5. Klein, L. A. ITS Sensors and Architectures for Traffic Management and Connected Vehicles. Taylor and Francis, Boca Raton, FL, 2018.
6. Davis, G. A., and Y. Guan. Bayesian Assignment of Coverage Count Locations to Factor Groups and Estimation of Mean Daily Traffic. *Transportation Research Record,* No. 1542, pp. 30–37, 1996.
7. Fricker, J. D., C. Xu, and L. Jin. Comparison of Annual Average Daily Traffic Estimates: Traditional Factor, Statistical, Artificial Neural Network, and Fuzzy Basis Neural Network Approach. Presented at 87th Annual Meeting of the Transportation Research Board, Washington, DC, 2008.
8. Fekpe, E., D. Gobalakrishna, and D. Middleton. *Highway Performance Monitoring System Traffic Data for High-Volume Routes: Best Practices and Guidelines*. Office of Highway Policy Information, Federal Highway Administration, U.S. Department of Transportation, 2004.
9. Klein, L. A., D. Gibson, and M. K. Mills. *Traffic Detector Handbook, 3rd Edition*. FHWA-HRT-06- 108 (Vol. I) and FHWA-HRT-06-139 (Vol. II). Federal Highway Administration, U.S. Department of Transportation, October 2006.
10. Klein, L. A. *A Summary of Vehicle Detection and Surveillance Technologies Used in Intelligent Transportation Systems, Vehicle Detector Clearinghouse*. Intelligent Transportation Systems Program Office, Federal Highway Administration, U.S. Department of Transportation, 2007.
11. Klein, L. A. *Sensor Technologies and Data Requirements for ITS*. Artech House, Boston, MA.

Managing Large Traffic Datasets

Aaron Moss
Colorado Department of Transportation

Steven Jessberger
Federal Highway Administration

Jason Breault
Midwestern Software Solutions (MS2)

Lawrence (Larry) Klein
Klein and Associates

Xu Zhang
Kentucky Transportation Center

STATE OF THE PRACTICE

Cities, MPOs, and state DOTs collect traffic counting data for sections of roadway to represent travel patterns on their surface transportation network. Large traffic datasets support capital investment programs and budgets and also effective design and maintenance programs. Data from continuous and short-duration locations can be part of these large datasets. Continuous counting involves collecting traffic data continually within a period of more than one week in any given location and can include up to 365 days per year of hourly count data. Continuous traffic data (volume, classification, speed, and weight) represent the temporal data needed for developing seasonal, axle and growth factors to annualize short-term counts. Short-duration traffic data counts are normally taken over 1–7 days and represent the spatial data sets for an agency's roadways.

Both continuous and short-duration counts are typically taken at the hourly level. However, agencies may determine the need to collect data more frequently, such as in 15-minute intervals. The 2022 FHWA *Traffic Monitoring Guide* (TMG) contains example formats for storage and submittal of traffic data in individual vehicle record (IVR) format in time intervals as small as 1/100th of a second (*1*).

Traffic data stored in large databases represent the temporal and spatial information from the continuous and short-duration counts, station information, volume, speed, classification, weight, and other attributes such as metadata. The TMG is intended to assist state and local transportation agencies and others involved in traffic data acquisition, storage, and reporting programs. The American Association of State Highway and Transportation Officials (AASHTO)

Guidelines for Traffic Data Programs (*2*) contains recommended counting procedures and national traffic monitoring techniques that reflect the current practice.

Although agencies collect both a variety of continuous and short-duration traffic data per their specific needs, there are several similarities across the nation's traffic monitoring programs. For example, the Travel Monitoring Analysis System (TMAS) and Highway Performance Monitoring System (HPMS) programs are universal data submittal formats that state DOTs use to submit their continuous traffic data monthly and the year-end traffic statistics annually. Every traffic counting program must store data for reporting purposes and access by their customers. Data management can be as simple as using a spreadsheet or as sophisticated as a customized software management system. Traffic monitoring programs must be familiar with the various protocols and needs of the agencies they interact with to ensure the best data are used for decision-making. Many agencies document methods for providing QA, QC, and database structures that ensure the collected data are beneficial and relevant to those who use it. Both AASHTO's *Guidelines for Traffic Data Programs* (*2*) and FHWA's TMG (*1*) recommend periodic review of an agency's traffic counting program, which includes manipulation of data and storage in large databases.

A large traffic dataset can lead to better consistency and data integrity, avoiding duplication and ensuring data accuracy through the dataset's design and a series of constraints. Therefore, the more data that are collected at a greater frequency, the more accurate the dataset becomes. The tables in a relational database are linked through primary and foreign keys that function as identifiers in each table to uniquely pinpoint a row. Each table has one or more primary key columns. Other tables linking to the first table contain a foreign key column whose value matches the first table's primary key. Providing these integrated aspects of the large dataset to those who use them to perform other tasks is important. The traffic dataset becomes even larger when datasets are disaggregated. It also becomes larger when data are used to calculate other parameters, such as traffic volume factors for adjusting short-duration traffic counts that represent an AADT statistic. If some data within the large dataset are dependent on external sources or other internal sources, then the limitations of and rights to such data use must be contained in the large dataset's documentation and metadata. Including this information is required to create, document, and make the large dataset available to those managing and using it.

STATE OF THE ART

Traffic monitoring programs maintain their traffic data and related metadata for use by in-house staff, contractors, and public agencies. These data are spread across different databases, often in diverse locations, in various formats, and in large volumes ranging from megabytes to terabytes or even larger. Today, data are often generated both by humans (e.g., manual counts) and by data recorders in large amounts. State-of-the-art programs deploy systems to manage the data and perform the numerous functions needed for success of the traffic monitoring program. Those functions include obtaining the counts from the devices, running initial QC checks to determine completeness of the data, running advanced QC checks to confirm the data meets acceptable ranges, and using established methods to both store and report the traffic data.

Once the quality review is complete, the data are made available in large traffic databases for reporting, are uploaded to other servers or platforms for use in other systems or are placed on publicly viewable websites for public consumption and reporting. Timeliness, quality, completeness, and transparency concerning the methods employed and the data calculation results and reviews are necessary for customers to have access to accurate datasets. Spatial representation and integration of large traffic datasets is increasingly important; new and more advanced methods to visualize the data are now available with added methods to spatially QC the data. What once were link-based (or line-based) data now tell a story though spatial representation as traffic professionals consider safety implications, corridor studies, state-to-state travel patterns, and other applications.

Characteristics of Big Data

When does a traffic database become big data? In 2013, anything over 500 gigabytes (GB) was considered big data; now big data is generally considered more than 1 terabyte (TB) (*3*). The term "big data" is qualitative and difficult to quantify. Hence, big data is identified by characteristics known as the five Vs: volume, velocity, variety, veracity, and value (Figure 2). These characteristics provide a way to consider various characteristics of structuring an application. Considerations include the time increments needed and reported, the structure of the databases, the quantity of data available, the cost/benefit of the data, the control features within the databases that enable full management and exploitation of the large datasets, and the uncertainty of the data. Volume, velocity, variety, veracity, and value are discussed further below.

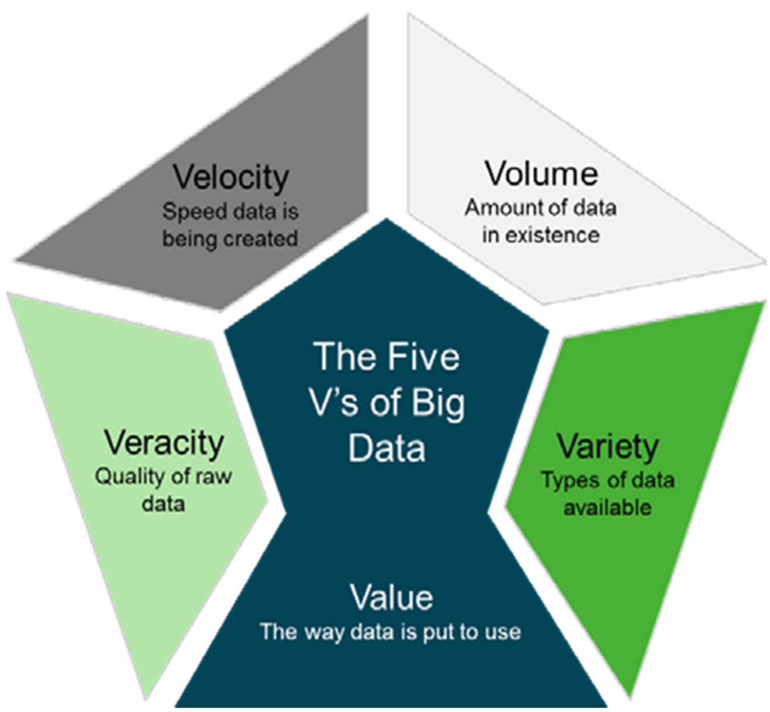

FIGURE 2 The five Vs of big data (*7*).

- **Volume** refers to the size of the dataset. Data availability and scale are growing rapidly. For example, five years of crash data from Florida represents less than 50 megabytes (MB). However, data generated by 300 turning movement count field devices is currently estimated at approximately 635 GB per year, and, if stored, the data from 300 closed circuit television cameras would require hundreds of terabytes of storage each year (*4*). Data from emerging sources such as connected vehicle data is also expected to generate terabytes of data per year.
- **Velocity** characterizes the speed at which the data are generated, processed, and transmitted. Different applications have different latency requirements. In today's competitive world, decision makers want the necessary data and information as quickly as possible. For example, many ITS locations process data in one-minute increments, which some scenarios consider near real time or real time. In different fields and different areas of technology, data acquisition can be generated at different speeds.
- **Variety** signifies the different forms of data and the formats in which they are stored. Today, large volumes of unstructured data are generated apart from the structured data generated in enterprises. Advancements in big data technologies have encouraged industries to develop powerful and reliable tools to extract, transform, and analyze the voluminous unstructured data common today. Current traffic monitoring applications mean that, to stay competitive, organizations must not only rely on structured data from enterprise databases and warehouses, but also consume vast quantities of data generated outside of the enterprise, such as intersection loop data, probe-based data, and Bluetooth and automatic license plate reader data. Using license plate readers for collecting vehicle movements between locations is another example of collecting volumes of data.

  Another aspect of traffic data that place them in the large dataset category is the assortment of traffic-related statistics derived from a volume count. Apart from the traditional flat files, spreadsheets, and relational databases, large quantities of unstructured data are stored in the form of images, video traffic counting files, traffic site maintenance web logs, portable sensor data files, and individual vehicle records.
- **Veracity**. Amassing a lot of data does not mean the data are clean and accurate. Data concerning traffic monitoring must remain consolidated, cleaned, consistent, and current to make the correct decisions. Furthermore, not all data are good. In fact, unfiltered data or data that are not quality controlled are more likely to be bad than good. Although data quality and usability depend largely on the source, big data users need to always be wary of problematic data. Data unreliability may make agency personnel reluctant to rely on using some data even though significant opportunity exists with these data sets. Instead of discarding the unmatched potential of big data, public agencies and companies should work harder on implementing the right technology and people for its management (*5, 6*).
- **Value** denotes how big data datasets contribute to improving the status quo. Value involves determining a benefit and estimating the significance of that benefit across any conceivable circumstance. Value may be the most important of the five Vs, as investments in big data initiatives require a clear understanding of the benefits and

associated costs. Before any attempt to collect or leverage big data, business cases need to be developed to assess the benefits and costs associated with the data collection and analysis efforts (*3*).

Attributes of Traffic Database File Types

Data storage in the form of large traffic databases can occur in flat files or spreadsheets, hierarchical data, relational data, cloud data, data lakes, and GIS databases and single layers, as indicated in Table 2. A spreadsheet or flat file is a simple method for storing data. Individual records have different data in each field with some fields serving as a key (header) to locate a particular record. For example, a traffic station ID number may be one of the key fields in a record for a site's classification data. For some traffic data records, there could be hundreds of fields associated with the traffic station ID. When the number of fields becomes lengthy, a flat file is cumbersome to search and use manually. Also, the key field is usually determined by a programmer and searching by other determinants may be difficult for the user. Although this type of database is simple in its structure, expanding the number of fields usually entails reprogramming. Additionally, adding new records is time consuming, particularly when there are numerous fields.

Hierarchical files store data in more than one type of record, usually described as a "parent–child" or "one-to-many" relationship. One field is key to all records, but data in one record do not have to be repeated in another record. This system allows records with similar attributes to be associated together. The records are linked to each other by a set of key fields in a hierarchy of files. Each record, except for the master record, has a higher level record file linked by a key field pointer. This allows one record to lead to another and so on in a descending pattern, resulting in an efficient data structure when the relationship is clearly defined, and queries follow a standard routine. The database is arranged per its use and customer requirements. Access to different records is readily available but also easy to deny to a user by not furnishing that unique file of the database. One of the disadvantages of hierarchical files is the need to access the master record with the key fields determinant in order to link downward to other records.

Relational files connect different files or tables (relations) without using internal pointers or keys. Instead, a common data link joins or associates records together. The link is not hierarchical. A "matrix of tables" stores the traffic information. If the tables have a common link, they may be combined by the user to form new inquires and data output. This is the most flexible system and is particularly suited to structured query language (SQL). SQL is the most common relational database language in use today. Queries are not limited by a hierarchy of files, but instead are based on relationships that the user establishes from one type of record to another.

Cloud data allows users to access traffic data from anywhere with an internet connection. Since cloud data uses the existing infrastructure of established providers, it can be cost effective since agencies are not required to purchase large servers. Scaling the data is also possible by using existing cloud service providers. Backup data and replicas of the databases can be stored across multiple geographical locations, ensuring continuity in case of a site failure.

A data lake is a storage repository that holds a vast amount of raw data in its native format. Data lakes are often implemented on variations of the Hadoop Distributed File system, where a networked cluster of physical drives redundantly stores data as a single logical volume. The adoption of data lakes allows big data platforms to store any type of data before that data are prepared to fit a specific type of analysis. This approach allows analysts to shape and refine the stored data to fit their needs without impairing other analyses (*8*). Data lakes require regular data governance to manage and maintain data integrity; a data lake can become a data swamp with unorganized and unusable data that lack clear identifiers or metadata information (*9*).

GIS datasets allow the integration of spatial and attribute data into large traffic datasets. For example, the HPMS datasets produced annually by state DOTs are submitted in a geospatial format, allowing users to spatially analyze the roadway network across different roadway statistics. GIS datasets rely heavily on the accuracy and quality of input data. If the spatial data are incomplete, inconsistent, or contain errors, they can lead to incorrect analysis and flawed results. Ensuring data quality and integrity requires careful data validation, cleaning, and updating processes. Failure to maintain data quality can undermine the reliability and usefulness of the spatial database (*10*).

TABLE 2 Attributes of Traffic Database File Types

Database Type	Advantages	Limitations
Flat files or spreadsheets	Simple methods and easy to use Fast data extraction and use	Larger datasets more difficult to process Adding new fields is more difficult Must know primary key(s)
Hierarchical data	Multiple associations to other datasets Fast data retrieval Adding or removing data fields is easy	Each association requires repetitive data Pointers require large data storage space Pointer path restricts access
Relational data	Flexible and easy to use with new queries Physical data storage can change Easy access with only minimal training Can add or remove relationships and data	Adding new relations can require reprocessing Sequential access is rather slow Easy to make logic mistakes with queries Disc storage affects process time
Cloud data	Accessibility from anywhere Cost savings compared to in-house data servers Ability to scale storage capacity Minimized risk of data loss in natural disasters and with hardware issues	Dependency on network and Internet connectivity Data security
Data lakes	Data stored in native format Allows shaping and refinement of stored data without impairing other analyses	Data quality issues Complexity of large volumes of data
GIS datasets and layers	Integration of spatial and attribute data into traffic data Visualization of patterns, key findings	Reliance on the accuracy and quality of input data

EMERGING TRENDS AND DRIVERS OF CHANGE

The growth in data quantities, cost of data storage, and the ways and means of providing traffic datasets to other agency and internal partners all continue to drive the change of managing large traffic datasets. By 2025, 470 million connected vehicles, each producing roughly 25 GB of data per hour, may be on highways around the world (*13*). Likewise, the cost of data storage has decreased over the years. For example, the cost of storing a TB in a disk format was $70 per TB in 2009. In 2022, that cost had decreased to $14.30 per TB (*11*) and by mid-2025 there are predictions the cost could be $10 per TB (*12*).

Traffic data collection programs continue to be essential in providing traffic data resources to an agency's departments, including the performance and safety divisions. FHWA now allows state DOTs to submit volume, class, and speed data from continuous counters at smaller time increments, as small as 1/100th of a minute. Having speed data by class can give users speed variances to calculate speed differentials used in safety analysis. Bridge departments have also seen the need for data at smaller increments than the traditional one-hour data submittals state DOTs upload onto FHWA's Travel Monitoring Analysis System (TMAS).

The Transportation Research Informatics Platform (TRIP) is an informatics-based system designed to manage massive amounts of transportation data and provide researchers an efficient way to conduct analytics on big data. The objectives of TRIP include creating the ability to handle massive amounts (e.g., terabytes) of transportation data; to use open-source technologies and tools to ingest, store, align, and process data; to accept structured, semi-structured, and unstructured datasets from any source; to provide an efficient way to query data without in-depth knowledge of metadata; to integrate with open-source and consumer off-the-shelf analytics products; and to provide visualization tools to offer greater insights into data. TRIP specifically incorporates large traffic datasets into its crash/incident management platforms (*14*).

The Regional Integrated Transportation Information System (RITIS), developed by the CATT Laboratory at the University of Maryland, is an automated data sharing, dissemination, and archiving system. It includes many performance measure, dashboard, and visual analytics tools that help agencies to gain situational awareness, measure performance, and communicate information among agencies and to the public. Use cases of RITIS include monitoring and managing known events, measuring corridor performance, and predicting holiday travel patterns (*15*).

The freeway Performance Measurement System (PeMS) from Caltrans collects real-time traffic data from sensors and generates performance measures of vehicle miles traveled, hours traveled, and travel time (*16*).

GAPS IN PRACTICE AND KNOWLEDGE

Traffic monitoring systems continue to integrate multiple equipment and data sources into their programs, creating large traffic datasets. As ITS data, probe data, and traffic monitoring data become integrated into searchable databases, agencies will need to improve their ability to manage these large datasets. Consistently verifying equipment is key to managing large traffic datasets. Potential errors in using continuous and short-duration sites as ground truth data to probe-based data could occur if equipment is not calibrated, field checked, and verified on a consistent basis.

Structured and unstructured data are available from traffic monitoring stations. In traffic monitoring terms, a structured data set is one that conforms to existing national, state, or local procedures. Local counts, counts conducted to support individualized research, and counts for another individualized need often do not contain the structured aspects that support use in a large database. Individualized data collection and counting for a specific project occurs frequently. The manual counts performed at rest areas and ramps, by private industry for local needs, or by agencies to support new construction or corridor improvements are often not used or inserted into larger systems or datasets. These data also sometimes lack the ability to be fully used for annualization applications but can serve as a QC dataset to the larger annualization datasets.

In order to fully use the rich and large traffic datasets available today, documentation of the data dictionary and governance, key fields, database relationships, and data availability needs to be increased. Agencies may also not use the same definitions and standards for traffic data collection that conform to existing national, state and local procedures. With proper documentation and consistency of definitions and standards, large datasets can be used more fully and can be managed more easily.

CURRENT AND PROPOSED NATIONAL RESEARCH INITIATIVES

Authors for this chapter determined there is no current or proposed research relevant to this E-Circular.

REFERENCES AND OTHER RESOURCES

1. *Traffic Monitoring Guide*, Federal Highway Administration, U.S. Department of Transportation, December 2022. https://www.fhwa.dot.gov/policyinformation/tmguide/. Accessed Aug. 28, 2023.
2. *AASHTO Guidelines for Traffic Data Programs, 2nd Edition*. AASHTO Publications Order Department, Atlanta, GA, 2009. https://bookstore.transportation.org/item_details.aspx?ID=1393.
3. Pecheux, K. K., B. B. Pecheux, and G. Carrick. *NCHRP Research Report 904: Leveraging Big Data to Improve Traffic Incident Management*. Transportation Research Board, Washington, DC, 2019. http://www.nap.edu/25604. Accessed Aug. 28, 2023.
4. Gettman D., A. Toppen, K. Hales, A. Voss, S. Engel, and D. El Azhari. Integrating Emerging Data Sources into Operational Practice—Opportunities for Integration of Emerging Data for Traffic Management and TMCs. FHWA-JPO-18-625 (Final Report). Federal Highway Administration, U.S. Department of Transportation, Washington, DC, 2017.

5. The Four Vs of Big Data. Business Analytics and Dashboards Cloud Services Internet of Things. May 21, 2015. http://www.plasmacomp.com/blogs/the-four-vs-of-big-data. Accessed Aug. 28, 2023.
6. Lukasik, D., Hale, D., Ma, J., Shibley, P., Malone, T., Chandler, A., Cleary, C., Matout, N., and Adebisi, A. (2020). *Enhancing Active Transportation and Demand Management (ATDM) with Advanced and Emerging Technologies and Data Sources* (FHWA-HOP-19-010). Leidos. https://rosap.ntl.bts.gov/view/dot/52797.
7. *Integrating Emerging Data Sources into Operational Practice: State of the Practice Review*; FHWA JPO-16-424. U.S. Department of Transportation. 2016.
8. Pecheux, K. K., B. B. Pecheux, G. Ledbetter, J. D. Schneeberger, J. Hicks, B. Burkhard, and M. Campbell. NCHRP *Web-Only Document 282: Framework for Managing Data from Emerging Transportation Technologies to Support Decision-Making*. Transportation Research Board, Washington, DC, 2020.
9. What Is a Data Lake? Pros and Cons of Data Lakes, Masterclass. https://www.masterclass.com/articles/what-is-a-data-lake. Accessed August 28, 2023.
10. Spatial Post, Advantages and Disadvantages of Spatial Database | Explained. https://www.spatialpost.com/advantages-disadvantages-of-spatial-database/. Accessed Aug. 28, 2023.
11. Historical Cost of Computer Memory and Storage, Our World in Data; https://ourworldindata.org/grapher/historical-cost-of-computer-memory-and-storage. Accessed Aug. 28, 2023.
12. Mellor, Chris. Blocks and Files; Backblaze predicts HDD prices will fall to 1¢/GB by mid-2025. https://blocksandfiles.com/2022/12/12/backblaze-sees-hdd-pricing-falling-to-a-penny-a-gig. Accessed Aug. 28, 2023.
13. Connected Vehicles Shift an Industry, Deloitte. https://www2.deloitte.com/us/en/pages/advisory/articles/connected-vehicles-shift-an-industry.html. Accessed Aug. 28, 2023.
14. Applications of Knowledge Discovery in Massive Transportation Data: The Development of a Transportation Research Informatics Platform (TRIP); FHWA-HRT-19-008, U.S. Department of Transportation.
15. Regional Integrated Transportation Information System, RITIS Introduction. https://ritis.org/intro. Accessed Aug. 28, 2023.
16. Caltrans Performance Measurement System, https://pems.dot.ca.gov/, Accessed Aug. 28, 2023.
17. Exploring Non-Traditional Methods to Obtain Vehicle Volume and Class Data, Transportation Pooled Fund. https://www.pooledfund.org/Details/Study/636.
18. Travel Survey State of the Practice, FHWA NHTS Report, March 2023. https://nhts.ornl.gov/assets/NextGen%20NHTS_State%20of%20Practice_032423.pdf

Traffic Monitoring and Performance Measures

Jijo K. Mathew
Purdue University

Lin Zhang
Elite Transportation Group

Shraddha Sagar
University of Florida

Dhivyabharathi Bhaskaran
Trinity College Dublin

Mena Lockwood
Virginia Department of Transportation

Christopher Vaughan
North Carolina State University

Scott Vockeroth
Alaska Department of Transportation

Lawrence A. Klein
Klein and Associates

STATE OF THE PRACTICE

The concepts of traffic monitoring and performance measurement are closely related and highly dependent on each other. Traffic monitoring is a long-standing program within state DOTs and plays a vital role in gathering information to understand current and past performance of the transportation system and to predict future performance. Despite the proliferation of new private-sector probe data sources, monitoring of traffic volumes and interpretation of the data will continue to be a critical need for state DOTs and other transportation agencies. Barriers to successful monitoring and interpretation include the need for evaluation of private-sector data and the ongoing need to ensure data quality.

Legislation and Requirements

The Moving Ahead for Progress in the 21st Century (MAP-21) Act of 2012 and 2015's Fixing America's Surface Transportation (FAST) Act legislation introduced performance management into the Federal Highway Administration (*1, 2*). State DOTs and MPOs must now report on and make progress toward targets they set against a number of national performance measures. The objective of this new aspect of the federal programs is to focus federal funds on achieving national goals, increasing accountability and transparency, and improving investment decision-making through performance-based planning and programming.

MAP-21 and FAST Act performance areas include safety, infrastructure condition, system reliability, freight movement and economic vitality, congestion reduction, and environmental sustainability. The implementation of the legislation aims to change the way state DOTs and MPOs conduct transportation planning. Resource allocation decisions will be based on outcome-based measures. MAP-21 and the FAST Act are clear in their intent to require a performance management approach to federal investment in the nation's highways. Although performance management has gained momentum among state and local transportation agencies for several years, and many of them have already implemented exemplary programs, the rulemaking has accelerated the process significantly.

The FAST Act defines four main categories of measures: safety, pavement and bridges, mobility, and air quality. Traffic data (volumes) are required to support both safety and mobility measures. The safety measures to be reported by states and MPOs include the number of fatalities on all public roads, rate of fatalities per 100 million vehicle miles traveled (VMT) on all public roads, number of serious injuries on all public roads, and rate of serious injuries per 100 million VMT on all public roads. The key data items are total VMT on all roads and annual number of fatalities in crashes involving a motor vehicle. The mobility measures include percentage of person-miles traveled that are reliable and annual hours of peak hour excessive delay (PHED) per capita.

A congestion management process (CMP) is a systematic approach for managing congestion that provides accurate, up-to-date information on transportation system performance and that assesses alternative strategies for congestion management that meet state and local needs (*3*). A CMP is required in metropolitan areas with population exceeding 200,000, known as Transportation Management Areas (TMAs). CMPs are required to be developed and implemented as an integrated part of the metropolitan transportation planning process. The CMP uses an objective-driven performance-based approach to planning for congestion management. This approach involves screening strategies which use objective criteria and rely on system performance data, analysis, and evaluation.

Performance Measures

According to *NCHRP Report 706*: *Uses of Risk Management and Data Management to Support Target-Setting for Performance-Based Resource Allocation by Transportation Agencies*, performance measures are a set of metrics used by organizations to monitor progress toward achieving a goal or objective (*4*). The Federal Highway Administration (FHWA) Performance Measures and System Monitoring program defines performance measures as indicators of how well the transportation system is performing (*5*). The criteria for selecting performance measures often include feasibility, policy sensitivity, ease of understanding, and usefulness in actual decision-making.

Performance management is a business process that links organization goals and objectives to resources and results. Performance measures, used along with well-defined and well-communicated targets, provide transparency and clarity to the resource allocation decision-making process. Performance-based resource allocation in any organization relies on the availability of timely, accurate, high-quality data easily accessible through a framework known as data management. Such a program usually includes the functions of data collection, analysis, and reporting. In the case of a DOT, performance measures may include traffic volume, speed, travel time/delay, reliability, safety, and pavement conditions.

Performance Measures and Traffic Data

The FHWA TMG states that traffic counts are fundamental to almost every task a highway agency performs and are critical to a comprehensive performance measurement system (*6*). The timely delivery of high-quality data can serve as a critical framework for effective decision-making. The ability to describe traffic volumes and vehicle types on any link in the transportation system reflects positively on the agency's ability to effectively perform its responsibilities and manage its budget.

FHWA defines Transportation Performance Management (TPM) as a strategic approach that uses system information to make investment and policy decisions to achieve national performance goals (*7*). State DOTs are mostly concerned with performance in the areas of safety, mobility, preservation, and economic competitiveness. Traffic data and their associated performance measures, including AADT and VMT, are a cornerstone of mobility measures.

Data Sources

In the past, traffic monitoring and performance measures relied on traditional data sources, including detector stations (either temporary or permanent), survey data, cameras, vehicle occupancy counts, and travel time runs. With the proliferation of new private-sector probe data sources such as INRIX, HERE, and StreetLight Data, more data sources are available for local and regional traffic monitoring. More discussion of these data sources is in Probe Data for Traffic Volume Estimation.

Data Visualization

Traffic statistics are a crucial component of mobility measures and can be visualized in a variety of ways. Many states have developed dashboards to display their measures. A comprehensive resource for details on the rulemaking, requirements, and noteworthy practices of some states is FHWA's TPM website (*8, 9*).

The TPM Digest (*10*) includes the latest information concerning online state dashboards, performance reports, mobility, performance-based planning, safety, events, workshops, webinars, research, and innovation. Noteworthy practices are included for the following topics: getting started, data collection and management, target setting, project prioritization and decision-making, reporting, and collaboration. External links are also provided to TPM dashboards for the states of Alaska, Colorado, Delaware, Idaho, Iowa, Kansas, Michigan, New Hampshire, North Carolina, Ohio, Oregon, South Carolina, Virginia, Washington, West Virginia, and Wisconsin, along with the cities of Seattle; Washington, DC; and San Francisco.

Traffic Safety

Besides traditional safety measures using crash rates per 100 million VMT, cameras or light detection and ranging (LiDAR) now extract near-miss crash data and imagery as a surrogate metric for traffic safety. The trajectory of each road user (e.g., vehicle, pedestrian, bicyclist) can be extracted from roadside cameras or LiDAR via data processing algorithms, and the time to collision or even sudden braking can be captured via this technology.

Performance Measure with Missing Data

Several methods are available to estimate AADT when data are missing for at least one hour or at least one day. A study compared various methods and found that the Highway Policy Steven Jessberger-FHWA and Battelle (HPSJB) Method provided more accurate and reliable AADT under various missing data scenarios (*11*).

STATE OF THE ART

Traffic monitoring and performance measures typically rely on traditional data sources, such as inductive loops and other fixed sensors, survey data, cameras, vehicle occupancy counts, and travel time measurements using probe vehicles. With the introduction of private-sector probe data sources, such as INRIX, HERE, and StreetLight, more data sources have become available to augment or replace traditional data sources that are typically location-constrained and unable to provide continuous spatial coverage. Most recently, data generated and wirelessly exchanged among connected and automated vehicles (CAVs) have demonstrated potential use in traffic monitoring and performance measures (*12–15*), with more research and development ongoing. Using CAV data for traffic monitoring is covered in more detail in the section Integrating Traffic Monitoring with Connected Vehicle Data.

Automated traffic signal performance measures (ATSPM) for arterials with signalized intersections have been developed and implemented over the last five years. ATSPM is defined as “a suite of performance measures, data collection, and data analysis tools to support objectives and performance-based approaches to traffic signal operations, maintenance, management, and design to improve the safety, mobility, and efficiency of signalized intersections for all users (*16*).” ATSPM provides data analytics tools and approaches that automatically collect and convert high-resolution traffic signal data into actionable performance measures such as traffic volumes, travel times, arrivals on green, progression ratio, split failures, and device and communication uptime (*17*).

EMERGING TRENDS AND DRIVERS OF CHANGE

With the recent advancements in technology and communication standards, it is now possible to have a constant feed of connected vehicle and truck data. Nearly 1 in every 20 vehicles in the United States is estimated to provide some form of telematics-based connected vehicle data through one of the commercial data providers (*18*). Several connected vehicle data providers and tools have emerged in recent years (*19*). Table 3 illustrates a consolidated summary of the applications and performance measures that can be derived from several emerging data types in various areas of the transportation system.

Segment-based reporting of crowdsourced data has been available for over a decade. Now the current frontier is the trajectory-based data available every 1–3 seconds with attributes including speed, geolocation, heading, and timestamp. Data providers are beginning to enhance trajectory data further by providing event data such as hard braking, hard acceleration, and pothole detection, as well as road condition, including friction and pavement roughness. In December 2022, approximately 503 billion connected vehicle records were reported in the United States through one of the data providers (*20*).

TABLE 3 Emerging Data Types and Performance Measures

Category	Data Type	Applications/Performance Measures
Mobility	Probe Data (Waypoint + Segment)	Real-time incident detection Queue length and duration Congestion monitoring Speed, travel time and reliability Route diversions
Safety	Vehicle Telemetry Data	Hard braking and hard acceleration events Roadway friction and roughness
Origin-Destination	Probe Data (Waypoint)	Origin-destination studies
Freight Movement	Probe Data (Waypoint + Segment) Commodity Flow Data	Speed, travel time, and volume Parking availability and use Commodity movement
Traffic Signals	Probe Data (Waypoint + Segment) Camera/Video Data	Levels of service Arrivals on green Split failure Spillback or downstream blockage Turning movement counts Pedestrian activity
Economic/Transactional	Micromobility Data Transit Data	Origin-destination Speed, travel time, and reliability Ridership and costs
Forecasting	Modeled and Forecasted Data	Travel behaviors Urban planning Traffic volumes and transit ridership

Figure 3 shows a commonly used traffic heatmap tool (*21*) developed using enhanced trajectory data for incident detection, traffic monitoring, and after-action assessment. The plot depicts an incident involving a primary crash (callout P) and a secondary crash (callout S) on interstate I-70 in Indiana. This incident impacted traffic for nearly 2.5 hours (callout i), which resulted in a 12-mile queue (callout ii) and road closure for approximately 30 minutes (callout iii). Hard-braking events (red dots) between the congested and uncongested regimes indicate the critical areas where drivers decelerated to avoid back-of-queue or rear-end collisions. Diverting trips (callouts D1-D9) between the exits can also be extracted to understand the impact on various diversion routes during incidents.

Emerging connected vehicle data have also been instrumental in work zone related safety measures (*22*). Maryland DOT developed several analytical and visualization tools to understand the mobility impact of work zones, including amount of queueing and delays (*23*). The bottleneck ranking analysis tool estimates the length and duration of the bottlenecks, as well as the delay resulting from the bottleneck. These estimates allow agencies to identify key trends and issues in work zones. Indiana DOT used commercially available hard braking event data to identify key safety and design issues on work zones (*24*).

Other emerging tools and performance measures include:

- **Tethered Drones/Unmanned Aerial Vehicles (UAVs)**: The use of tethered drones for traffic operations has been gaining popularity. Several studies and data providers have outlined the feasibility of using tethered drones to estimate several performance measures including AADT, level of service, gap acceptance, and queue lengths (*25–27*).

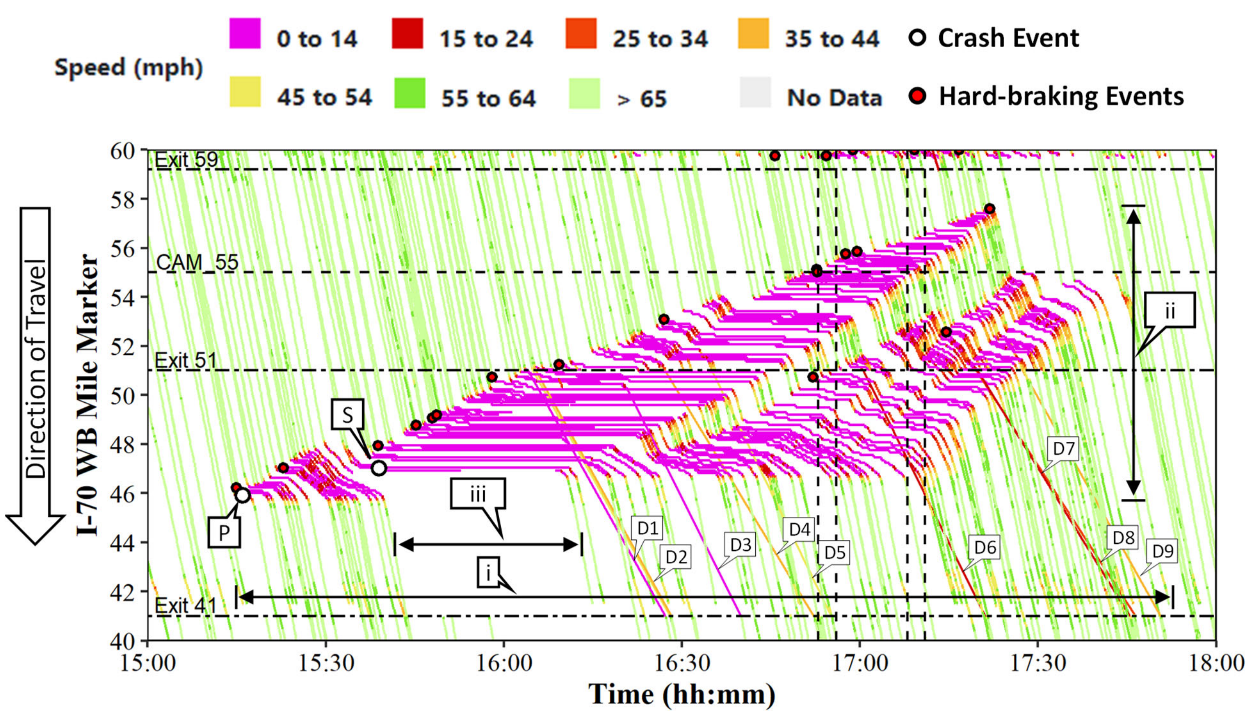

FIGURE 3 Traffic heatmap from vehicle trajectories colored by speed overlaid with crash events and hard-braking events.

- **Queue Warning Trucks and Digital Alerts**: A few states (*28–29*) have implemented a “protect the queue” program in work zones by deploying queue protection trucks or queue warning trucks that move with the queue to alert motorists of upcoming queues and slowdowns. These trucks are also equipped with technology from digital alert providers, such as HAAS Alert and iCones, that sends alerts to Waze, Apple Maps, and other in-cab navigation systems to complement road signs. Studies have found that this system was effective in reducing hard-braking events and crash risks (*30*).
- **Stationary LiDAR Equipment**: Enhancing pedestrian and traffic safety at intersections is a critical objective for agencies. Researchers have used LiDAR sensors to develop innovative tools that collect and study the performance of traffic and pedestrian safety at intersections (*31–33*).

GAPS IN PRACTICE AND KNOWLEDGE

With new traffic data sources steadily emerging, the potential has grown for traffic engineers and planners to have more insight than ever into the performance of the transportation system. However, most modern data sources involve some type of probe vehicles. Validation of such data sources is lacking due to the presence of minimal counting stations, especially on local roads. Also, low traffic volume local roads have smaller probe vehicle penetration, leading to unreliable data and performance measures. A lack of counting stations discourages the data fusion strategies that are popular for improving volume counts. As road crashes, especially for vulnerable road users, rise in the United States, crash rates are critical for identifying trends and finding solutions. Hence, there is a need to study the factors affecting data fusion techniques targeted at improving the accuracy of AADT on local roads.

Agencies consider reductions in delay when prioritize projects, as these reductions improve mobility for drivers and freight haulers. However, a significant number of current programs seek to improve peak hour traffic flow using measures such as peak hour profiles or, more generally, hourly profiles of traffic volumes. Prioritization based on hourly profiles is therefore needed for return on investment calculations. Temporary or limited counting stations, typically used for hourly profile factoring, may not accurately reflect all conditions. Hourly profiles can be site-specific as they are heavily influenced by land use and other causes over time, such as unusual situations like COVID-19.

The upsurge of teleworking and remote learning greatly affected travel patterns and peak traffic times, resulting in shifts in hourly profiles. Many locations observed a much lower morning peak period, a slight increase in midday, and similar evening peak hour percentages. More robust data sources are needed to understand these changes. Accurate hourly profiles are not readily available for much of the arterial network but are critical for project prioritization, project design, signal timing, and performance monitoring. Leveraging data from ATSPM systems can provide greater network coverage but is not sufficient to fill all the gaps.

Geospatial data analytics play an important role in combining data sources. Conflation techniques are essential to merge two geospatial data sets, as most performance management program and data analyses require data from different sources. For example, delay per vehicle becomes more useful when it is combined with volume; this combination can represent a total delay over a specified period or over multiple locations, such as a series of road segments or corridors. Conflation, which is needed to combine these data sets, is time consuming and demands specialized skill sets that many public transportation agencies do not have. Automating the conflation process would help reduce the needed time and resources. Inconsistency in data sources, naming conventions, map versions, and facility names, along with a lack of clear conflation algorithms and the need for (extensive) manual work are some of the roadblocks in successfully implementing conflation in practice.

Many state agencies locate their data (asset, infrastructure, traffic characteristics and flow) on their own linear referencing systems (LRSs). Conflation to those LRSs is needed to use the agencies' data in a meaningful way. Therefore, research that offers guidelines and practical examples for conflating different types of data sets, especially newer ones derived from CAVs, would be beneficial.

To spatially aggregate datasets, traffic message channels carry out segmentation created and maintained by third-party entities. Neither data providers nor data users have much input in the process. MAP-21 mandates the use of the National Performance Management Research Data Set (NPMRDS) traffic message channels segmentation for reporting federal system performance measures. To attain better spatial aggregation, data providers like INRIX, TomTom, and HERE have developed their own proprietary road network link systems. INRIX developed its own XD network, which provides greater coverage and typically shorter segments to allow for refined analysis. For analysis like queue length estimation, shorter segment lengths are relevant. Some probe providers do allow for dynamic sub-segmentation, but this needs to be requested in advance and could be subject to additional fees. Segmentation should be developed based on logical ground control points within which geometric conditions vary minimally. Overall, guidelines for segmentation are lacking and standards for optimum segment length have not been established.

In the pursuit of exploring alternate data sources, probe vehicle data sources are becoming popular, but consistent validation programs are lacking for some products (like volume) while others (like speed) have been more thoroughly vetted. Probe-based speed and travel time data products have been validated successfully for many years. The Eastern Transportation Coalition started a program to validate volume data from probe data and is interested in building an Origin-Destination Data Validation program. Practical tools to validate volume, origin-destination, connected vehicle waypoint, and event data are needed for states and localities to gain trust in probe vehicle data. With more trust, they can incorporate these data into activities such as queue length and duration estimation, performance measure calculation, and many other applications.

Traditional data collection methods (road sensors, side-fire and overhead radar, and video detection systems) provide accurate lane-level volume and occupancy. In some cases, vehicle classification provides insights into micro-level facility interactions and allows detailed evaluations of managed lanes, weaving maneuvers, speed harmonization among lanes, safety analysis, and more. Emerging probe volume data sources do not typically offer such lane-level spatial granularity or ramp-level data. In addition, better turning movement counts for intersection analysis, volume, and speed data at on-ramp locations are also needed for studies such as incident detour and ramp metering analyses. Though these alternative probe data sources offer many advantages over traditional data collection methodologies, the lack of lane-level granularity is problematic.

ATSPMs offer additional opportunities to collect volume and turning movement count data for statewide traffic monitoring needs. However, ATSPMs and signal detection systems typically do not provide lane-level data as that represents more granularity than signal systems require. Research is needed to develop algorithms that could fill gaps in lane-level or even in TMC-level data. Estimating lane-specific volumes is more difficult than estimating volumes for the entire approach. However, it would be helpful to investigate the accuracy of lane-specific volume estimates, whether and how these volumes can be developed using existing data (e.g., event-based data from signals), and the most effective configurations of signal data collection systems.

Data bias is another crucial element to consider while integrating data from multiple sources. The optimum proportion of data from any given source during data fusion needs to be established and standardized. Policy makers should be aware of potential data bias, particularly if connected vehicles are a significant component of the raw data. Connectivity is a feature of newer-model vehicles that are more common among communities with higher incomes. Use of connectivity-based data could therefore emphasize the needs of these drivers over those who drive older vehicles.

Research on CAVs has gained attention and pilot implementations are now prevalent. However, more information needs to be documented and available to public agencies. A wide variety of test cases under different scenarios are needed, especially those involving vulnerable road users. CAV trajectory data would help to identify areas with harsh braking and acceleration that could be indicative of safety hot spots. Ramp metering, dynamic detour rerouting, variable speed limits, and other active traffic management strategies could be evaluated or simulated in a more automated, data-driven fashion. The potential applications of using vehicle trajectory data from CAVs for traffic monitoring and performance measurement, such as supplemental traffic counts, traffic mobility calculations, and safety measures, are not well studied yet. Handling trajectory data is complex, and agencies often may not have the skill sets and the resources for data storage and computation.

CURRENT AND PROPOSED NATIONAL RESEARCH INITIATIVES

This section describes existing and proposed research to address gaps in practice and knowledge for pavement engineering applications related to traffic monitoring.

Current Research

Performance Measures

An ongoing study sponsored by the Kentucky Transportation Cabinet uses advanced machine learning techniques and enhanced probe vehicle data to generate more precise traffic estimates, particularly for unmonitored and local roads. It is titled Estimating Traffic Volume Using Ubiquitous Probe Vehicle Data.

Use of Measures in Traffic Data Programs

The FHWA Office of Operations developed a Data Business Plan Guide, aimed at providing support to state DOTs and transportation agencies in effectively managing and governing their traffic data programs. It encompasses aspects such as speed and volume (*34*). This comprehensive guide can serve as a valuable resource in the development and implementation of programmatic performance measures pertaining to traffic monitoring.

In addition, there is a need for research to investigate how state DOTs currently use performance measures to efficiently manage and optimize their traffic data programs. By studying these practices, valuable insights can be gained to enhance the effectiveness and efficiency of traffic data management strategies and processes.

Setting Performance Targets

Numerous states and regions have made significant progress in the implementation of performance measurement systems. A few states, such as Virginia (*35*), have conducted research to establish specific performance standards or targets beyond traditional areas like asset management and safety. However, several states that have given limited attention to setting targets. The implementation of MAP-21 and the FAST Act drives much of the surge in activity in setting targets, posing technical challenges and necessitates increased coordination and cooperation among agencies. Furthermore, the coordination of targets between DOTs and MPOs requires clear guidance and direction to ensure effective collaboration. Coordination across agencies benefits from innovative approaches to ensure that performance targets are properly aligned and enable efficient and impactful decision-making in transportation planning and management.

Proposed Research

This section summarizes general issues in need of research followed by a list of specific ideas for future research proposals.

Technical Issues—Conflation and Segmentation

To effectively address technical issues related to conflation and segmentation in traffic monitoring, it is crucial to focus on state-specific challenges. The NPMRDS offers a standardized and consolidated dataset that incorporates information from probe vehicles, traffic sensors, and other relevant sources. Although the NPMRDS mitigates many conflation and segmentation issues, further research and analysis are necessary. Conducting a comprehensive study, potentially using a standardized approach, would help identify and address these technical challenges.

One proposed study would be to develop a comprehensive guide on the application of spatial segmentation for travel time reliability measures.

Tools

Research and development efforts are needed to create additional tools to support congestion management processes. These tools can assist transportation agencies in analyzing and managing traffic congestion more effectively. Additionally, there is a need for improved tools specifically designed for forecasting heavy vehicle traffic, as these forecasts play a crucial role in freight transportation planning and operational decision-making.

Training

Research and development of skill sets and educational programs focused on data analytics are necessary to enhance the capabilities of transportation professionals.

Nonmotorized

Research, synthesis, and development of capabilities for nonmotorized data collection and estimation, such as bicycle and pedestrian data, are essential. Additionally, research on nonmotorized level-of-service measures can help evaluate the quality and convenience of facilities for pedestrians and cyclists. One ongoing project that aligns with this objective is a research effort sponsored by MnDOT that explores the use of mobile-device data to predict pedestrian and bicyclist flows on specific roads (*36*).

Vehicle Occupancy (Number of People in a Vehicle)

Research and methods are needed to accurately measure and estimate vehicle occupancy. An ongoing project sponsored by the University Transportation Centers program is titled "Pilot Application of Biometric-Based Vehicle Occupancy Detection on Managed Lanes for Congestion Reduction." The project aims to use biometric-based vehicle occupancy detection technology for congestion reduction on managed lanes (*37*). It addresses the need to optimize travel efficiency by incentivizing carpooling and vanpooling through accurate occupancy counting.

Forecasting

Research focusing on identifying the key variables for accurate forecasting is needed. This includes considering factors such as overall mobility, passenger versus freight transportation, and differentiating between types of vehicles and commodities.

Data Sources

It is important to harness the potential of emerging data sources and technologies to enhance traffic monitoring and improve performance measures. These sources include probe data from companies like INRIX and Waze, as well as data from LiDAR, cameras, and computer vision systems, all of which offer extensive traffic operations information. Recently completed and ongoing studies are actively exploring the use of these data sources. These include:

- *NCHRP Synthesis 611: Use of Probe Data for Freight Planning and Operations,* https://nap.nationalacademies.org/catalog/27249/use-of-probe-data-for-freight-planning-and-operations
- Investigation of LiDAR sensing technology to Improve Freeway Traffic Monitoring, https://rosap.ntl.bts.gov/view/dot/67550
- NCHRP Project 08-157, "Best Practices for Data Fusion of Probe and Point Detector Data," https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=5135.
- Integration of the Lane-specific Traffic Data Generated from Real-time CCTV Videos into INDOT's Traffic Management System, https://rosap.ntl.bts.gov/view/dot/66097

The emergence of CAVs has introduced new possibilities for traffic monitoring and performance measures. While ongoing research endeavors are using these data sources, further investigations are required to enhance our understanding and maximize the potential of these data sources. This topic, including the potential benefits of this data source and the obstacles to using these data, is discussed in further detail in Integrating Traffic Monitoring with Connected Vehicle Data and Probe Data for Traffic Volume Estimation.

In addition, it is crucial to conduct an analysis of potential biases present in third-party vendor data. A proposed research initiative aims to address the issue of Equity Bias in Vendor Data Used for Transportation Decision-Making. This study aims to investigate the potential equity bias introduced by third-party vendor APIs through variations in their geographic coverage, which could have implications for planning, safety, and other aspects of projects.

Enhancing Traffic Estimation on Unmonitored Roads Using Machine Learning Techniques and Probe Vehicle Data

Modern data sources primarily rely on probe vehicles, but the limited number of counting stations on local roads poses a challenge to the widely used data fusion strategies aimed at improving traffic volume counts. To enhance traffic estimation on these less monitored roads, this research would test advanced machine learning techniques while incorporating probe data, focusing on studying and improving the accuracy of AADT estimations on low-volume roads.

Best Practices in Effective Use of Performance Measures

The FHWA Office of Operations has developed a Data Business Plan Guide to efficiently manage and optimize traffic data programs. Conducting a synthesis would identify best practices in effective construction and use of performance measures. Further, this project aims to investigate how state DOTs currently use these measures and seeks to enhance the effectiveness and efficiency of traffic data management strategies and processes.

Congestion Management and Incident Detection Tools

This research idea seeks to use new data and tools to better predict congestion and related incidents. Using new data like vehicle telematics, probe data, and LiDAR, this research could highlight how to develop tools to more quickly detect congestion and related incidents so that incident response and other mitigation measures could be taken more swiftly.

Leverage Emerging Data Sources to Enhance Traffic Monitoring and Performance Measures

The focus of this idea is to leverage emerging data sources like trajectory data, LiDAR, cameras, and computer vision to enhance traffic monitoring and performance measures. This includes exploring innovative technologies such as tethered drones and queue warning trucks, both of which represent new frontiers in traffic data collection and analysis. These sources of data may provide advanced information to significantly improve the accuracy and effectiveness of traffic monitoring systems.

Re-Examine FHWA Definition of Peak Hours Post-COVID

The initiative aims to re-examine the FHWA definition of peak hours in the post-COVID era. This involves studying the changing patterns of peak hour traffic due to the pandemic's impact, understanding the evolving dynamics of traffic flow, and subsequently adjusting the definitions of peak hours and related performance measures to reflect these new realities.

REFERENCES AND OTHER RESOURCES

1. Moving Ahead for Progress in the 21st Century (MAP-21), Federal Highway Administration. https://www.fhwa.dot.gov/map21/. Accessed Apr. 26, 2023.
2. Fixing America's Surface Transportation Act or the FAST Act, Federal Highway Administration. https://www.fhwa.dot.gov/fastact/. Accessed Apr. 26, 2023.
3. Congestion Management Process (CMP)—Organizing and Planning for Operations, FHWA Office of Operations. Washington, DC, 2011.
4. Cambridge Systematics Inc., Boston Strategies International, Inc., Gordon Proctor and Associates, and M. J. Markow. *NCHRP Report 706: Uses of Risk Management and Data Management to Support Target-Setting for Performance-Based Resource Allocation by Transportation Agencies*, Transportation Research Board of the National Academies, Washington, DC, 2011.
5. *Performance Measures and System Monitoring*. Federal Highway Administration, U.S. Department of Transportation. Accessed Apr. 26, 2023. https://ops.fhwa.dot.gov/plan4ops/focus_areas/analysis_p_measure/sys_monitoring.htm.
6. *Traffic Monitoring Guide*. Federal Highway Administration, U.S. Department of Transportation, Dec. 2022. https://www.fhwa.dot.gov/policyinformation/tmguide/. Accessed Apr. 26, 2023.
7. What Is TPM? Federal Highway Administration, U.S. Department of Transportation. https://www.fhwa.dot.gov/TPM/about/tpm.cfm. Accessed Apr. 26, 2023.
8. FHWA/DOT TPM–Related Links. Federal Highway Administration, U.S. Department of Transportation. https://www.fhwa.dot.gov/tpm/links_fhwa.cfm. Accessed Apr. 26, 2023.
9. *Transportation Performance Management*. Federal Highway Administration, U.S. Department of Transportation. https://www.fhwa.dot.gov/TPM/. Accessed Apr. 26, 2023.
10. *TPM Digest*. Federal Highway Administration, U.S. Department of Transportation. https://www.fhwa.dot.gov/tpm/resources/digest/current.cfm. Accessed Apr. 26, 2023.

11. *Assessing Roadway Traffic Count Duration and Frequency Impacts on Annual Average Daily Traffic (AADT) Estimation.* Publication FHWA-PL-015-008. Federal Highway Administration, U.S. Department of Transportation, 2014.
12. Xie, K., D. Yang, K. Ozbay, and H. Yang. Use of Real-World Connected Vehicle Data in Identifying High-Risk Locations Based on a New Surrogate Safety Measure. *Accident; Analysis and Prevention*, Vol. 125, pp. 311–319, 2019. https://doi.org/10.1016/j.aap.2018.07.002.
13. Kurkcu, A., and Ozbay, K., A Hierarchical Clustering Based Travel Time Estimation Model in a Connected Vehicle Environment. *J. of Traffic and Logistics Eng.*, 2017. Department of Civil and Urban Engineering and Center for Urban Science + Progress, Tandon School of Engineering, New York University, NY. https://doi.org/10.18178/jtle.5.2.54-59.
14. Hadi, M., V. Sisiopiku, and S. Srinivasan. *Performance Measurement and Management Using Connected and Automated Vehicle Data*. University of Florida Transportation Institute and U.S. Department of Transportation Office of Research, Development and Tech., 2021.
15. Gao, K., F. Han, P. Dong, N. Xiong, and R. Du. Connected Vehicle as a Mobile Sensor for Real Time Queue Length at Signalized Intersections. *Sensors*, Vol. 19, No. 9, p. 2059, 2019. https://doi.org/10.3390/s19092059.
16. Automated Traffic Signal Performance Measures web page. Federal Highway Administration, U.S. Department of Transportation. https://ops.fhwa.dot.gov/arterial_mgmt/performance_measures.htm. Accessed Apr. 29, 2023.
17. Lattimer, C. *Automated Traffic Signals Performance Measures.* Publication FHWA-HOP-20-002. Federal Highway Administration, U.S. Department of Transportation, 2020.
18. Sakhare, R. S., M. Hunter, J. Mukai, H. Li, and D. M. Bullock. Truck and Passenger Car Connected Vehicle Penetration on Indiana Roadways. *Journal of Transportation Technologies*, Vol. 12, No. 4, pp. 578–599, 2022. https://doi.org/10.4236/jtts.2022.124034.
19. Transportation Data Marketplace. The Eastern Transportation Coalition. https://tetcoalition.org/projects/transportation-data-marketplace/. Accessed Apr. 22, 2023.
20. Desai, J., J. Mathew, H. Li, R. S. Sakhare, D. Horton, and D. Bullock. National Mobility Analysis for All Interstate Routes in the United States: December 2022. *Indiana Mobility Reports*, 2022. https://doi.org/10.5703/1288284317591.
21. Mathew, J. K., J. C. Desai, R. S. Sakhare, W. Kim, H. Li, and D. M. Bullock. Big Data Applications for Managing Roadways. *ITE Journal*, Vol. 91, No. 2, pp. 28–35, 2021.
22. Data Driven Work Zone Process Review Fact Sheet. FHWA. https://ops.fhwa.dot.gov/wz/prtoolbox/pr_toolbox.htm. Accessed Apr. 22, 2023.
23. Probe Data Analytics Suite. CATT Laboratory. https://pda.ritis.org/suite/. Accessed Apr. 22, 2023.
24. Desai, J., H. Li, J. K. Mathew, Y.-T. Cheng, A. Habib, and D. M. Bullock. Correlating Hard-Braking Activity with Crash Occurrences on Interstate Construction Projects in Indiana. *Journal of Big Data Analytics in Transportation*, Vol. 3, No. 1, pp. 27–41, 2021. https://doi.org/10.1007/s42421-020-00024-x.
25. Ozguven, E. E., A. Karaer, M. Koloushani, R. Moses, M. A. Dulebenets, and T. Sando. *Feasibility Analysis of Real-Time Intersection Data Collection and Processing Using Drones*. Report 2021-09-01, 2021.
26. Kinero, A. Exploring the Use of Drones for Conducting Traffic Mobility and Safety Studies. UNF Graduate Theses and Dissertations, 2021.
27. Alden, A. S., H. Park, J. Coggin, and Virginia Tech Transportation Institute. *Developing a Plan for Using Unmanned Aerial Vehicles for Traffic Operations Applications in Virginia*. Publication FHWA/VTRC 22-R24. 2022.
28. Protect the Queue. https://www.tn.gov/tdot/traffic-operations-division/transportation-management-office/protect-the-queue.html. Accessed Apr. 24, 2023.

29. Whaley, W. Cabinet Remind Drivers 'Slow Is Safe When Traveling through Work Zones,' 2023. Gov. Beshear, Transportation https://www.wbko.com/2023/04/17/gov-beshear-transportation-cabinet-remind-drivers-slow-is-safe-when-traveling-through-work-zones/. Accessed Apr. 24, 2023.
30. Sakhare, R. S., J. C. Desai, J. Mahlberg, J. K. Mathew, W. Kim, H. Li, J. D. McGregor, and D. M. Bullock. Evaluation of the Impact of Queue Trucks with Navigation Alerts Using Connected Vehicle Data. *Journal of Transportation Technologies*, Vol. 11, No. 04, pp. 561–576, 2021. https://doi.org/10.4236/jtts.2021.114035.
31. Zhao, J., Y. Li, H. Xu, and H. Liu. Probabilistic Prediction of Pedestrian Crossing Intention Using Roadside LiDAR Data. *IEEE Access*, Vol. 7, pp. 93781–93790, 2019. https://doi.org/10.1109/ACCESS.2019.2927889.
32. Li, T., S. Kothuri, X. (Terry) Yang, University of Texas at Arlington, Portland State University, University of Utah, and National Institute for Transportation and Communities. *Pedestrian Behavior Study to Advance Pedestrian Safety in Smart Transportation Systems Using Innovative LiDAR Sensors*. Publication NITC-RR-1393. 2023.
33. Tarko, A., K. Ariyur, M. Romero, V. Bandaru, and C. Lizarazo. TScan: Stationary LiDAR for Traffic and Safety Studies—Object Detection and Tracking. *JTRP Technical Reports*, 2016. https://doi.org/10.5703/1288284316347.
34. Roadway Safety Data Program, Federal Highway Administration, U.S. Department of Transportation. https://safety.fhwa.dot.gov/rsdp/manage.aspx. Accessed Oct. 2, 2017.
35. Babiceanu, S., and S. Lahiri. Methodology for Predicting MAP-21 Interstate Travel Time Reliability Measure Target in Virginia. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2676, pp. 253–266, 2022. https://doi.org/10.1177/03611981221083290.
36. Stern, R., Levin, M., and Lindsey, G. *Mobile-Device Data, Non-Motorized Traffic Monitoring, and Estimation of Annual Average Daily Bicyclist and Pedestrian Flows*, Center for Transportation Studies, University of Minnesota. https://www.cts.umn.edu/research/project/mobile-device-data-non-motorized-traffic-monitoring-and-estimation-of-annual-average-daily-bicyclist-and-pedestrian-flows. Accessed Aug. 23, 2023.
37. Xiaopeng L. Pilot Application of Biometric-Based Vehicle Occupancy Detection on Managed Lanes for Congestion Reduction. https://trid.trb.org/View/1885418. Accessed Aug. 23, 2023.

Pavement Engineering Applications

BARBARA OSTROM

WSP

OLGA SELEZNEVA

Applied Research Associates

STATE OF THE PRACTICE

Pavement engineering uses traffic parameters and traffic counting data for pavement analysis, design, and management in the following ways:

- Traffic data and parameters enable detailed characterization or study of traffic loading effects on pavement structure (e.g., pavement analysis, research, and forensic studies).
- Summary traffic parameters support high-level analyses (empirical pavement design, pavement performance or pavement maintenance modeling, or pavement management applications) (*1*).
- Other parameters are used in specialized pavement analysis and design software such as the *Mechanistic–Empirical Pavement Design Guide* (MEPDG) (*1,2*).

Parameters that provide detailed characterization of traffic loading are used for mechanistic and mechanistic–empirical pavement response and performance modeling. Analysis and modeling of pavement response require information about:

1. Wheel and axle load magnitude,
2. Load position and configuration (i.e., axle configuration and position of wheels on the pavement) as indicated in Table 4,
3. Area of load application or tire footprint,
4. Load duration, and
5. Time history of load application (i.e., changes in load magnitude over time).

Pavement performance modeling requires traffic loading history (i.e., the number and magnitudes of loads reported for specified time increments used in the analysis) for the entire analysis period. This information is typically collected by the traffic data collection staff within a DOT. Traffic loading is one of many uses of traffic data, making it challenging, but achievable, for the data collection staff to meet all the traffic monitoring needs for pavement engineering applications.

Summary traffic parameters are used in empirical pavement response and performance analysis and modeling, in empirical pavement design procedures, and in high-level analyses supporting pavement management models and decision support tools. For these analyses, a single traffic summary statistic is desired; this statistic may be the equivalent single-axle load

(ESAL), average annual daily truck traffic (AADTT), cumulative truck volume, or total load. These summary statistics are also used to identify and group sites in categories that represent different levels of traffic.

Another set of traffic parameters are direct inputs to specialized pavement analysis or design software, such as the traditional AASHTO 1993 and the newer MEPDG pavement software, AASHTOWare Pavement ME Design Software (*1, 2*).

Traffic Parameters Used in Empirical Pavement Analysis, Design, and Pavement Management Applications

The current state of practice in pavement engineering relies on empirically derived relationships between traffic summary statistics and pavement performance. Many studies of pavement response and performance use empirical methods or statistical models to correlate pavement performance parameters (for example, road roughness) monitored over time with traffic and environmental loads, site conditions, material properties, and construction practices. These studies frequently use a single traffic summary parameter to describe traffic. These analyses may require a complete history of changes in the selected traffic summary parameter (computed annually for the duration of analysis period), a single cumulative value aggregated over the analysis period, or one representative traffic summary value. The most frequently used traffic summary parameters are AADTT and ESAL (*2*).

ESAL as a Traditional Summary Traffic Loading Statistic

ESAL has been used as a summary traffic loading statistic for pavement design and analysis applications since the 1960s (*3*). It is a concept developed from data collected at the American Association of State Highway Officials (AASHO) Road Test to establish a pavement damage relationship by comparing the effects of axles carrying different loads. In ESAL computations, load equivalency factors (LEFs) convert a mixed stream of traffic, consisting of different axle loads and axle configurations predicted over a design or analysis period, into an equivalent number of 18,000-lb single-axle load applications summed over that period. Thus, ESAL is a cumulative traffic loading summary statistic. There is general understanding and consensus in the pavement engineering community that ESALs or LEFs do not precisely describe the relationship between axle load and specific pavement distresses like rutting or cracking. However, ESAL continues to be a convenient statistic for sizing and quantifying traffic loading levels for empirical pavement analysis and design.

Generic ESAL

The generic ESAL (GESAL) is a parameter computed similarly to ESAL. It uses LEF values for flexible pavements with the structural number equal to 5 and the terminal serviceability index equal to 2.5 (*4*). Because LEF values are held constant, GESALs are independent of pavement type and thickness and of level and type of pavement distress. Therefore, any changes in GESAL values can be attributed directly to changes in traffic loads. This makes GESAL a more-desired summary traffic loading statistic for comparison of loads or effects of loads on pavement performance between different sites. GESAL is more sensitive to heavy loads on pavement performance than it is to average load or total load summary statistics. However, use of constant LEF parameters keeps GESAL from being applicable as a direct input to pavement design.

Traffic and Truck Volume Summary Parameters

For pavement analyses focused on characterizing traffic or truck volumes at a given location, the most widely used traffic volume parameters are AADT and AADTT. AADTT is more relevant for pavement analysis and management applications than is AADT because trucks contribute much more to pavement damage than do the lighter vehicles that make up most of AADT. Other traffic volume statistics used in pavement analyses are total annual truck volume, annual truck volume by vehicle class, cumulative volume of class 9 vehicles, and cumulative volume of heavy trucks (vehicles in FHWA classes 4 and 6–13).

STATE OF THE ART

MEPDG Traffic Parameters

The pavement engineering world is undergoing a paradigm shift from the empirical to mechanistic-empirical (ME) design methods. The goal is to eventually develop mechanistic methods for pavement design. For over 50 years, the empirical pavement design method included one traffic summary parameter (i.e., ESAL), In contrast, the newer ME method requires extensive use of traffic data. Many ME pavement performance analyses are performed using the MEPDG method and software products. Table 4 describes the traffic parameters required for analyses and design based on the MEPDG method.

Traffic Loading Defaults for MEPDG

Recognizing the emerging state of WIM technology and the need for research-quality WIM data to support Long-Term Pavement Performance (LTPP) research, the LTPP program installed and maintained WIM equipment at select Specified Pavement Study (SPS) test sections in 22 states. This effort proved that collection of consistent high-quality WIM data (satisfying ASTM E1318 WIM Type I performance requirements [*5*]) for more than 10 years is possible with proper maintenance and calibration (*6*). The data from the LTPP SPS WIM sites have been used to develop the new generation of traffic loading defaults for use with the MEPDG method (*7*). These defaults were included in AASHTOWare Pavement ME Design software for use nationally and internationally. FWHA published a guide for selection and use of these defaults (*8*). LTPP expanded its WIM data collection to include the program's warm-mix asphalt pavement experiment. Several states are also working on or have completed development of their own MEPDG traffic loading defaults. Table 4 describes the various MEPDG inputs and input definitions.

TABLE 4 Traffic Input Parameters Required by the AASHTOWare Pavement ME Design Software

MEPDG Input Parameter	Parameter Description
Axle load distribution factors (ALDF)	ALDF represents a percentile axle load distribution for a typical day for each calendar month for a typical design–analysis year. One set of ALDF is provided for each vehicle class (classes 4−13), axle group type (single, tandem, tridem, quad), and calendar month (January–December). ALDF remains constant between analysis years. One representative percentile distribution of vehicles in each of the classes 4–13 is provided to represent an average vehicle class distribution for the base design or analysis year.
Monthly adjustment factors	One representative set of 12 monthly coefficients is provided for each vehicle class 4–13 to represent differences in truck volume by calendar months for the base design or analysis year.
Hourly distribution factors	One representative set of 24-hourly factors showing the percentage of total truck traffic for each hour. Values are the same for all truck classes and only apply to truck volume. This input parameter only applies to Portland cement concrete (PCC) pavements.
Number of axles per truck	One representative set of values showing the average number of single, tandem, tridem, and quad axles for each truck class (classes 4–13).
Base (first) year AADTT for design lane	One value representing average annual daily volume of vehicles in classes 4–13 for the base design–analysis year. If this input parameter is used in MEPDG software in place of two-way AADTT, enter the following values also: percent trucks in design direction = 100% and percent trucks in design lane = 100%. Alternative input: MEPDG base (first) year two-way AADTT.
Base (first) year two way	Two-way AADTT computed for the base design or analysis year.
Percent of trucks in design direction (%)	Percent of trucks traveling in design direction (direction of LTPP lane) for the base design or analysis year.
Percent of trucks in design lane (%)	Percent of trucks traveling in design lane (LTPP lane) for the base design or analysis year.
Vehicle class annual volume growth rate by vehicle class (%)	Growth rate (%) for each truck class (classes 4–13). Applied together with the growth function (linear or compound) to estimate truck volume over analysis or design period from the base design or analysis year AADTT values.
Vehicle class growth function	Type of truck volume growth function (linear or compound) by vehicle class 4–13. Applied together with the growth rate to estimate truck volume over analysis–design period from the base design or analysis year AADTT values.
Operational speed (mph)	Value defined as posted speed limit or the average speed of the heavier trucks through the project limits.
Axle spacing for tandem, tridem, and quad axles (in.)	Average representative axle spacing for tandem, tridem, and quad axles.
Average wheelbase length and corresponding percentage of trucks	The average wheelbase length and the corresponding percentages of trucks with wheelbases that fall in the following three categories: short (≤1 2 ft), medium (> 12 and ≤ 15 ft), and long (> 15 and ≤ 20 ft). For multi-unit and combination trucks, only wheelbase of the truck power unit (i.e., first unit) is considered. Used for top-down jointed plain concrete pavement cracking model only.

MEPDG Input Parameter	Parameter Description
Average axle width	The distance in feet between two outside edges of an axle. Constant between all truck classes. Only needed for rigid pavement designs.
Mean wheel location	The mean distance in inches from the outer edge of the wheel to the pavement marking. This parameter is constant between all truck classes and does not change with time.
Truck wander standard deviation	Standard deviation from the mean wheel location, computed in inches based on measurements from the lane marking.
Dual tire spacing	This parameter is constant between all truck classes and does not change with time.

Traffic Parameters for Mechanistic–Empirical Pavement Performance Prediction

Research on ME pavement performance analysis and modeling focuses on analyzing and predicting pavement distress that develops over time. Most pavement distress (cracking, rutting, faulting) develops from incremental or cumulative changes in pavement structure due to material aging, environmental impacts, and traffic loading. Therefore, for traffic loading characterization, information about individual traffic loads must be known, but also the sequence and the cumulative total number of traffic load applications that lead to pavement deterioration over time.

Axle Loading Characterization

Traffic loads are summarized in the form of an axle load spectrum (or axle load distribution) to track and summarize traffic load applications over time (*7*). The axle load spectrum represents a frequency distribution of axle loads in which counts of axle load applications are summed and reported using predefined load bins. Recognizing the importance of load configuration, separate axle load spectra are used to summarize axle load counts for typical axle load groups: single, tandem, tridem, and quad. Depending on the intended application, load spectra can be created for an individual truck class or for all truck classes combined. The axle load spectrum input provides information about the axle load magnitudes, number of axle load applications over a specified period, and load configuration (i.e., the number of axles in each axle load group). Such detailed characterization of traffic loading allows modeling of pavement responses. This type of modeling uses methods where each axle load application on the pavement, whether expected or observed, is modeled and its effect on pavement response and performance is predicted. In addition to the axle load spectrum, information about the relative position of axle loads on the pavement is also important, especially for jointed rigid pavements.

Relative Pavement Performance Impact Factor

Traffic loading summary statistics include two new parameters: relative pavement performance impact factor (RPPIF) and annual total truck load (ATL). The RPPIF statistic is computed similarly to ESAL (*7*), but instead of the LEF factors based on the data from the AASHO Road Test, it uses W-factors determined through MEPDG analysis. The W-factors are based on the globally calibrated distress prediction models included in the NCHRP Project 01-37A MEPDG report and software (*9*). The analysis leading to W-factor development considered both jointed concrete and flexible asphalt concrete pavements located in four different climatic regions (wet–freeze, wet–no freeze, dry–freeze, and dry–no freeze). The goal was to develop a single-value summary loading statistic—RPPIF—that could help highway agencies determine the number of default axle load distributions necessary to support MEPDG pavement design implementation. Differences in RPPIF of 10% or more may lead to differences in pavement design thickness of the top structural layer of 0.5 in. or more; such differences would justify the need for multiple axle loading defaults. The primary purpose of the RPPIF statistic is to compare axle loading distributions between different sites. As with GESAL, RPPIF values are independent of pavement type, thickness, and distress level. This statistic could also be used to identify and group sites with similar traffic loading levels.

Annual Total Load

The ATL statistic is a simple estimate or summary of all traffic loads accumulated over a year. The main advantage of ATL is that it is independent of any empirically derived relationships that relate load to damage. However, it cannot help infer whether trucks are empty or loaded. It can also be unclear whether an ATL value results from the number of trucks or from the weight of trucks; a small number of heavy trucks and large number of light trucks may produce the same ATL value. This makes ATL less desirable for analyses of pavement responses or performance that have a nonlinear relationship with the load magnitude.

GAPS IN PRACTICE AND KNOWLEDGE

Traffic Parameters for Pavement Response Prediction Based on Mechanistic Models

As pavement engineering evolves from the empirical to mechanistic-empirical and then to fully mechanistic methods, the demand for more accurate and more detailed traffic loading characterization continues to rise. The emerging mechanistic pavement response analysis and modeling studies focus on stresses, strains, and deflections that pavements experience under each traffic load application. Pavement responses could be predicted using static or dynamic mechanistic modeling methods.

Pavement responses predicted by static models (elastic, viscoelastic, and elastoplastic) depend on the following traffic loading parameters:

- Load magnitude;
- Load configuration (i.e., location and number of wheel loads simultaneously applied on the pavement surface);
- Sequence of loads;
- Time and date of load application;
- Area of load application and shape of load distribution under each wheel (i.e., over the tire footprint); and
- Position of the wheels and axles relative to the pavement edge or concrete slab edges.

Pavement responses predicted by dynamic models consider the dynamic effect of the applied loads. In addition to the parameters listed above for static modeling, dynamic models require the following additional inputs:

- Load duration,
- Rate of load application (number of load applications per time unit measure), and
- Time history of load application (change in load magnitude or pressure under tire footprint over time, as each wheel passes over the specific pavement location).

Existing traffic monitoring technologies, especially WIM, can provide most but not all the above parameters. New advancements are required to take WIM measurement ability beyond the estimation of the static equivalent of axle or truck load weight. This ability will need to include accurate recording of the full-time history of load application, including accurate measurement of the dynamic forces applied by the tire to the pavement and quantification of the exact area of load application (load or tire footprint) and position of each tire footprint for each truck, relative to the pavement edge.

Accuracy of Weight Data

The emerging mechanistic and MEPDG methods currently being implemented demand an accurate measurement of traffic loads. To provide accurate prediction of stresses, strains, and deflections in pavement structure, weight measurements should be as accurate as those used for weight enforcement. For MEPDG methods, accuracy of WIM data should satisfy performance requirements of ASTM E1318 Type 1 WIM systems (*10*).

Big Data

Instrumented vehicles and advances in weigh-in-motion and sensor technology permit the acquisition of larger volumes of truck loading and classification data from new sources. Vendors are collecting various types of information, including probe data, although currently trucks remain a very small part of the sample set. Acquisition of these data for pavement purposes may be secondary to their use in system performance, freight modeling, safety analyses, and noise studies.

Data collected by non-traditional means require validation of accuracy, consideration of privacy issues, and methods to review and analyze the data quality. Integrating agency collection and independent validation of data provided by outside sources will be a common issue across multiple data types, beyond just pavement needs. More discussion of probe data sources for traffic volume is in Probe Data for Traffic Volume Estimation.

CURRENT AND PROPOSED NATIONAL RESEARCH INITIATIVES

This section describes existing and proposed research to addresses gaps in practice and knowledge for pavement engineering applications related to traffic monitoring.

Current Research

Study to Facilitate Use of LTPP Traffic Data in Pavement Applications

The FHWA–LTPP program completed research reviewing nationwide traffic volume, classification, and WIM data available through the LTPP database (*11*). These data were used to develop traffic parameters suitable for pavement research, design, and analysis applications, and to rank these parameters for use by pavement researchers and practitioners. LTPP traffic volume, vehicle classification, and WIM data evaluations consider data availability, accuracy, and applicability for different pavement applications. New database tables with analysis-ready traffic parameters, along with the pavement analysts' *Guide to LTPP Traffic Data,* are available. The guide helps users to quickly select the most appropriate LTPP traffic statistics for the pavement analyses they are performing. In addition, the LTPP program has developed traffic inputs for all its traffic sites in the format compatible with the AASHTOWare Pavement ME Design software. This facilitates LTPP traffic data use in MEPDG applications.

LTPP Study to Estimate Traffic Inputs for Network Level Pavement Applications

The LTPP program recently completed a study using LTPP monitored traffic and a variety of spatial information covering location and demographic characteristics; this data generated models that estimate some of the traffic inputs for pavement applications (*12*). With these models, a user with nothing more than traffic volume data and information derived from geospatial datasets can estimate truck volumes, truck growth rates for HPMS truck groups, truck distributions across FHWA TMG truck classes, and axle load distributions by axle group and vehicle classification. Additional information may include monitored vehicle classification versus traffic volume data, ESAL estimates versus no loading information, or limited axle monitoring data versus ESALs alone. All of this additional information gives a more representative estimation for the network location. The models are being used either for LTPP sections with no loading information or for those with no loading information beyond ESAL estimates (FHWA Contract DTFH6114C00023).

WIM and Freight Data Use for MEPDG

About half of state highway agencies are in various stages of implementing the MEPDG method for pavement design and conducting traffic data studies to ready their state for MEPDG implementation. Two representative studies are being conducted in Michigan and Pennsylvania.

Michigan DOT is performing research to develop a methodology for traffic loading characterization for all state roads based on available WIM, road inventory, and freight data. The results will be used to develop traffic loading defaults and the methodology of selecting these defaults for pavement design using the MEPDG method. The Pennsylvania DOT (PennDOT) has begun to implement the MEPDG into its routine pavement design practice by characterizing traffic and material inputs and by verifying/calibrating MEPDG transfer functions.

To fill the gaps in traffic data, PennDOT developed default values based on analysis of traffic data from its WIM and continuous automated vehicle classification traffic monitoring sites. However, for several truck traffic input parameters, supporting data were not available within their truck traffic database. For these parameters, PennDOT found that defaults based on the research-quality WIM data collected by the FHWA Transportation Pooled Fund TPF5(004) study were applicable for its designs. The traffic default values were included in the PennDOT Pavement ME Design data library. The PennDOT-specific traffic inputs will be used in the state's regional verification and calibration of transfer functions. The inputs are required for predicting distresses in both flexible and rigid pavements.

Proposed Research

Enabling Detailed Traffic Loading Data Collection

Research and development are needed for sensors capable of capturing detailed and accurate traffic loading history. They are also needed for information about location and size of the loading area (tire footprint and load distribution) to enable mechanistic pavement response and performance modeling. As new technologies capable of measuring truck wander, position of individual tires within travel lane, tire footprint, and tire inflation pressure come on the market, data collection and data reporting protocols should be developed. These protocols should include specifications for data reporting formats compatible with current FHWA individual vehicle reporting (IVR) formats. The results of these research studies would aid in developing a anew generation of mechanistic pavement analysis, design, and management methods and tools.

Improving Accuracy of Traffic Inputs

Accuracy challenges are not limited to existing technologies. The accuracy of estimation and validation of expansion techniques for vendor-provided data is an upcoming issue. Probe data and other samples have limited publicly available information for expansion and validation of the estimates. The current data options are in their infancy, but experience with similar products for the passenger fleet hint at potential concerns over proprietary data. Research is needed to evaluate the applicability and limitations of different traffic data sources. Research is also needed on techniques for estimating traffic inputs for pavement engineering and pavement management applications, including development of guidance and methodologies for pavement users.

Advanced Methods for Project-Level Traffic Loading Estimation

While site-specific axle loading information is ideal for pavement design, the high cost of WIM data collection makes it impractical to have a WIM site at every pavement design location. Therefore, new methods are required to:

1. Accurately estimate site-specific axle loading from the limited number of WIM sites maintained by state highway agencies,
2. Obtain information about freight carried by trucks on specific highway roads,

3. Gather data from connected vehicles (e.g., onboard truck sensors capable of transmitting truck or axle weight data), and
4. Acquire other readily available data.

In addition, research is needed to explore the feasibility of using inexpensive portable WIM data collection equipment and methods to help estimate traffic loads for pavement design in combination with other data sources. Georgia DOT has found portable WIM data to be a viable source of general traffic loading information; such data could help select default values for high-significance pavement designs that lack site-specific axle loading distribution data (*10*).

Freight and the Urban Environment

Existing truck weight and axle load measurement technologies, like WIM, require high-speed travel, not stop-and-go traffic conditions. Existing technologies are thus not applicable for roads susceptible to congestion or stop-and-go traffic. Alternative methods for estimating traffic loading inputs for these roads are necessary. The development of estimation methods for facilities other than rural highways is likely to be related to freight performance metrics and modeling. As urban distribution centers increase in number and as real-time vehicle tracking technologies mature, the ability to provide accurate traffic loading data for site-specific pavement evaluation and design will increase.

REFERENCES AND OTHER RESOURCES

1. *AASHTO Guide for Design of Pavement Structures*, American Association of State Highway and Transportation Officials, Atlanta, GA, 1993. https://www.fhwa.dot.gov/engineering/geotech/pubs/05037/ac.cfm;https://habib00ugm.files.wordpress.com/2010/05/aashto1993.pdf.
2. *Mechanistic–Empirical Pavement Design Guide, Interim Edition: A Manual of Practice*. American Association of State Highway and Transportation Officials, Washington, DC, 2008.
3. Yoder, E. J., and M. W. Witczak. *Principles of Pavement Design*. Wiley, New York, 1975. https://doi.org/10.1002/9780470172919.
4. Hajek, J. J., O. Selezneva, G. Mladenovic, and J. Jiang. *Estimating Cumulative Traffic Loads.* Volumes I and II, FHWA-RD-03-094. Federal Highway Administration, U.S. Department of Transportation, March 2005.
5. E1318-09, Standard Specification for Highway Weigh-In-Motion (WIM) Systems with User Requirements and Test Methods. American Society for Testing and Materials, West Conshohocken, PA, 2017.
6. Walker, D., and D. Cebon. The Metamorphosis of Long-Term Pavement Performance Traffic Data. *TR News*, No. 277, November–December 2011.
7. Selezneva, O. I., M. Ayres, M. Hallenbeck, A. Ramachandran, H. Shirazi, and H. Von Quintus. *MEPDG Traffic Loading Defaults Derived from LTPP Traffic Pooled-Fund Study.* Final Report, FHWA-HRT-13-090. Applied Research Associates, Inc., Elkridge, MD, April 2016.
8. Selezneva, O., and M. Hallenbeck. *Long-Term Pavement Performance Pavement Loading User Guide*. FHWA-HRT-13-089. Federal Highway Administration, U.S. Department of Transportation, October 2013.

9. Applied Research Associates, Inc. Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures. Final Report for NCHRP Project 01-37A, Transportation Research Board, Washington, DC, 2004.
10. Selezneva, O., and H. Von Quintus. Traffic Load Spectra for Implementing and Using the Mechanistic–Empirical Pavement Design Guide in Georgia. FHWA-GA-14-1009. Atlanta, GA, February 2014.
11. Selezneva, O.I. and M.E. Hallenbeck, *Facilitating Analysts' Use of Traffic Data from the Long Term Pavement Performance (LTPP) Program*. FHWA-HRT-22-074. Federal Highway Administration, U.S. Department of Transportation, June 2022. https://highways.dot.gov/sites/fhwa.dot.gov/files/2022-06/FHWA-HRT-22-074.pdf.
12. Ostrom, B., P. Serigos, and J. Bryce, *Predicting Truck and Axle Loading Patterns*. Federal Highway Administration, U.S. Department of Transportation, June 2022.

Traffic Monitoring Statistics, Data Quality, Usage, Integration, and Equipment Calibration

STEVEN JESSBERGER
Federal Highway Administration

DEBORAH WALKER
Federal Highway Administration

STATE OF THE PRACTICE

Cities and towns, metropolitan planning organizations (MPOs), and state departments of transportation (DOTs) collect highway traffic monitoring data (i.e., vehicle volume, classification, and weight) for sections of roadway that represent travel on their surface transportation network. Continuous and short-term counts of vehicles in the 13 FHWA classes are collected. Continuous counting involves collecting data continually either hourly, in more frequent time increments, or on an event or individual vehicle record (IVR) basis. Continuous counting is similar to short-term counting but extends over an interval of more than one week in any given location. It also often includes 365 days per year of binned in hourly (or 15 min, IVR) traffic counting data. Short-term traffic data are normally acquired over one to seven days and represent the spatial data sets for the agency's roadways.

Traffic data collection devices gather traffic information and are installed in either a portable or permanent manner. Both manners require reliable and timely calibration and the ability to provide accurate data in support of roadway network decision processes. Equipment calibration can be simple when performed for volume counting (completed in one hour); it can also be complex when calibrating weigh-in-motion (WIM) sites (completed in a day). Quality control and assurance methods are most often integrated throughout the process from field data collection to office processing and eventual production of the final data products. Accurate reporting of vehicle classification is important: class data are used in many ways from pavement design to studying the environmental impacts of highways. These data also provide an understanding of goods movement in the United States.

Field calibration of traffic equipment must be performed at least annually, as detailed in the FHWA *Traffic Monitoring Guide* (TMG) (*1*). The TMG provides guidance for state, local, and other transportation agencies involved in traffic data collection programs. For example, it is a recommended method for volume calibration to calibrate to the most detailed level the data is capable of supporting, for instance by lane. Ground truth recommendations include manual counts performed by human observers, video recordings with post-processing counts by human observers, or comparisons with a gold standard counter used for calibration of volume counting devices. Equipment used for field counts is typically accurate within ± 5% to 10%. Standard forms and methods should be used for recording the field counts and calibration results over multiple years for tracking trends in the calibration results. Example forms for calibration of WIM sites can be found in the FHWA's Weigh-in-Motion Pocket Guide.

Parameters calibrated at vehicle classification sites include speed, vehicle length (bumper-to-bumper), and inter-axle spacing. All classification checks are best performed on an individual lane basis. Speed calibration is best performed using a laser speed gun. Ensuring the site's speed data are accurate will also ensure the accuracy of axle spacing measurements made at the site. Accurate inter-axle spacings are needed for proper vehicle classification. Typically speed verification tolerances are within 1 to 2 mph. Validation is usually performed over a few-minute interval by measuring the speed of different vehicle classes, by length or axle spacing, in each of the lanes. If the speed of the vehicle is accurate (within 1 to 2 mph), the counting and axle spacings are generally also accurate (provided the site has a good classification method defined). As a result, vehicles are likely to be properly classified.

A second method for verifying the accuracy of classification sites involves either having a vehicle of known bumper-to-bumper length for length calibration or using the known axle spacing(s) of vehicles as they travel across the sensor array in each lane. A third method for verification of vehicle class involves calibrating by lane using either manual class counts or video data that is later manually counted to compare with the classification device under calibration. This third method is among the best for checking the accuracy and should check each vehicle against the visual vehicle classification. Performing class calibration can be done by time period for a given number of hours with stated accuracies for each vehicle class. It can also be done by any time period as long as a certain sample size of each classification is obtained to verify that each vehicle class meets the accuracy requirements.

For WIM sites, calibration often requires using vehicles of known axle weight, length, and inter-axle spacing between each axle. WIM calibration methods are specified in ASTM Standard E1318-09(2017) (*3*) and elsewhere (*4, 5*). The Long-Term Pavement Performance (LTPP) program's method for calibrating WIM sites recommends using two different types of vehicles: one smaller vehicle and one fully loaded five-axle semi-truck (*2*). WIM site calibration is critical because inter-axle spacings and weights from the sensors can vary over time and thus affect the accuracy of the measurements. There are various methods to monitor the calibration of WIM sites including:

1. Front axle weights of specific vehicle classes (e.g., five-axle semi-trucks [FHWA class 9]);
2. Tandem axle group weight distributions; and
3. Gross vehicle weight (GVW) over time for a given time period or specific day of the week.

For WIM sites with left and right weights obtained by wheel path, the in-balance can also be used to check system calibration of individual wheel path sensors.

Although each agency often collects both continuous and short-term count data for their specific needs, there are several similarities across the nation's traffic monitoring programs in regard to calibration and quality control. State DOTs normally document procedures, standards, and specifications. Every traffic counting program must provide its traffic data to its customers. The goal of every count is to represent what actually took place on the roadway for the reported period of time. Proper annual, or more frequent, calibration of equipment and sites ensures that reliable and accurate results are provided to customers and users of the traffic data.

STATE OF THE ART

State-of-the-art programs use automated systems to manage the data quality (i.e., reliability, precision, and accuracy). They provide near real-time information resulting from the effective deployment of resources and corrective actions. Automated systems provide count; initial QA of data; initial QC checks to ensure completeness, precision, and accuracy of the traffic counts; and advanced QC checks to ensure the data meet nationally acceptable ranges. Transparency concerning the methods employed and the results from the data calculations and reviews often occurs through comprehensive documentation of current practices.

The American Association of State Highway and Transportation Officials developed the *AASHTO Guidelines for Traffic Data Programs*, *2nd Edition* (*6*), the reference guide for counting. This document establishes recommended national traffic monitoring practices that reflect current practice. Advances in traffic monitoring technologies and data collection methods are described in FHWA's 2022 TMG. Many state DOTs and other agencies find that the Highway Performance Monitoring System (HPMS) (*7*) requires more detailed collection and reporting of traffic data and are therefore trending toward collecting IVR data at more locations. The detailed IVR data support improved methods to check data quality and to correct issues as they are found. IVR data support advanced QC methods when post processing the data. They also result in a rich data set that provides gap/headway analysis, speed by vehicle class, weight from WIM sites, and travel by vehicle type that is more detailed (when using vehicle signatures) than the 13 vehicle classes specified by FHWA.

Many agencies have documented methods for their QA, QC, and database structures. These methods include per-lane automated feedback along with emails of daily downloads, completeness, quality issues, and status of each day's data. *Seven Deadly Misconceptions About Information Quality* (*8*) assists agencies in understanding data issues, what parts of data affect data quality, and how to best account for detected issues. This information is helpful for repairing, troubleshooting, and identifying those sites that may warrant further scrutiny. The AASHTO traffic monitoring guidelines and FHWA TMG both recommend periodic review of procedures used by traffic monitoring programs in the areas of QA, QC, and calibration, along with daily reviews of the traffic data no matter the source.

There are also National Institute of Standards and Technology (NIST) methods and documents that describe calibration of data collection systems. New NIST standards on the use of WIM data are being investigated by FHWA. As previously mentioned, FHWA has done studies that contain data quality checks for both classification and weight data collection.

The TMG lists numerous issues that affect calibration of traffic counting sites, including QC methods employed in the FHWA Travel Monitoring Analysis System (TMAS). These issues are in the TMG's Appendix B and in other references that describe data quality issues that agencies experience. Solutions to these issues include applying best practices to check volume, classification, and weight data. Additional information about quality and calibration aspects of collecting WIM data is in the *Weigh-in-Motion Handbook* from Iowa State University's Center for Transportation Research and Education (*9*) and other sources found in the Travel Time, Speed, and Reliability Data section of this circular.

EMERGING TRENDS AND DRIVERS OF CHANGE

Agencies are trending toward collecting more travel monitoring data. Where they used to collect only classes 4–13 for WIM data, many agencies now collect data for all vehicle types at WIM sites. This expands the use of the data to topics such as axle correction factors. Over 15 state DOTs now collect all individual vehicle weight data. Data collected include speed data, nonmotorized (micromobility) data, IVR data, signalized intersection data, crowd sourced data, or real-time data.

In the past, data was collected by site. Data are now collected and stored by travel direction or by lane. In addition, data traditionally were summarized over a given time period, such as in 60, 15, or 5-min intervals. Agencies are moving toward collecting and storing traffic data in the 2022 TMG IVR format that retains the rich information from each lane's array of sensors. Many state DOTs and portable device vendors collect event and IVR data. In dealing with funding issues and reduced staff to complete quality data reviews, agencies are leaning toward automated systems to obtain, review, store, analyze, and report traffic data. This has led to several companies providing QC services to agencies. Additionally, some agencies now pay for high-quality complete data instead of having in-house staff perform such work. The 2022 TMG recommends that agencies shift to IVR data recording and storage for all motorized traffic counts because IVR adds fidelity to collected data, added QC methods, and the ability to provide more detailed information to traffic data users.

One reason for the transition to new data collection methods and sources is that the limitations from 10 years ago (cost of field data storage, transmission rates, and device CPU speeds) have either been overcome or the issues have been reduced acceptable levels. With increased data resolution through widespread availability of IVR data, automated processing has improved significantly. Whether purchased from a software company or gathered by agency staff, traffic data are reviewed and processed faster than ever before. Many agencies are being pressed to provide data online. Such public review and feedback have led to improved information in support of better decision-making.

Another driver of change is the shift from the AADT process developed by AASHTO in the 1980s, which used daily volumes to compute AADTs, to the improved FHWA AADT calculation method (*10*), which uses data in any time increment (1 min, 5 min, 15 min, or hourly) (1). The newer AADT method improves the accuracy and use of datasets from non-traditional sources (primarily Intelligent Transportation Systems or ITS). Even those datasets containing gaps that would preclude using the older AASHTO AADT can now be used for monthly average daily traffic and AADT values. This leads to improved AADT quality because the new AADT method is more accurate and does not have the negative overall bias of the AASHTO method. It also leads to larger and more integrated traffic datasets, additional comprehensive calculations, and an increase in detailed information for reporting purposes.

Collecting traffic data once (correctly) and using the data many times is a trend that numerous data collection agencies, including FHWA, have long supported. For example, WIM data have been used to create spectra for site-specific axle loads used in pavement management applications (*11*). State DOT agencies' continuous count WIM data are also provided to the FHWA through TMAS. The TMAS weight data are a key dataset for national research projects. In addition to weight data, TMAS has processed and stored micromobility data at the national level since 2017.

Most agencies map traffic counts and provide the ability to visualize the counts with other data on geographic information system (GIS) layers to make informed decisions. Geolocation data are important in that they are larger, can be analyzed for quality and integrated with other datasets in new ways, and allow for a greater use of the data when merged with other large datasets (i.e., safety, roadway management, or operations). The spatially represented data are in alignment with processes that collect data once correctly and use it many times. A FHWA pooled fund study examined advanced methods to visualize data and review the information spatially and temporally (*12*). These techniques allow improved analysis and use of the rich class and weight data that traffic monitoring sites provide.

Traffic data collection programs are also key in providing information to an agency's overall asset management system. Several publications, including *Data Systems and Asset Management* (*13*), describe data collection systems as part of the asset management function.

GAPS IN PRACTICE AND KNOWLEDGE

Traffic monitoring program staff continue to integrate multiple equipment and data sources into their programs. However, traffic monitoring program data quality requirements hinder some interagency departments from sharing traffic monitoring sites, installation, equipment, and data. This is a gap in both knowledge and practice; some DOTs have yet to overcome data integration challenges due to technology configuration, cultural differences, institutional coordination issues, and a lack of common data definitions that all hamper sharing opportunities. By integrating datasets, the quality and calibration of traffic site data can be independently verified as with FHWA data from HPMS, TMAS, LTPP, and other data users. These sets of volume data afford many opportunities to check the data using non-traditional sources to ensure that trends and reported values represent the vehicle mix actually traveling on the roadway network.

Issues with using ITS data for traffic monitoring result from an absence of complete datasets. This absence is due to a lack of technical knowledge about how to install and configure a site for both traffic monitoring and ITS operational purposes. With the development of FHWA's new AADT calculation method, many ITS locations can be processed and potentially used as continuous counting sites if data from each time increment for each day of the week are present for each month of the year.

Manual data analysis becomes more complex as additional data are collected. Automated methods to verify the quality of the data are therefore needed. Transitioning from manually reviewing count gaps and reviewing pages of hourly counts from permanent counters to now receiving reports and summaries of travel trends over various time periods has led to significant improvements for many agencies. This is beneficial for agencies that have modernized, but not all agencies have yet done so. Thus, there is cause for concern with inconsistent QA/QC methods employed by state, MPO, city, and local agencies.

Documentation of the methods used, key fields, database relationships, and availability of the data is lacking to fully use the rich traffic datasets available today. With proper documentation, large datasets with good metadata can be more fully used and more easily managed. Public agencies are working to overcome the hurdle of sharing their vast amounts of data that have different spatial data, attributes, and structures. Nonmotorized (micromobility) calibration, accuracy, and QA/QC methods need to be established to assist in establishing consistent data programs that support multimodal analysis.

Most data acquisition and reporting programs specify a tolerance that expresses the acceptable variability in the data. However, the tolerance is often not related to a confidence interval or level. Thus, if a technician installs a road tube to obtain a needed count and it reports 9,250 axle hits, the agency divides by two axels to report a traffic count of 4,625 and may even designate it as the count representing a typical day. It is inaccurate to calculate an AADT without proper factors for axles, hour of day, day of week, and month of year to ensure the AADT's accuracy. Establishing standard methods for obtaining counts (IVR is recommended) and the associated accuracy of the different methods would contribute to informed decision-making based on good quality data instead of data that may be questionable. The traffic counting industry should work toward establishing methods for obtaining the required accuracies, from detailed data such as IVR and associated confidence intervals for reported volume, speed, classification, weight, and micromobility counts.

The collection of data for and knowledge of e-devices also needs to be better understood. The rapid adoption of electric bikes, scooters, and other devices that use paths and roadways has led to safety concerns given speed differentials and the extended distances these new devices can cover. Some assumptions will need to be adjusted to properly record and report e-devices in order to help agencies provide necessary data on micromobility adoption and e-device travel patterns.

CURRENT AND PROPOSED NATIONAL RESEARCH INITIATIVES

This chapter describes existing and proposed research to addresses gaps in practice and knowledge for pavement engineering applications related to traffic monitoring.

Current Research

Several agencies are exploring data visualization and integrating data with internal and external partners. The FHWA Pooled Fund TPF-5(195) on Web-based Traffic Data Visualization and Analysis Tools (*11*) developed methods to display WIM data on maps to view classification and weight data and other data quality issues in new ways.

A Small Business Innovation Research initiative that explored inductive loop detector signatures for vehicle classification concluded in the fall of 2017 (Phase II). This work improved the quality of classification counts and provided speed and detailed class of up to 200 unique vehicles from single-loop data. By using more detailed classes and possibly re-identification of vehicles in the traffic stream, one may be able to gather additional information about the vehicle classes in the traffic stream and improve the quality of the data and calibration between traffic counting sites. FHWA pooled fund TPF-5(520) (*16*), Improving Traffic Detection Through New Innovative i-LST Technology Demonstration Pilot, began at the end of 2023 and will examine how to turn point-based permanent count data into link-based data for various roadway corridors across the United States.

In 2017, FHWA began reviewing all traffic terms in HPMS Field Manual, TMG (*1*), Model Inventory of Roadway Elements (MIRE), Highway Capacity Manual, and AASHTO Guidelines for Traffic Programs (*6*) to ensure that methods to calculate a given traffic item are complete and correct and to reconcile variances among documents. The Traffic Data Computation Pocket Guide (*14*) now provides cross references to these documents and examples of how to calculate these items.

FHWA also has initiated research projects to support MIRE for local AADT data collection and HPMS for local VMT calculations. Both are ongoing and are being led by offices within FHWA.

NCHRP Project 20-50(20), “LTPP Data Analysis: Develop Practical Tools and Procedures to Improve WIM Data Quality,” developed six practical tools and guidance for collecting high-quality WIM data. The project used LTPP road inventory, pavement, and WIM data as primary sources for analyzing factors affecting WIM data quality (*14*).

The FHWA National Bikeway Network data processing software was completed in 2021 and now contains over 32 agencies’ GIS bikeway networks. Any agency can submit its bikeway network to this online reporting tool (*15*). FHWA is currently conducting nonmotorized research on collection methods and consistent reporting of nonmotorized counts for the U.S. FHWA’s goal is to establish a national bicycle network compatible with the GIS HPMS roadway system.

Numerous state DOTs are enhancing their GIS networks to include bidirectional travel, more detailed data integration within the state, better quality controls, and more accurate data for the annual HPMS data submissions and items that support the FHWA Transportation Performance Management (TPM) 1, 2 and 3 initiatives.

Proposed Research

Determine Accuracy of Different Types of Traffic Counts for Diverse Applications

There is a need for a study to determine the accuracy of traffic counts (volume, class, speed, weight, and identification of nonmotorized travel) needed for different applications of traffic data. Outcomes would include accuracy and other attributes of research-grade data, accuracy of WIM sensors for highway design, and accuracy of classification data for environmental and freight management applications. For example, is accuracy within 10% good enough for vehicle classification data? Should the concept of a confidence interval be incorporated into ASTM Standard E1318? If so, how? How many sites does are needed in a state or metropolitan planning district for a sufficient number of vehicle classification data samples? Should the number of sites be based on the type and funding resources of the agency providing the data, its service area, and the variability of the roadways?

Innovative Approaches for Automated Site Calibration: Exploring the Use of Cell Phone and Roadside Readers to Verify Classification Site Accuracy

What new methods can be employed to help automate site calibration? Is there a way of using cell phone or roadside readers that collect transponder, GPS, and Bluetooth information to verify classification site accuracy? The FHWA new pooled fund TPF-5(520) (*16*) looks to provide one such method using loop signatures to do this on major corridors across the United States.

Develop Data Imputation Methods for Missing Traffic Count Data

Can data be identified to fill in missing counts? If so, how much data are needed and how should the data be tagged, if at all? Past FHWA guidance and truth-in-data guidelines discouraged this practice.

Data Collection and Funding Methods to Obtain AADT for Lower Functional Class Roadways

To support the MIRE requirement of having AADT values for all paved public roads for safety analysis by 2026, agencies need collection guidance and funding methods to obtain this data. Most roadways where AADT values are not known are on the lower functionally classed roadways (FHWA functional class 6, minor collector, and 7, local).

REFERENCES AND OTHER RESOURCES

1. *Traffic Monitoring Guide*, Federal Highway Administration, U.S. Department of Transportation, Washington, DC, October 2022. https://www.fhwa.dot.gov/policyinformation/tmguide/.
2. *LTPP Field Operations Guide for SPS WIM Sites, Version 1.0*, Federal Highway Administration, U.S. Department of Transportation, Washington, DC, 2012.
3. E1318-09(2017), Standard Specification for Highway Weigh-In-Motion (WIM) Systems with User Requirements and Test Methods, ASTM, West Conshohocken, PA, 2017.
4. Weigh-in-Motion of Road Vehicles, European Cooperation in Science and Technology (COST): Action 323, 1998, Paris, 2002.
5. NMi International WIM Standard: Specification and test procedures for weigh-in-motion systems, NMi Certin, Dordrecht, Netherlands, 2016. http://www.nmi.nl/nmi-wim-standard/.
6. *AASHTO Guidelines for Traffic Data Programs*, 2nd Edition, AASHTO Publications Order Department, Atlanta, GA, 2009. https://store.transportation.org/Item/PublicationDetail?ID=1393.
7. *Highway Performance Monitoring System Field Manual*, Federal Highway Administration, U.S. Department of Transportation, Washington, DC, December 2016. https://www.fhwa.dot.gov/policyinformation/hpms/fieldmanual/.
8. L.P. English, *Seven Deadly Misconceptions About Information Quality*, 1999. https://www.niss.org/sites/default/files/english.pdf.
9. *Weigh-in-Motion Handbook*, Sections 6 and 7: Calibration, Accuracy and Quality Assurance, Center for Transportation Research and Education, Iowa State University, Ames, IA, 1997. http://www.ctre.iastate.edu/research/wim_pdf/index.htm
10. Jessberger, S., R. Krile, J. Schroeder, F. Todt, and J. Feng. Improved Annual Average Daily Traffic Estimation Processes. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2593, Transportation Research Board, Washington, DC, 2016. https://journals.sagepub.com/doi/10.3141/2593-13
11. Hajek, J.J., Selezneva, O.I., Mladenovic, G., and Jiang, Y.J., *Estimating Cumulative Traffic Loads–Volume II: Traffic Data Assessment and Axle Load Projection for the Sites with Acceptable Axle Weight Data*, Final Report for Phase 2, FHWA-RD-03-094, Federal Highway Administration, U.S. Department of Transportation, March 2005.

12. *Web-based Traffic Data Visualization and Analysis Tools*, TPF-5(280), Federal Highway Administration, U.S. Department of Transportation, Washington, DC, November 2016. http://www.pooledfund.org/Details/Solicitation/1335.
13. Schofer, Joseph L. Moving the Goods: Performance Measures and the Value Proposition for Transportation Projects, *Transportation Research Record: Journal of the Transportation Research Board*, No. 2460, pp. 1−11, 2014.
14. Federal Highway Administration. (2018). Traffic Data Computation Method Pocket Guide (Publication No. FHWA-PL-18-027). August 2018. https://www.fhwa.dot.gov/policyinformation/pubs/pl18027_traffic_data_pocket_guide.pdf.
15. National Bikeway Network Data Portal. https://nmtdev.ornl.gov/.
16. Improving Traffic Detection Through New Innovative i-LST Technology Demonstration Pilot, TPF-5(520), Federal Highway Administration, U.S. Department of Transportation, Washington, DC, 2024. https://www.pooledfund.org/Details/Study/752 Accessed Mar. 27, 2024.

Integrating Traffic Monitoring with Connected Vehicle Data

WEIMIN HUANG
HERE

CHRISTOPHER VAUGHAN
North Carolina State University

NILOO PARVINASHTIANI
Iteris

ALAN CHACHICH
U.S. DOT Volpe Center

ALICAN KARAER
Iteris

This chapter discusses the technologies that enable communications between connected vehicles (CVs), the data generated by these communication activities, and the use of such data for traffic monitoring. The State of the Practice section focuses on existing technologies and their applications. It first introduces background information on CVs, then examines the technology development, including the latest status on mobile communications like cellular vehicle-to-everything (C-V2X) and dedicated short range communication (DSRC). It also discusses the data available from existing CV pilot programs and the applications of such data (basic safety messages, signal performance and timing [SPaT] data) in traffic monitoring. The State of the Art section introduces ongoing and new communications technologies that might replace DSRC for CVs in near future, along with the applications of the data they produce. The Emerging Trends and Drivers of Change section discusses technology-related business procedures and management issues and existing practice or knowledge gaps that need to be examined. The section Gaps in Practice and Knowledge describes current research projects and proposes a set of recommended research topics.

STATE OF THE PRACTICE

Overview of Connected Vehicles and Communication Technology Development

A CV, as defined by the United States Department of Transportation (U.S. DOT), is a vehicle that is equipped with vehicle-to-everything (V2X) communications technology that enables it to communicate with other vehicles, infrastructure, and devices. The CV concept uses data collected from V2X to determine what other travelers are doing and identify potential hazards.

Two of the most studied V2X categories are vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I). Some of the data or messages in past V2V/V2I studies include:

- Cooperative awareness messages (CAM) and basic safety message (BSM): Status information about traffic flow, vehicle position, driving speed, driving direction, and other vehicle status.
- SPaT: Signal phase and timing information for each traffic signal phase at an intersection.
- Emergency vehicle alert (EVA): Advanced warning to drivers of the approach of an emergency vehicle.

While the main purpose of these data and messages is for real-time traffic operation and safety, the data can also be used for system performance measures and monitoring.

In the past, V2X communication had relied on DSRC based on wireless local area network technology. DSRC radios use the IEEE 802.11-2012 wireless protocols that also specify WiFi communications. In 1999 the U.S. Federal Communications Commission (FCC) allocated 75 MHz in the spectrum of 5.850–5.925 GHz (also referenced as the 5.9 GHz band) for DSRC use only. In 2016, the National Highway Traffic Safety Administration published a Notice of Proposed Rulemaking that would have seen DSRC-based V2X become mandatory on all new cars and trucks in the United States. The first commercially available connected vehicles in the United States were introduced in March 2017.

However, wide adoption of DSRC did not materialize in vehicles or infrastructure in the following years. Instead, cellular-based LTE and 5G communication technologies, known as Cellular V2X (C-V2X), have emerged as an alternative to WiFi-based DSRC. Besides using the 5.9 GHz band for direct communication for V2V and V2I, C-V2X also operates in the mobile operator's cellular spectrum for vehicle communications to servers. On November 18, 2020, the FCC reassigned 45 MHz in the 5.9 GHz band to unlicensed uses such as WiFi. The remaining 30 MHz was kept for C-V2X. On August 12, 2022, a federal court ruled that FCC can implement this bandwidth reassignment.

While C-V2X differs from DSRC in communication protocols, C-V2X is supposed to keep all the data features from DSRC. Data and messages such as BSM, CAM, SPaT, and EVA would also have been implemented in C-V2X systems. Transportation agencies can still have access to these CV data despite the DSRC to C-V2X changes.

Past DSRC Deployment, Data Collection

U.S. DOT has a safety pilot model deployment (SPMD) that tested over 2,700 connected vehicles in Ann Arbor, Michigan from October 2012 to April 2013. It also has several other recent pilot studies that were conducted or are currently ongoing. These include southeast Michigan and several analysis, modeling, and simulation testbeds that are evaluating the active transportation and demand management (ATDM) applications shown in Table 5. Although not indicated in the table, an additional testbed in San Diego, CA is analyzing potential ATDM

TABLE 5 ATDM Testbed Applications

ATDM Strategy	Application	San Mateo	Phoenix	Dallas (ICM[1])	Pasadena	Chicago
Active traffic management	Dynamic shoulder lanes	—	—	√	√	√
	Dynamic lane use control	—	—	—	√	√
	Dynamic speed limits	—	—	—	√	√
	Queue warning	—	—	—	√	—
	Adaptive ramp metering	—	√	√	√	—
	Dynamic junction control	—	—	—	√	—
	Adaptive traffic signal control	—	√	√	√	√
Active demand management	Predictive traveler information	—	√	√	—	√
	Dynamic routing	—	√	√	√	√
Active parking management	Dynamic priced parking	—	—	√	—	—
Dynamic mobility	DMA[2] program evaluation	√	—	—	—	—

[1] Integrated corridor management.
[2] Dynamic Mobility Applications: Intelligent Network Flow Optimization (INFLO) application consisting of queue warning, speed harmonization, and cooperative adaptive cruise control, and the Multi-Modal Intelligent Traffic Signal Systems application.

applications that include queue warning, speed harmonization, intelligent signal control, dynamic lane use control, dynamic speed limits, dynamic merge control, predictive traveler information, managed lanes, and dynamic routing.

In most of the previous or current test programs, the primary data is conveyed by the broadcast of a BSM by each vehicle several times a second. This message includes the GPS coordinates of the transmitting vehicle. Other nearby vehicles use this information to avoid crashes. The most common V2I messages sent by infrastructure transmitters are SPaT, which broadcast information for a traffic signal, and MAP, which provide the geographic description of the intersection. Applications in the vehicles use the information to improve safety, fuel efficiency, and reduce emissions, among other things.

The BSM can contain as few as a dozen or over a hundred data elements. The most useful data elements for counting and traffic management are the GPS coordinates, heading and speed of vehicles contained in Part I of the BSM. An extended BSM can also include data describing vehicle dimensions, vehicle class, and trailers. All identifiers in messages from a DSRC radio are changed frequently to prevent tracking and protect the anonymity of drivers. However, these changes also impede the use of BSMs to measure travel time and origin-destination flows.

Therefore, in addition to the BSM, the SAE J2735 standard defines probe vehicle data (PVD) and probe data management (PDM) messages with the purpose of sending snapshots of data for use by infrastructure applications. Besides weather-related status flags, vehicle position, and vehicle class, these data include an optional vehicle identifier for selected types of vehicles.

The BSM is designed for safety applications and hence transmits over a short range with low latency and small size. Short range requires many infrastructure receivers to use them for traffic data collection and management. Because the PVD message was designed for mobility rather than safety applications, it collects the snapshots and transmits them only when in proximity of a V2I receiver, making it a better choice for traffic data collection in a wider range of environments.

The U.S. DOT has made data from some connected vehicle pilot programs available to researchers via the Research Data Exchange, or RDE. Data currently available include those collected in the SPMD. This information was uploaded in December 2016 and includes both the entire dataset captured and a sample dataset collected on April 11, 2013. Other SPMD data from April 5–7, 2013 were scheduled to be uploaded in December 2016, but have not yet been included in the RDE. This April 5-7 dataset is called the Enhanced Operational Data Environment (E-ODE) and includes vehicle responses to emulated road weather warnings, incident zone warnings and other unique circumstances.

Other CV data available on the RDE include simulation results from Phoenix and Dallas testbeds, among other locations. BSM data are also included in the RDE from a number of studies, including the SPMD. Likewise, data captured by roadside equipment during pilot studies are available on the RDE.

One of the biggest benefits of the RDE is its influence on data formats and language. Each dataset imported into the RDE must meet specific standards to avoid compatibility issues among datasets. This results in the standardization of language across datasets originating from entities and agencies across the country, which will ensure interoperability and communication in the future between CV infrastructures as CVs begin to infiltrate the market at a higher rate.

The U.S. DOT migrated the existing data in the RDE to a new open portal called "ITS DataHub," which contains all the historical and newly available data from CV pilot programs and other U.S. DOT ITS programs. ITS DataHub is described later in this chapter under Current and Proposed National Research and Initiatives."

STATE OF THE ART

Current Testing and Research on C-V2X

The FCC issued a public notice on August 6, 2021, about the issuance of waivers requested by ITS licensees allowing them to test and operate C-V2X technology in the upper 30 MHz portion of the 5.9GHz band. These waivers will ensure coordination between licensees or independent licensee operation of technology in this band in approved geographic areas. The waiver process is intended to facilitate testing of C-V2X technology in place of DSRC until the FCC, Wireless Telecommunications Bureau, and the Public Safety and Homeland Security Bureau provide a ruling regarding appropriate bands for C-V2X communications.

Audi of America, in conjunction with other vendors, has already begun testing C-V2X technology in suburbs of Atlanta, GA as a part of the Audi C-V2X Pilot Project. One part of the pilot project specifically focuses on school zones and with school buses. This technology is in the form of roadside units that can inform drivers when near school zones or school buses to protect people outside of vehicles. The focus of this portion of the pilot project is on the use of LTE and 5G networks around the Atlanta suburbs. Likewise, another portion of the pilot project is using enhanced safety vests to alert construction workers and nearby vehicles of each other's presence. Audi is also experimenting with a traffic light information service that can provide drivers with information like a countdown to a red light while in the amber phase. This can help drivers make better decisions about slowing for an impending red light. Audi also promises that unique identifiers will be randomly assigned to vehicles in cities where this technology is active. This identifier is independent of other identifying information, like license plates, allowing for anonymity.

One of the bigger hurdles currently for C-V2X testing on LTE and 5G networks is latency. These types of networks have not proven to be able to provide adequate response signals for improved vehicle safety. As such, advances in communication technology, like faster networks, or other avenues for implementing C-V2X need to be explored. However, C-V2X also uses a dedicated 5.9 GHz band that is independent of a cellular network, particularly in V2V (vehicle-to-vehicle) and V2P (vehicle-to-pedestrian) applications. While this can enhance the communication among vehicles and pedestrians, the range is not as long as what can be available when using the full cellular network. The full network provides a larger communications range, even in comparison to DSRC, as long as the cellular network is operating.

Future C-V2X Development and Their Data Applications in Traffic Monitoring

The message data produced by DSRC tests and initial C-V2X tests focus on each individual vehicle's safety status, speed, and location. While these types of data can be used for traffic monitoring, future C-V2X development could greatly broaden their applications. Based on the roadmap published by 5G Automotive Association, more advanced use cases have been proposed for future 5G C-V2X:

- **Vehicle cooperation**: Enabled by the support for unicast between two vehicles, and multicast among a group of vehicles (as opposed to broadcast communications only under DSRC and current V2X). Two vehicles or a group of vehicles would be able to coordinate their behaviors for a safer and more efficient driving environment.
- **Pedestrian interaction**: Newly established V2P realm would share the intentions from vulnerable road users such as pedestrians to vehicles, and vice versa, leading to safer roads for pedestrians.
- **Data collection and sharing**: Improved latency, throughput, and reliability in future C-V2X would enable a much richer set of data to be collected and shared among road users. Vehicles will be able to report on a wide range of road objects and send the information to servers for making and improving HD maps. Vehicles can also process the road object information and share relevant results with other vehicles in real time.

These types of new use cases would generate various data both in real time and archives to help practitioners operating and maintaining roadway systems.

EMERGING TRENDS AND DRIVERS OF CHANGE

Privacy, Equity, Data Ownership and Management, and Funding

Privacy continues to be one of the biggest public concerns about CAVs. Drivers are right to have a reasonable expectation of privacy. Specifically, people do not want their destinations to be discoverable to private or public entities. Therefore, there has been a strong push to anonymize captured data. Anonymization has been successfully accomplished in the industry, but it creates a problem collecting travel data on these vehicles, particularly with travel times. As a result, many have worked to find a solution to this issue. These solutions come in many forms, either through the promise of maintaining privacy to the public or by allowing individual drivers to compromise by ensuring faster travel times as a reward for sharing their information.

The growing cyber threat is another critical trend highlighted by attention to hacking vehicles and traffic signals at DEFCON and Black Hat (the largest hacker conferences) in 2016 and 2017. In addition to the cybersecurity issues faced by traditional ITS, the distributed communications network and equipment ownership of CAVs opens additional attack vectors. Potential attacks include sending incorrect data, impersonating another vehicle, or even creating false data to impact safety and operational CV applications. Validation and network security issues in CV applications closely mirror issues described in other Internet of Things applications; CV implementations could benefit greatly from incorporating lessons learned in these systems. In a possible distributed denial of service attack, many falsified messages are sent to flood the server to prevent real messages from being received. Vehicles can be similarly overwhelmed by too many false broadcast messages.

Regarding equity, a fear among some is that CVs and AVs will only be available or beneficial to the wealthy driving public. Retrofitting older vehicles is likely out of the question, due to the cost of adding the required hardware and software capabilities, and not everyone can or will use these vehicles. FHWA is aware of this issue and is attempting to address it before CVs are mass developed. The Accessible Transportation Technologies Research Initiative is researching ways in which CV and AV can assist disabled individuals or those whose travel is typically nonmotorized by providing information about vehicles near pedestrians or bicyclists. This could greatly reduce the stress sometimes involved in nonmotorized travel.

Data ownership and management has been a persistent issue among transportation agencies even before CVs. Many agencies (local, county-level, or statewide) have been unwilling to share and coordinate data with neighboring agencies. Additionally, these agencies struggle to modify their data management systems and policies on freedom of information. Many are unsure how to use the potentially enormous amount of data in a timely manner while dealing with limited resources for program expansion. Concerns also continue over data ownership, an unwillingness to reformat existing data, or simply a fear of being judged by peers regarding current practices. These hurdles must be crossed if CV technologies are to succeed; this will likely start by educating transportation and safety agencies on the benefits of data sharing (safety improvements, congestion mitigation, cost savings). Fortunately, researchers have already begun informing agencies of the value of interagency data coordination and sharing, mainly through presentation of benefits of such coordination directly to several state and local agencies.

Another driver of change is the need for a self-sustaining business model or funding source to pay for the infrastructure portion of the CV system. Related questions are: Will this involve infrastructure funding for other types of data acquisition, like traffic counting? Can counting programs use CV data as planned, or do they need to ask for something specific to be included in the CV data transmissions?

GAPS IN PRACTICE AND KNOWLEDGE

Some of the critical gaps for CV practices and data include the following:

- Regulatory and technological uncertainty
- Limited V2X comprehensive testing (interferences)
- Telecom technology and data science developing at a faster pace than the transportation industry
- Standardization on data procurement

The primary concern currently for CV data is technological uncertainty. However, the FCC is allowing requests of waivers to test C-V2X using specific frequency bands, which will at least allow the research to continue. As this research is ongoing, another concern will be latency, as the LTE and 5G networks do not seem to provide adequate speeds for data transfer. Unfortunately, these are the state of the practice for network speeds, so more development of faster networks will be necessary.

Audi of America has begun testing of C-V2X but other C-V2X testing has been limited. While the FCC waiver does account for prevention of interference between multiple C-V2X tests, there has not yet been a circumstance where multiple C-V2X deployments have occurred. This is likely to happen in the future.

Likewise, while V2V and V2P applications may not rely on public cellular networks, the range of these technologies are limited. This may or may not be a significant issue, but with a shorter range in communication abilities, latency becomes a much bigger factor, particularly on higher speed corridors. Also, other V2X applications may operate on publicly available cellular networks, which can provide more-than-adequate range for notifications to vehicles, drivers, and pedestrians. However, three potentially major hardware issues exist with using public cellular networks, outside of the obvious hacker and malware concerns. First is network failure; second is increased latency due to bandwidth issues with more devices being added to the network; and third is technology obsolescence with advances in cellular networks. As most people have likely experienced, cellular networks can drop due to tower failure, satellite failure, or simple signal issues. Another common experience for many people is a bandwidth issue as networks become overloaded, such as at large sporting events or concerts. As such, the addition of thousands or millions of new devices (roadside and onboard vehicles) may cause similar issues. However, the data packet sizes of these devices should be much smaller than those for cell phones and computers. Lastly, while enhancements in technology are generally welcomed, many have also experienced frustration with technology upgrades, like updating from 3G to 4G LTE. The update can be somewhat seamless initially, but as older technology like 3G is abandoned, the roadside units and vehicles devices must be capable of adapting to technological advancements.

CURRENT AND PROPOSED NATIONAL RESEARCH INITIATIVES

This chapter describes existing and proposed research to addresses gaps in practice and knowledge for pavement engineering applications related to traffic monitoring.

Current Research

Relevant projects and resources that are of national significance, recently completed or currently underway are listed below.

- **ITS JPO Open Data Portal.** ITS DataHub provides a single point of entry to discover publicly available, reusable, and open ITS research projects and data management tools. ITS DataHub enables free access to research datasets and associated documentation in near real time and decreases the time from research to insight. Dataset areas that can be browsed include CV Message, Application Message, Trajectories, Field Test, Sensor Data, Research Results, Connected Equipment, and Weather. The portal can be accessed at https://www.its.dot.gov/data/
- **The CV Pilot Deployment Program** is a national effort to deploy, test, and operationalize cutting-edge mobile and roadside technologies and to enable multiple CV applications. Three pilot locations were New York City, Wyoming, and Tampa, Florida. The New York City pilot is focused on the safety of travelers and pedestrians in the city through the deployment of V2V and V2I CV technologies. The report from the New York City study indicated a large portion of the data had to be omitted due to various data storage issues and other “data issues.” However, more positively, the viable data points did indicate an increase in driver safety, including increased speed limit compliance, reduced red light running, and reduced unsafe lane changes, among other safety improvements. The Wyoming pilot site focused on the needs of commercial vehicle operators along I-80 in Wyoming. It sought to develop applications that use V2I and V2V connectivity to support advisories including roadside alerts, parking and inclement weather notifications, and dynamic travel guidance. Analysis was limited in the initial report on the Wyoming results. For instance, while the majority (65%) of alerts were weather related, they were not specific to the study corridor of I-80 but were instead statewide. As such, analyses were not conclusive on travel speed reductions as a direct result of these alerts. More information can be found at https://www.its.dot.gov/pilots/overview.htm
- **NCHRP Project 20-24(098), “Connected/Automated Vehicle Research Roadmap for AASHTO.”** This research cataloged open issues that need to be resolved to enable successful deployment of CAVs. The issues that will affect agencies and the public were organized into four areas: institutional, legal, policy, and operational. The researchers then narrowed the catalog to critical issues suitable for near-term research. Last, the issues were consolidated into research projects needed to address the highest priority

issues and described in a roadmap that estimates the time and resources required for each. The projects are grouped into four general subject clusters: institutional and policy, infrastructure design and operations, transportation planning, and modal applications. (Draft Connected/Automated Vehicle Research Roadmap for AASHTO https://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP20-24(98)_RoadmapTopics_Final.pdf).

- **NCHRP Project 08-119, "Data Integration, Sharing, and Management for Transportation Planning and Traffic Operations,"** is focused on developing tools, methods, and guidance for improving data integration, sharing, and management practices. These would enable transportation agencies, in collaboration with private-sector and public-sector stakeholders, to make better planning and operations decisions. CAV-related resources of this publication can be found at https://data.transportationops.org/connected-autonomous-vehicles-cav-data.
- **The CV Pooled Fund Study** is a research and development program to support state and local transportation infrastructure owner-operators in preparing to effectively deploy and operate CV systems infrastructure and applications. The program is in its fifth phase (March 2021-July 2023). The projects underway include guidance documents for MAP (SAE J2735 Map Message) Messaging, Connected Intersection Program, Connected Intersections Message Monitoring, and Model Connected Vehicle Data Architecture.
- T**he ITS JPO V2X Communications for Deployment** webpage includes resources on various topics including V2X testing, materials from the V2X Communications Summit, and near-term actions and future deployment visions for V2X. This resource can be found at https://www.its.dot.gov/research_areas/emerging_tech/htm/Next_landing.htm.

Proposed Research

Research on the following topics is most likely to have the greatest impact on the potential for counting and traffic management programs to exploit CV data.

- **Integration of CV data into traffic monitoring programs.** What interface should counting programs and traffic operations use for CV data? How can local and state traffic authorities be cleared to connect?
- **Develop guidance and requirements for advanced messaging for CVs.** Will the planned messaging for CVs meet the needs of traffic operations and counting programs? If not, what requirements are needed for advanced messaging research?
- **Maximizing benefits and value of CV data.** How can traffic authorities maximize the benefits of access to CV data? What potential does access unlock? What skill set changes are required?
- **Privacy issues related to CV data.** How can unintended privacy issues be avoided in applications that use and store CV data?
- **Impact of big data analytics on extracting value from CV data and risks of compromising privacy**. What is the impact of big data development analytics on extracting value from connected data and the risk of compromising privacy? This research would study the positives of big data analytics versus the privacy risks while also considering the impact of mitigation strategies.

Weigh-in-Motion

OLGA SELEZNEVA
Applied Research Associates

HANS VAN LOO
Corner Stone International

STEVEN JESSBERGER
Federal Highway Administration

STATE OF THE PRACTICE

Description of Weigh-in-Motion

Weigh-in-motion (WIM) is a traffic monitoring technology capable of capturing, recording, and communicating vehicle class, weight, speed, and axle configuration information as vehicles drive roadways at operational speeds. Based on the WIM sensor outputs, the WIM system provides an estimate of static wheel or axle loads and also gross vehicle weight (GVW) of a moving vehicle. It provides these estimates by measuring and analyzing the dynamic forces transferred by vehicle tires on the sensors (*1*). An in-road WIM system (installed in the roadway or bridge) is also capable of determining other parameters related to the vehicle and its passes over the WIM system and can store this information in an individual vehicle record (IVR). An IVR includes a unique record number, location of the in-road WIM system including the lane and direction of travel, date and time (time stamp to a 1/100th second) of passage, vehicle speed, vehicle length, inter-axle spacings (distances), wheelbase or vehicle length (bumper-to-bumper), vehicle classification, pavement temperature, and weight/load (wheel path, axle, axle group load, and gross vehicle weight) information. Depending on the application, a WIM system may be combined with other sensors or devices (e.g., digital cameras to record images of the vehicle and its license plate[s]). In such cases, the information included in the IVR may be extended.

WIM systems are designed to collect axle load and GVW data for the vehicles moving both at highway speed and at low speed. For the traffic monitoring community, the high-speed (HS) free flow WIM systems designed for capturing vehicle data at highway speeds are of most interest. WIM systems that are capable of reporting the weight of vehicles traveling at prevailing highway speeds make the weighing process more efficient and less disruptive than pull-out to permanent or portable static weigh stations that require the vehicle to stop. Thus, HS WIM technology improves road safety and the efficiency of highway freight transportation. The sensor output is converted to an estimate of the static weight of the vehicle.

Several different in-road WIM sensor technologies exist, including strain gauge sensors, hydraulic pressure gages, and others that measure various piezoelectric properties of different materials. Under the bridge sensors measure the structural response (e.g., bending) of bridge structural members caused by the passing of the vehicle(s). In-road sensors are the most widely used technology in the United States. Sensor selection is typically driven by data accuracy requirements, reliability, pavement type, traffic and environmental conditions, installation and operation costs, intended length of data collection, and temperature dependency.

WIM Data Applications

WIM technology addresses the need to characterize trucks and loadings both accurately and reliably for a wide range of applications. Accurate weights of vehicles from 2,000 lbs to over 150,000 lbs are needed from WIM systems; the required weight ranges may depend on the specific application. WIM data are used by practitioners for transportation planning; pavement and bridge design; pavement and bridge management; load rating; routing and permits for overloaded vehicles; freight planning and analyses; and safety, environmental, legislative and regulatory studies. Additionally, WIM data support network analyses to improve operational efficiencies and enable data-driven transportation asset management decisions and more efficient use of funding.

Motor vehicle enforcement authorities use truck weight and axle load data to plan enforcement activities and to identify specific vehicles that violate federal and state size and weight laws during real-time on-site monitoring. In some countries, WIM technology identifies transportation companies that have overloaded trucks and allows for direct enforcement and fines for truck weight and size violations without the need for a secondary static or low speed weighing.

For traffic monitoring statistics, WIM systems often produce more accurate and detailed vehicle classification data than traditional vehicle classification technology. They do so by providing axle weight data in addition to data about axle quantities and spacing.

In addition to the traditional WIM data elements, emerging “Super WIM” sites are being pilot tested. These sites include the combination of various sensing technologies that can capture and store additional vehicular and traffic flow elements listed in Table 6 and Table 7. These additional data expand the traditional applications of WIM data and increase the WIM data user base.

TABLE 6 Vehicular Data Collected by WIM Sites

Vehicular Data	WIM System Type or Application
Vehicle type (class of vehicles)	All
Number of axles and axle-to-axle spacing	All
Vehicle weight	All
Load of axle group, individual axle, and wheel	All
Vehicle length and/or wheelbase	All
Vehicle image (top and/or lateral)	Super WIM sites only
Vehicle DOT registration	Super WIM sites only
Vehicle license plate number	Super WIM sites only
Vehicle 3D-profile	Super WIM sites only
Dangerous goods identification plate	Super WIM sites only

Vehicular Data	WIM System Type or Application
Tire footprint, inflation, flat tire	Stress-in-Motion (SIM) sites only
Overloading or unsafe loading condition	All, typically for enforcement applications
Vehicle signatures linked to detailed vehicle type classification	WIM systems with inductive loop signature card installed

TABLE 7 Traffic Flow Data Collected by WIM Sites

Traffic Flow Data and Summary Statistics	WIM System Type or Application
Vehicle presence in lane and direction of travel	All
Date and time of passage	All
Speed	All
Headway	Through post-processing of data
Gap	Through post-processing of data
Vehicle lateral position in lane (lane wander)	Detailed position with SIM sites only
Traffic volume by lane (by day, hour, minute, second)	Through post-processing of data
Traffic composition by vehicle class (or type), by lane, and direction (by day, hour, minute, second)	Through post-processing of data
Traffic loading by vehicle class or type by lane, direction (by day, hour, minute, second)	Through post-processing of data
Overloading statistics	Through post-processing of data
Assessment of the structural safety of the bridges	Bridge WIM systems only

WIM System Components

The major components of a WIM system include:

1. Sensors embedded in the roadway surface, placed on the surface, or placed on/under bridge decks to detect and measure the vehicle characteristics.
2. Electronics to control the WIM system collection, processing, and storage of sensor measurement signals.
3. WIM infrastructure, including conduits, cabinet, and junction boxes.
4. Support devices, such as AC or solar power equipment to power the WIM electronics and communication devices to transmit the collected data to a remote server.
5. Firmware installed in the WIM electronics to process sensor measurements and analyze, format, and temporarily store collected data.
6. For most WIM sensor types, the pavement or bridge structure acts as a part of the WIM system and has a direct influence on WIM sensor output. For in-road WIM sensors, pavement segment (e.g., 250 feet) prior to and (e.g., 50 feet) after the WIM system should meet specific thickness, strength, and smoothness requirements (FHWA LTPP Guide 2009). For bridge WIM systems, the bridges or culverts have to meet specific additional criteria regarding type, length, and skew.

WIM Sensor Types by Technology

Several types of WIM sensors are available on the market, including the following:

- **Load cell scales**: In a WIM scale, one or more load cells are mounted between a rigid steel plate (on top) and the support frame (at the bottom). The load cell measures the vertical forces transferred from the wheel/axle to the plate and then to the frame.
- **Hydraulic plates** (hydraulic load cells): two steel plates with elastically deformable tubular spring elements are evenly placed between the plates. The elements are filled with a hydraulic liquid and connected to a gauge measuring the changes in volume resulting from a deformation of the elements. A force applied to the plate will cause a deformation in the tubular spring elements that is linear to the weight on the plate.
- **Bending plates**: the plate is simply supported by the frame at its edges and is instrumented with stain gauges, which measure the bend of the plate while a wheel or axle is crossing it. The bending strains are proportional to the applied vertical force. Combining several strain gauges, the wheel(s) weight(s) are estimated.
- **Piezoelectric cables**: consist of a flat coaxial cable with a thick brass outer sheet, with a piezo-electric film spiral-wrapped around a silver-plated copper wire. A force applied to the cable strip results in a signal between the core and sheath of the sensor. Piezoelectric cables include piezo-polymer and piezo-ceramic sensors.
- **Piezo-quartz strip sensors** consist of a row of quartz discs mounted in an aluminum profile. When a load is submitted to the sensor, a charge is generated that is proportional to the applied load.
- **Strain-gauge strip sensors** consist of several strain gauges that measure the deformation (strains) of the bars as a result of the vertical load applied on the sensor. The resistance of a strain gauge will change when it is deformed in a certain direction.
- **Fiber optic cables** or strips measure the force acting on the cable by measuring the changes in the characteristics of the light beam. This may be based on different principles: phase shift, change in the polarization, changes in the spectral characteristics, or changes in the optical path intensity (amplitude).
- **Bridge WIM systems** use an existing bridge as a weighing scale to estimate the axle loads and weight of the passing vehicles. This is done by measuring strains or displacements in a bridge superstructure (girders or slab) that are induced by vehicles crossing overhead. The most common sensors employed are strain gauges and strain transducers. During the WIM installation, the characteristics of the bridge (dimensions, structural response, construction, and material) need to be implemented in the system. Suitable types of bridges include short culverts, common beam-and-slab and slab bridges, and long-span orthotropic deck bridges.

WIM Standards and Performance Requirements

Several WIM standards are available. In the United States, the ASTM Standard Specification for Highway WIM Systems with User Requirements and Test Methods E 1318-09 is the primary accepted WIM standard (ASTM E1318-09(2017)). This document categorizes system types by their performance characteristics and intended applications. System accuracy is defined by the error tolerances, with errors computed as percentile differences from the static weight. The 95% compliance rate defined in ASTM E1318-09(2017) provides the minimum percentage of measurements that should be within the specified tolerances to satisfy the performance requirements.

The FHWA LTPP specification for TPF 5(004) WIM sites may be of interest to highway agencies collecting WIM data for pavement engineering applications. LTPP specification uses a statistically computed 95% confidence interval of error to verify if the tolerances listed in Table 8 have been satisfied. The mean error (measurement bias) is also to set a target value of less than 2%. The LTPP performance requirements include provisions regarding performance testing at three temperature and speed ranges.

Another U.S. specification of interest is National Institute of Standards and Technology (NIST) Handbook 44 (NIST 2017). Other WIM standards recognized around the world are the European Cooperation in Science and Technology (COST) 323 Standard (COST 2002), Netherlands Measurement Institute (NMi) WIM Standard (NMi, 2016), and the International Organization of Legal Metrology (OIML) R134 International Recommendation (OIML 2006).

Best Practices

Examples of best practices for WIM data collection include:

- Judicious site selection that follows ASTM E1318-09(2017) requirements or smoothness requirements specified in the AASHTO M331-13 Standard Specification, Smoothness of Pavement in WIM Systems, and/or Optimum WIM Locator Software (ASTM E1318-09(2017), AASHTO M331-13).
- Installation of in-road permanent sensors in smooth and structurally sound pavements.
- Use of WIM sensors that are not sensitive or that successfully mitigate changes in the environment, especially changes in temperature or seasonal changes in pavement support.
- Continuous collection and storage of IVRs that enable monitoring of data quality, fast identification of data quality issues, and evaluation of seasonal variations in truck weights.

TABLE 8 Functional Performance Requirements for WIM Systems (ASTM E1318-09(2017))

Function	Tolerance for 95 % Compliance[A]				
	Type I	Type II	Type III	Type IV	
				Value ≥lb (kg)[B]	±lb (kg)
Wheel Load	±25 %		±20 %	5000 (2300)	300 (100)
Axle Load	±20 %	±30 %	±15 %	12 000 (5400)	500 (200)
Axle-Group Load	±15 %	±20 %	±10 %	25 000 (11 300)	1200 (500)
Gross-Vehicle Weight	±10 %	±15 %	±6 %	60 000 (27 200)	2500 (1100)
Speed			±1 mph (2 km/h)		
Axle-Spacing and Wheelbase			±0.5 ft (0.15 m)		

[A] 95 % of the respective data items produced by the WIM system must be within the tolerance.
[B] Lower values are not usually a concern in enforcement.

- Routine data quality checks in the office (daily, weekly, bi-weekly, and/or monthly), typically using weights of Class 9 vehicles (GVW, single axle weights, tandem axle weights, and tandem axle spacing).
- Initial and annual field validation of WIM system functional performance parameters and calibration using heavy trucks of known weight to establish reference values (annually, as recommended by FHWA in the 2022 TMG and the FHWA WIM Pocket Guide). Additionally, there are procedures documented in *NCHRP Synthesis 386* (*9*) and *LTPP Field Operations Guide for SPS WIM Sites* (FHWA TMG 2022, FHWA LTPP Guide 2009).
- Real-time field observations and validations, at the time of calibration or independently, to visually check reasonableness of the system output for capturing the range of vehicle types in the traffic stream (vehicle classification checks). Included here is how agencies use weight station data as feedback to HS WIM for live calibration adjustment as vehicles travel the roadways.
- Routine preventive maintenance per WIM sensor manufacturers' recommended schedule.
- Proper documentation of installation, maintenance, calibration, repair, and replacement activities.

STATE OF THE ART

This section highlights recently completed research studies and technological developments that are ready for implementation by highway agencies and for further advancement by academic and research institutions.

Advances in WIM Technology

Advancements in data processing speed, data storage, power requirements, and data communications have enabled WIM system deployment at remote locations, real-time data monitoring, real-time data analysis at the edge, and retention of all vehicle records (versus a filtered set in the past, due to storage capabilities).

WIM Systems with Add-On Capabilities

Recent advances include WIM systems with add-on capabilities that provide the following information:

- Video images of the vehicles linked to the measurement data for visual verification of vehicle characteristics.
- License plate numbers, dangerous goods identification shields, and vehicle registration information from automatic readers.
- Vehicle signatures from the inductive loop cards that identify the vehicles passing over the loops. Re-identification of the same vehicles and possible vehicle matching is possible at downstream sites with vehicle signature capabilities.
- Infrared images of the vehicles to detect overheating of the brakes or tires.

- Traffic wander (i.e., truck drifting within the lane), axle, and wheel position data to provide in-lane vehicle behavior information to roadway and pavement designers. The MnDOT and FDOT US301 test roads are equipped with tire footprint sensors.
- Systems linked to variable messaging signs, allowing rerouting of overweight or oversized vehicles for inspection or infrastructure protection.

Advances in WIM Data Usage

The section below details the recent advances in the use of WIM data.

Vehicle Re-Identification

The unique vehicle signatures obtained from the inductive loops, as well as the license plate and truck video images or RFID tags (radio frequency – identification transponders), can re-identify vehicles at downstream locations by a network of interconnected WIM sites. Comparison of the WIM measurements for the same vehicle at two or more locations could facilitate a quick comparison of sensor performance and detection of WIM system performance issues. The comparison can also provide valuable information for transportation planners and freight analysts. Vehicle re-id can also take place from WIM to non-WIM sites, making the rich WIM data available at volume, speed, and class sites.

Recent FHWA-led Small Business Innovation Research (SBIR) showed how WIM and non-WIM sites over a 200+ square mile area can be linked together with the re-id of vehicles. This linking uses vehicle signatures to transfer data between sites for site health monitoring, improved network data corridor travel characteristics, and WIM site and lane calibration checking. A follow-on FHWA-led pooled fund study TPF-5(520) is ongoing as of the writing of this E-Circular. Website: TPF - Study Detail (pooledfund.org)

WIM Data Use for Freight Planning

The freight community is seeking WIM data to improve characterization and modeling of truck movement and demands. These data can be coupled with other data sources to provide robust analyses and better understanding and decisions to enable improved movement of goods. The pooled fund study TPF-5(280) produced web-based traffic data visualization and analysis tools that offer data quality review and control functions along with data visualization capabilities and analysis. The study also produced data output controls to meet pavement design, freight analysis, and truck weight and load trend analysis. The TPF-5(280) study report contains information on how state agencies can use the tool to better understand the “cargo” component of WIM data (Lawson 2015). FHWA is using the TMAS data weight (500+ sites) and class (3,200+ sites) data in the Freight Analysis Framework (FAF) to help with model calibration and network results.

The 2023 TRB workshop on the fusion of WIM data and permits in different DOTs shows how WIM data can be merged with permit information to help agencies know how vehicles are using the roadway system. Recordings of the WIM/permit data fusion are also available from International Society of Weigh in Motion (ISWIM) for the follow-on training that was provided in 2023.

Effect of WIM Accuracy on Pavement Design (LTPP Findings)

LTPP has conducted sensitivity studies of the effect of WIM precision and bias on pavement design outcomes using the MEPDG method. The research findings recommend specifying ASTM 1318-09 Type 1 WIM systems as the means to collect WIM data for pavement design. Furthermore, the measurement bias should be kept as close to zero as possible through regular calibration. The FHWA LTPP TPF5(004) study shows that bias under 2% could be consistently achieved through regular WIM calibration for piezo-quartz, bending plate, and load cell WIM sensors. The results of sensitivity analyses also found that the increase in error due to bias is far more critical than the same error increase due to poor precision. Bias over 5% should be avoided in WIM data collected for pavement design purposes (FHWA-HRT-13-090, Final Report).

Truck Weight and Size Enforcement

The objective of weight enforcement is to achieve better compliance with truck weight and axle loading regulations. Better compliance would result in a reduction of overloading and its negative effects, including increased wear and tear of the road infrastructure, unfair trucking competition, and reduced road safety. Both low-speed (LS) and high-speed (HS) WIM systems offer a range of applications to assist in more efficient and effective weight enforcement. Applications include:

1. Roadside controls at fixed weigh stations;
2. Use of portable scales;
3. Use of WIM data for development of statistics and for planning of controls on overloading;
4. Pre-selection of vehicles that are likely overloaded for roadside controls;
5. Profiling and inspections of structurally overloaded companies; and
6. Direct weight enforcement using high accuracy HS WIM systems.

Use of WIM for truck weight enforcement is more widely used outside of United States. However, in recent years, U.S. interest in WIM data for enforcement has been steadily increasing. There is support for further development of NIST Handbook 44 and pilot projects exist in several states. The NYC DOT is piloting a program to use HS WIM for enforcement; this program is just getting started at the time of publishing this E-Circular.

WIM systems have been used for the enforcement of overloading in various European countries and by some U.S. states. In the UK, the Vehicle and Operator Services Agency has used WIM, in combination with automatic number plate recognition(ANPR) cameras, since 2007 as a pre-selection tool for roadside overload controls. In France and the Netherlands, the Transport Inspectorate uses the registrations of overloaded vehicles for the highway network of 15 WIM systems to identify companies with the highest and most frequent overloads. The company profiles are used for targeted inspections of companies with the most overloads.

In Slovenia, 15 Bridge-WIMs are used for an annual campaign of short-duration measurements at 150 locations on the national highway network. Most measurements are over 2–4 weeks, while a few are longer durations (6–12 months). The main end users are the

Slovene Ministry of Infrastructure, which is responsible for maintenance of the highway network, and the National Police Agency, which is responsible for weight enforcement.

In Switzerland, WIM systems verify the weight declarations by heavy goods freight vehicles for the highway usage fee for trucks (LSVA). The fee is based on the distance traveled and the loading capacity of each individual heavy vehicle (>3500kg).

The Czech Republic was the first European country to start using WIM in 2012. It uses WIM for direct automatic weight enforcement with a network that currently consist of 20 WIM systems. Another major implementation began in Hungary in 2019 with the installation of a network of more than 100 WIM systems (in total >240 lanes).

WIM Data for Screening and Sorter Applications

Use of WIM as sorters at weigh station facilities has advanced the ability to collect data and to validate system performance. This was documented in a Virginia Tech study (*19*) where researchers developed a tentative code for the use of WIM for screening and sorting applications and included the code in the NIST Handbook 44 (NIST 44).

ADVANCES IN WIM PROGRAM MANAGEMENT AND OPERATIONS

NCHRP WIM Tools to Improve Data Quality Using Holistic Approach

NCHRP has recently completed Project 20-50(20) to develop practical tools and guidance to improve accuracy of WIM data. The main products of this research study are six WIM tools and an accompanying practical guide. These products are designed to aid in activities such as WIM site assessment, sensor selection, site design, installation quality assurance, maintenance, performance troubleshooting, calibration, data quality review, and data acceptance (*27*). These tools are based on the findings from a survey of the best WIM practices in selected state highway agencies, in-depth analysis of the procedures developed and implemented by the FHWA LTPP TPF 5(004) study, guidance from the FHWA WIM Pocket Guides, and the requirements set forth in ASTM E1318-09(2017). The tools use analytical models developed from the analysis of WIM, pavement, road inventory, and climate provided by the FHWA LTPP program and participating state highway agencies.

Pavement Smoothness and WIM Smoothness Index

Pavement smoothness is critical to achieving acceptable system performance and is considered part of the WIM system design, installation, and maintenance process. FHWA-LTPP conducted the research investigation that led to the development of the AASHTO M331-13 Standard Specification, Smoothness of Pavement in Weigh-in-Motion (WIM) Systems. The pavement smoothness analysis is accomplished using the profile data (collected with a high-speed profiler) and the Optimal WIM Locator. This locator was developed as part of the FHWA-LTPP ProVal data analysis software to identify optimal sensor placement locations to maximize WIM performance (AASHTO M331-13).

To minimize the effect of vehicle dynamics on the WIM data accuracy, FHWA recommends using the double threshold sensor array with an 18-ft WIM sensor array spacing (2022 TMG).

LTPP Method for WIM Validation and Calibration

The FHWA LTPP program implemented a WIM validation and calibration method to ensure research-quality WIM data collection. The method computes the 95% confidence interval of measurement errors and compares it with the ASTM 1318-09 tolerances for Type I WIM systems. In addition, any systematic bias (mean measurement error) is computed and evaluated against a target value of no more than 2%. This LTPP method for WIM validation and calibration also incorporates the requirements for temperature, speed, and pavement smoothness. Virginia, Maryland, Wisconsin, and Arizona are testing or implementing the LTPP method, or some variation of the method, in their WIM validation and calibration.

WIM Data Integration

Efforts are underway to develop business processes at state highway agencies allowing WIM data sharing between multiple users (planning, design, freight, safety, environmental). Examples include:

- WIM data sharing between law enforcement and pavement design offices.
- Integration of WIM data in pavement warranties and related disputes.
- The FHWA Database for Air Quality and Noise Analysis (DANA) tool for the integration of WIM data for environmental studies.
- Use of WIM data to assess roadway network needs and loadings for many applications.
- Integration of WIM data into enterprise data plans to assist in vital DOT functions, including asset management, design, environmental studies, planning, research, safety, and weight enforcement.

GAPS IN PRACTICE AND KNOWLEDGE

Institutional Challenges

- Many state highway agencies collect WIM data for submittal to FHWA. However, there is still a limited application of these data sets for improving other functions at the state level, such as pavement and bridge design, load rating, freight studies, congestion, safety analyses, and agency-level transportation management decisions.
- Agencies are slow to take advantage of available WIM data or they lack the dedicated funding needed to establish and maintain modern WIM programs. The result is the application of default data in pavement design, bridge design, and environmental studies. This leads to increased uncertainty in the magnitude of over- or under-designs stemming from variability in traffic data estimates on different roadways.
- Despite considerable interest, Bridge Weigh-in-Motion (BWIM) capabilities are not being fully used in the United States. Further initiatives are needed to quantify BWIM benefits and limitations (as may exist due to U.S. bridge designs), to define its applications and document any limitations, and to advance BWIM use in United States beyond research.

- WIM data collection requires considerable technical skills and knowledge. The new NCHRP WIM tools (*27*) address many of these needs. Institutional and organizational support is needed to maintain the awareness in WIM community about the NCHRP WIM tools benefits and to provide tools for training and technical support.
- Despite AASHTO having identified it as priority technology over a decade ago, not all agencies have allocated support to build and sustain effective WIM programs. This includes doing short-term measurements (1 week to several months) to obtain an indication of the (over-) loading situation and doing structural strength/safety assessments of certain bridges. Sufficient management support and prioritization are necessary to obtain the needed staffing, skills, and resources to make WIM a standard production tool. The FHWA Office of Highway Policy Information is actively working to address this across the U.S. with recent successes in TN, NH, and UT.
- Linking WIM data by state agencies and sharing those data with commercial vehicle operations is currently lacking, including WIM data analysts' access to information regarding vehicles with special permits. This access is needed to assess validity of WIM data or to retrieve truck transponder data.
- Long-term use of weigh station data for activities other than enforcement needs to be improved. Weigh station data is some of the most accurate weight data available yet is rarely used for more than enforcement activities.
- Limited availability of automated programs to check data quality and reduce the dependence on staffing and institutional knowledge.
- Lack of coordination or transparency among end users of WIM data and the different organizations involved with the installation, operation, maintenance, and financing of the WIM systems.
- The methods to merge state DOTs' planned WIM data collection with weight enforcement activities should be explored and implemented. This would be a way of combining resources for more comprehensive use of WIM locations, such as those employed by the TN DOT as part of their new 30-site installations.
- Limited data sharing between collectors and users (including saving data from enforcement in-line sorter systems for other applications such as design and management of infrastructure). Find ways of using the U.S. DOT NTAD, FHWA TMAS (500+ sites), and FHWA LTPP data sets for WIM-needed activities. Improve the availability of weight data from enforcement scales. Even though it is a skewed data set, this would be a highly useful data set for agencies if it were only kept and made available.
- In practice, some states do not perform field validations of WIM performance (with enough test truck runs or frequency of calibration) due to high expense. There is a need for alternative solutions for optimized use of available limited resources. FHWA pooled fund TPF-5(520) hopes to address parts of this need.

- Encourage the storage of and more detailed data sets from WIM sites to make the data more useful. This could include LTPP data records of all vehicle data or the FHWA 2022 TMG IVR formats that allow for pavement temperature, vehicle signatures, 1/100th of a second time stamp, and axle weight data storage by wheel path. By increasing the detail and information available from WIM sites, the value of each WIM site is increased without the need for any changes for the in-road sensor side of WIM installations.

EMERGING TRENDS AND DRIVERS OF CHANGE

The recognition of the value of WIM data continues to grow in the transportation field. We live in the world of data analytics, data mining, data visualization, and instant data access from any location. These changes will affect how WIM data are collected and exploited.

Highway Design and Management Needs

The paradigm shift in U.S. pavement design practice from the empirical to mechanistic-empirical method serves as a driver for a renewed interest in WIM programs within state highway agencies as they adopt the MEPDG method into their traffic loading environment. This shift has driven interest in better defining WIM equipment performance requirements (for truck weight and axle load data accuracy) and has created a higher emphasis on WIM data quality. WIM data coverage requirements that enable agencies to acquire the truck weight and axle load data they need without over or underrepresenting the network also receive additional interest from practitioners.

Re-identification of specific vehicles between sites using Bluetooth, transponders, or vehicle signatures will aid in more advanced use of WIM site data and improved understanding of traffic patterns along corridors. FHWA pooled fund TPF-5(520) aims to accelerate adoption of one of these re-identification technologies (vehicle signatures with 100+ classes of trucks) to accomplish this goal.

Transportation Asset Management

In the United States, the ongoing implementation of the data-driven performance measures in transportation asset management is creating a renewed interest in WIM data and serves as a driver of change. The management initiatives that will be improved with the use of traffic loading and vehicle classification data collected by WIM include:

- Pavement design and management.
- Bridge design, management, and load rating.
- Commercial vehicle operations (enforcement).
- Safety and systems analyses.
- Freight planning and operations.
- Commerce.
- Environmental analyses for air quality and noise.

Effectiveness and sustainability of the WIM programs could be further improved by

- Development and preservation of institutional knowledge through creation of operations manuals, guides, and instructional and training materials.
- Improved efficiency of WIM operations through use of robust tools based on current standards and best practices.
- Better field staff training of how to properly install, collect, maintain, and calibrate WIM sites.
- Improved ability to collect and share data across institutional departments.
- Improved definition of coverage requirements (a minimum number of locations to meet the greatest number of needs).
- Upper-level management understanding of the need and support for WIM programs.

Opportunities Related to Smart Road Infrastructures, Connected, and Autonomous Vehicles

Highway transportation systems and infrastructure are evolving from physical systems to hybrid physical and digital systems. Transportation modalities are expanding to include connected and autonomous vehicles (CAVs). Several emerging technologies serve as drivers of change in how vehicular data will be collected in the future, including the Internet of Things, in-vehicle and onboard data collection, CAVs, smart road infrastructure, and ITS.

WIM Sites as Part of Internet of Things

The Internet of Things (IoT) is made up of smart electronic devices and various sensors connected to and sharing data through the internet. Existing and emerging WIM technologies can be integrated within the IoT universe and use IoT techniques for data processing and data transmission. Data collected at a WIM site (at the IoT edge) can then be shared with other smart and connected devices including onboard truck communication devices or CAVs identified within a certain proximity of a truck with detected issues (such as a flat tire). It could also be communicated to a connected police vehicle, allowing that vehicle to react and respond more quickly to identified issues with a particular truck or to a rapid change in traffic conditions. WIM sites have been used as an IoT “device” for over a decade. For example, the state of Maryland has implemented a technology that uses WIM sensors to detect an overloaded truck, capture a video image of the truck, and share the collected information about the truck (weight, truck image, and license plate) with a police patrol car via web-based application. These actions allow an immediate corrective action to be taken (FHWA 2017).

Use of In-Vehicle Data Collection Technologies

With the development of smart, connected, and autonomous vehicles, multiple in-vehicle sensor technologies are being developed and used. The most frequently used in-vehicle sensors include accelerometers, radar, LiDAR, and GPS. These sensors collect information on vehicle self-diagnostics, the environment, surrounding vehicles, traffic speed, headway, and location.

The collected information informs drivers about road and weather conditions, diagnoses problems with the vehicle, assists with vehicle operation and driving, prevents collisions, and avoids unsafe situations.

In the future, in-vehicle sensors could provide road managers with information, such as probe vehicle data, instantaneous vehicle speed, vehicle trip origin and destination, and a unique vehicle identifier. Data about a driver's behavior and the vehicle condition can also be used for forensic crash investigation. Depending on the application, data can be made anonymous to protect the privacy of the driver and vehicle, which is likely to increase participation from transportation companies. As CAV technologies develop and mature, these data will be collected and communicated by the systems and sensors integrated in the vehicle design. It is likely that, in the coming years, WIM data collection and reporting will be done by the CAVs themselves. In anticipation of this development, it is important to develop data quality assurance; data aggregation, storage, and sharing standards; data collection and reporting protocols; procedures to safeguard privacy of participating drivers and transportation companies; and strategies for technology implementation.

WIM Integration with CAV Technologies

In-vehicle sensor data could greatly improve knowledge of the travel stream loading characteristics. Automated enforcement is another application of WIM technology in the emerging CAV world. As CAV technology evolves, in-vehicle sensors integrated within CAVs can be used in addition to infrastructure-based sensors for automated weight enforcement.

Onboard WIM (OBW) technology provides a current example of how in-vehicle data collection technology can improve highway freight transportation efficiency. OBW uses a transponder to communicate the truck weight measured by OBW sensors to a central server. In the future, OBW technology can be integrated with CAVs to ensure the safety of CAVs. Truck weight data from OBW systems can be used to determine each truck's safe braking distance at different speeds. Surrounding CAVs' vehicle weight and speed data can be also used to calculate a safe distance among CAVs. In the case of non-connected vehicles, the vehicle weight data from OBW can be used with the adaptive cruise control's distance radar information to determine the safe distance among vehicles.

CURRENT AND PROPOSED NATIONAL RESEARCH INITIATIVES

This chapter describes existing and proposed research to addresses gaps in practice and knowledge for pavement engineering applications related to traffic monitoring.

Current Research

Emerging WIM Sensor Technology

Technology to capture various data elements related to WIM continues to advance. The section below documents some of these advancements and potential research initiatives.

Testing of Fiberglass WIM Sensors in the Netherlands Recent advancements in fiberglass WIM sensors aim to overcome traditional problems with durability under harsh highway conditions (Karabacak 2019). Several pilot projects are taking place in New Jersey (USA) and the Netherlands (EU).

Pilot with Stress-In-Motion Sensors in the Netherlands In the Netherlands, two pilot projects are running with International Road Dynamics' (IRD) Stress-In-Motion systems or TACS (Tire Anomaly and Classification System). The first project started in 2019 on highway A16 near the city of Dordrecht. The second project, on national road N279 (traject Asten-Veghel), started in 2020. The objective of both projects is to evaluate the quality of the measurements and their effectiveness in crash reduction. A study by the Stichting Incident Management Vrachtauto (STIMVA) showed that 41% of breakdowns with trucks were caused by insufficient tire pressure. During the pilots, between 1% and 2% of all trucks showed an irregular tire pressure; however, the direct effects on the reduction of crashes has not been demonstrated yet. In 2023 the Stress-in-Motion (SIM) system of the second pilot will be renewed; the first results are expected in 2024.

The innovative SIM systems expand traditional uses of WIM data to include traffic safety and efficient tire management. In combination with the existing uses of WIM data for traffic monitoring, pavement design, traffic loading, and weight enforcement, these new WIM data applications show high potential of using WIM systems for multiple-user groups and, thus, a high value of installing and maintaining WIM sites. The difficulty may lie in bringing together the end users from different organizations.

WIM Technology to Measure Truck Wander (IRD VectorSense) The MEPDG pavement design considers position of truck tires on the pavement. This is particularly important for determining thickness of concrete pavement slabs. Two state highway agencies (Minnesota DOT and Florida DOT) are installing additional new sensors capable of measuring tire position and type for each vehicle crossing the sensor. These agencies will collect truck wander data on highway pavement test roads (MnROAD and FDOT US 301 Test facility) and use the collected data to support pavement research studies.

Advances in WIM Data Usage for Weight Enforcement Europe continues to embrace use of WIM data for direct enforcement. The Walloon region of Belgium has approved its first WIM system for direct enforcement. Preparations are also under way in France, Poland, Germany, and Estonia. The United States has been slowly showing more interest for this application as well.

Research in WIM Program Management and Operations

WIM Sensor Testing Facilities Currently, at least two state highway agencies have WIM testing facilities. Florida DOT has a unique WIM testing facility that provides for side-by-side comparison of the performance of different WIM sensors, WIM controllers, power and communication devices, and installation practices. Minnesota maintains the MnROAD testing facility that includes WIM testing (http://www.dot.state.mn.us/mnroad/data/traffic.html). The long-term goal of this research is to evaluate the change in performance of the sensors and system over time and the lifecycle cost of the different systems. In addition, MnROAD set up a testing facility for analyzing benefits of multi-sensor WIM systems. MnDOT personnel will generate an annual report documenting the performance of the systems. Both FDOT and MnDOT currently conduct WIM experiments to test different WIM sensor arrays for collecting high accuracy WIM data; they also test traffic wander, lane position, wheel, and axle position sensors. Sensor arrays are being installed in 2023 and 2024.

NCHRP LTPP WIM Data Analysis Research Project and WIM Tools NCHRP Project 20-50(20), “LTPP Data Analysis: Develop Practical Tools and Procedures to Improve WIM Data Quality” has produced a practical guide and tools to improve WIM data accuracy and to increase reliability through enhanced processes for WIM system selection, site selections, installation quality assurance, calibration and maintenance, data analysis methods, and data QC/QA procedures. A follow-up project to implement these tools on the web and to integrate with FHWA web portal InfoPave is being planned.

Costs and Benefits of Site-Specific WIM Data Collection for Pavement Design As state highway agencies implement MEPDG method for their pavement designs, the costs and benefits of site-specific WIM data collection come into focus. FHWA has recently conducted a study that compared the costs associated with site-specific WIM data collection to the pavement costs resulting from use of WIM-based site-specific and default axle loading data. The researchers found that the costs to install, operate, collect data, and interpret WIM data results are small compared to pavement construction costs. For example, about 0.5 to 1 inch in asphalt layer thickness savings (from the use of site specific vs. default load data) for a 1-mi project with two 12-ft-wide lanes are enough to offset the cost of the WIM data collection. Similarly, 0.25 to 0.5 inches in asphalt thickness savings is enough to offset the cost of WIM data collection for a 2-mi project with two 12-ft-wide lanes (*28*).

Proposed Research

Remote WIM Sensor Calibration Using Connected Vehicles, On-Board WIM Systems, and Smart Road/WIM Infrastructure

Exploratory research is necessary in the area of WIM controller-to-vehicle connections. This research would explore the possibility of communication between a WIM controller and either onboard vehicle sensors or an onboard WIM communication device to facilitate remote WIM system calibration. This approach may eliminate the requirement for a WIM technician to be on-site to perform the calibration. This technology could reduce the cost of calibration, reduce the time a WIM technician spends in calibration activities, and improve safety of field personnel. Research conducted in Australia on the use of onboard vehicle weighing systems and WIM systems shows the benefits of using both types of devices; both types of WIM systems would operate within specifications and generate warnings when a system does not meet the requirements. A pilot site should be installed to investigate the practical procedures for data exchange and to show the benefits to all stakeholders. Additional Australian research on the use of OBW vehicle weighing systems and HS WIM systems supports the benefits of using both types of devices. Further, to address the need for practical procedures, research is needed to define data users and data needs, to establish common data quality and privacy standards, and to develop standard protocols for vehicle-to-smart-infrastructure communications and data sharing/exchange.

Road Safety Improvement Using Connected Vehicles and WIM with Tire Anomaly Detection Sensors

Road safety can be improved by expanding a standard HS WIM system with tire anomaly detection sensors and IoT for direct communication to road users. The main benefit of such a system would be the improvement of traffic safety for all heavy goods vehicles and for CAVs in particular. A pilot implementation of this technological solution is needed to test its practicality and to define a standard specification for practical implementation. The specification should include user needs; technical equipment specification; data collection, quality, and privacy requirements; and data exchange protocols for vehicle-to-smart-infrastructure communications. A pilot project would show the benefits to all stakeholders.

WIM to Support Autonomous Truck Weight and Size Enforcement for Connected Vehicles

Currently, advanced WIM systems provide the most comprehensive set of inputs for accurate classification of vehicles, including vehicle length, wheelbase length, number of axles, spacing between axles, weight per axle, and GVW. The additional technologies being implemented in "Super WIM" sites provide additional information about the vehicle, including 3D vehicle profile, video image, license plate, and DOT registration numbers.

As more detailed information about the vehicle becomes available, the WIM system can more precisely check for certain errors or infringements. For example, a smart weight and size enforcement system installed at a "Super WIM" site would be able to automatically verify if a vehicle meets federal and local legal load limits for that class of vehicles. If the WIM system can also identify the specific vehicle, the individual loading limits for that vehicle can be verified. If wireless general packet radio service (GPRS) communication between the WIM system and the vehicle is available, then the driver can be informed directly of the issue. In addition, the WIM system can perform an independent verification of the vehicle classification by comparing its data collected with the data accessed from the vehicle identification record. Research is needed to pilot-test implementation of this technology.

Portable WIM Solutions

Portable WIM systems are needed for various reasons to provide site-specific loading characteristics for local agency needs and DOT uses. In the United States, most agencies are not able to obtain acceptable WIM results using portable systems. Bridge WIM systems provide an option to perform short-term measurements (1–4 weeks). After a short measurement, a system can be moved to the next location for another short-term measurement. In this way, one system can generate an overview of the (over-) loading situation on a part of the road network instead of just one location.

WIM Data for Strength and Safety Assessment of Existing Bridges

During their lifetime, bridges deteriorate while their loading increases. Both factors affect bridge safety. Bridge failures are unacceptable as they can cost lives and undermine confidence in the entire transportation infrastructure. Bridge safety assessment requires the actual on-site traffic loading information to determine the corresponding response of the bridge structure. The actual on-site traffic loading information can only be obtained by BWIM systems (*26*). The measured loading information obtained from BWIM and from traditional in-road WIM systems located near bridges can be used to verify, calibrate, and, if necessary, adjust the load and resistance factors or safety factors used in bridge design codes. BWIM data can also be used for advanced bridge health monitoring and bridge management systems.

REFERENCES AND OTHER RESOURCES

1. LTPP Field Operations Guide for SPS WIM Sites, Version 1.0, Office of Infrastructure Research, Development, and Technology, Federal Highway Administration, U.S. Department of Transportation, McLean, Virginia, May 2012.
2. Cebon, D. *Handbook of Vehicle-Road Interaction*, Swets and Zeitlinger, Lisse, NL, 1999, p. 515.
3. E1318-09: Standard Specification for Highway Weigh-in-Motion (WIM) Systems with User Requirements and Test Methods, ASTM Committee E17 on Vehicle Pavement Systems, American Society for Testing and Materials (ASTM), West Conshohocken, PA, 2009.
4. European Cooperation in Science and Technology (COST): Action 323. 1998. Weigh-in-Motion of Road Vehicles, Paris, 2002.
5. NMi International WIM Standard: Specification and test procedures for weigh-in-motion systems, NMi Certin, Dordrecht, NL, 2016. http://www.nmi.nl/nmi-wim-standard/.
6. AASHTO M331-13 Standard Specification, Smoothness of Pavement in Weigh-in-Motion (WIM) Systems, American Association of State Highway and Transportation Officials, Washington, DC.
7. Implementation of ProVAL OWL—A Tech Brief, FHWA Surface Enhancement Team and ProVAL Support Team. www.roadprofile.com/download/Implementation-of-OWL.pdf. Accessed Sep. 28, 2017.
8. Transtec Group. *ProVAL User's Guide*, ProVAL 3.6 Profile Viewing and Analysis Software, Austin, TX, 2016. www.roadprofile.com/download/ProVAL-3.60-Users-Guide.pdf. Accessed Sep. 28, 2017.
9. Papagiannakis, A. T., R. Quinley, and S. R. Brandt. *NCHRP Synthesis 386: High-Speed Weigh-in-Motion System Calibration Practices*. Transportation Research Board of the National Academies, Washington, DC, 2008. https://doi.org/10.17226/23062. Accessed Sep. 28, 2017.
10. I-Loop Duo Signature Detection Card, CLR Analytics, Irvine, CA. http://clranalytics.com/content/i-loop-duo-signature-detector-card. Accessed Sep. 29, 2017.
11. Hallenbeck M. E., Selezneva, O. I., and Quinley, R. Verification, Refinement, and Applicability of LTPP Vehicle Classification Scheme, FHWA Report FHWA-HRT-13-091, Federal Highway Administration, U.S. Department of Transportation, Washington D.C., 2014.
12. Lawson, C. T. Web-based Traffic Data Visualization and Analysis Tools, TPF-5(280), Task 5 Final Report, Federal Highway Administration, U.S. Department of Transportation, Washington, DC, 2015. http://www.pooledfund.org/Details/Study/516. Accessed Sep. 28, 2023.
13. Selezneva, O. and Von Quintus, H. Traffic Load Spectra for Implementing and Using the Mechanistic-Empirical Pavement Design Guide in Georgia, FHWA-GA-14-1009, Atlanta, Georgia, February 2014.

14. Buch, N., Haider, S.W., Brown, J., and Chatti, K. Characterization of Truck Traffic in Michigan for the New Mechanistic Empirical Pavement Design Guide, Research Report RC-1537, Michigan State University Department of Civil and Environmental Engineering, East Lansing, MI, December 2009.
15. Bhattacharya, B., Selezneva, O., and Peddicord, L. Development of Traffic Inputs Library in Pennsylvania for the Use in AASHTOWare Pavement ME Design Software, Proceedings of the International Conference on Highway Pavements and Airfield Technology 2017, Philadelphia, PA, August 27–30, 2017.
16. Selezneva, O. and Hallenbeck, M. Long-Term Pavement Performance Pavement Loading User Guide (*LTPP PLUG*), FHWA-HRT-13-089, Federal Highway Administration, U.S. Department of Transportation, Washington DC, October 2013.
17. Selezneva, O.I., Ayres, M., Hallenbeck, M., Ramachandran, A., Shirazi, H., and Von Quintus, H. MEPDG Traffic Loading Defaults Derived from LTPP Traffic Pooled Fund Study, FHWA-HRT-13-090, Final Report, Federal Highway Administration, U.S. Department of Transportation, Washington DC, 2016.
18. ProVAL Software. http://www.roadprofile.com/proval-software. Accessed Sep. 28, 2017.
19. Katz, A.J. and Rakha, H.A. Final report of ITS Center Project: Weigh-in-Motion Evaluation, Virginia Tech Transportation Institute, Blacksburg, VA, Jan. 2002. https://vtechworks.lib.vt.edu/bitstream/handle/10919/55111/WIM-Evaluation-FINAL-REPORT.pdf?sequence=1. Accessed Sep. 28, 2023.
20. Specifications, Tolerances, and Other Technical Requirements for Weighing and Measuring Devices, NIST Handbook 44, Butcher, L. Crown, and R. Harshman (eds.), National Institute of Standards and Technology, Washington, DC, 2017. https://www.nist.gov/document/hb44-2017-webfinalpdf. Accessed Sep. 28, 2017.
21. Applied Research Associates, Inc. NCHRP Project 01-37A, “Development of the 2002 Guide for the Design of New and Rehabilitated Pavement Structures,” Final Report and Software (Version 0.70) Transportation Research Board, Washington, DC, April 2004.
22. MnROAD Traffic, Minnesota Department of Transportation, St. Paul, MN, 2017. http://www.dot.state.mn.us/mnroad/data/traffic.html. Accessed Sep. 28, 2023.
23. Koniditsiotis, C., Susilo, A., and Cai, D. Weigh-In-Motion and ITS: Heavy Vehicle On-board Weighing Using Intelligent Access Program, 1st International Seminar on Weigh-In-Motion, Florianópolis, Santa Catarina, Brazil, April 3–7, 2011. http://www.labtrans.ufsc.br/media/42949/s3pt09_chris_koniditsiotis.pdf.
24. Load and Resistance Factor Design (LRFD) for Highway Bridge Superstructures, NHI Course No. 130081, 130081A, and 130081B, Publication No. FHWA-NHI-15-047, Federal Highway Administration, U.S. Department of Transportation, Washington DC, Revised July 2015. https://www.fhwa.dot.gov/bridge/pubs/nhi15047.pdf. Accessed Sep. 28, 2017.
25. Federal Highway Administration. *Traffic Monitoring Guide*. Washington, DC: Federal Highway Administration, U.S. Department of Transportation, 2022. Available online: https://www.fhwa.dot.gov/policyinformation/tmguide/, last accessed June 6, 2023.
26. Žnidarič, A., Kreslin, M., Lavrič, I. and Kalin, J., 2012. Modelling Traffic Loads on Bridges - A Simplified Approach Using Bridge WIM Measurements. Dallas, Texas, London: ISTE; Hoboken: Wiley, pp. 418-428.
27. Selezneva, O., and D. Wolf. *NCHRP Research Report 1070: Tools for Assuring WIM Data Quality: Practical Guide*, Transportation Research Board, Washington, DC, 2023. Available online: https://doi.org/10.17226/27234, last accessed September 6, 2023.
28. Tech Brief: Exploring the Importance of Traffic Data Input Levels for Mechanistic-Empirical Pavement Design, 2021. FHWA Publication No.: FHWA-HRT-21-046. Available online: https://highways.dot.gov/media/1316, last accessed September 26, 2023.

Travel Time, Speed, and Reliability Data

Xu Zhang
Kentucky Transportation Center

Xiaofeng Li
University of Hawaii at Mānoa

Brian Yueshuai He
University of California, Los Angeles

Eugene Antwi Boasiako
University of Kentucky

Lawrence A. Klein
Klein and Associates

Mena Lockwood
Virginia Department of Transportation

Sam Granato
Ohio Department of Transportation

STATE OF THE PRACTICE

Travel time and speed are fundamental metrics for measuring and monitoring traffic conditions on roadways. Travel time refers to the duration a vehicle takes to travel from point A to point B, while speed is the rate at which the vehicle is traveling. Depending on data collection methods, speed can be categorized as either space mean speed, which is the average speed of vehicles traveling over a specific road segment, or as time mean speed, which is the average speed of all vehicles at a particular location during a certain time period. Travel time reliability (TTR) quantifies the consistency and predictability of travel times on specific routes or road segments over time under the influences of various non-recurring events. It often necessitates the availability of travel time data for an extended period, such as a whole year, to capture different day-of-week, monthly, and seasonal traffic patterns (*1*). Many transportation agencies have recently started using TTR to identify areas prone to non-recurring congestion or delays and to develop mitigation strategies (*2*). The travel time, speed, and reliability data are essential for the operation, monitoring, and planning of efficient transportation systems. Some of the elements in this section, namely the use of probe data for volume estimation, are expanded upon in the Probe Data for Traffic Volume Estimation section.

Advancements in data collection technologies have led to a rapid expansion of data sources for travel time, speed, and reliability data in recent years (*3*). Data collection methods can be categorized into three groups: point sensors (e.g., inductive loop detectors, microwave radar, infrared, Doppler, ultrasonic, acoustic sensors, and video detection systems), point-to-point sensors (e.g., automated vehicle identification systems based on Bluetooth, WiFi, toll tag devices, or license plate readers), and area-wide sensors (e.g., smart phones and GPS technologies) (*4*). Detailed descriptions of the characteristics of these data collection methods can be found in *Transportation Research Circular E-C227: Advancing Highway Traffic Monitoring Through Strategic Research* (*5*) and are omitted here to avoid repetition. Figure 4 also provides additional discussions into several common data collection methods and their specific applications.

The installment, operation, and maintenance of point and point-to-point sensors generally require significant resources from transportation agencies (*6*). Limited funding often results in major highways, such as freeways in urban areas, being well instrumented, while urban minor arterials and rural highways are less covered. A well-known example of point sensors is the Caltrans Performance Measurement System (PeMS), which collects real-time traffic information, including speeds from thousands of loop detectors, across the freeway system in California. In a different application, Georgia DOT installed 35 Bluetooth units along several main arterials around I-75 in Cobb County and Fulton County. Collected travel time data were used to monitor the regional arterial network for incidents and determine the impact of I-75's heavy congestion on surrounding arterials.

In contrast, GPS-based technologies collect data with much wider geographical coverage. As the GPS-enabled devices proliferate, the availability and quantity of speed or travel time data have greatly increased (*7*). The purchase and use of private-sector probe data have therefore become common practice in many transportation agencies. According to a recent survey discussed in *NCHRP Synthesis 561*, at least 88% of the forty-five responding state DOTs are using probe speed/travel time data; the most prevalent use case is for federal performance reporting using NPMRDS data (*7*). Figure 5 shows the number of state DOTs responding to the same survey regarding probe data sources.

Transportation agencies make extensive use of probe data in a variety of applications, as highlighted in two recent NCHRP Synthesis studies (*6, 7*). In real-time applications, the data is commonly used to provide the public with travel times on dynamic message signs, display traffic information on traveler information websites and on mobile apps, display highway conditions at traffic management centers, detect and alert to queues resulting from non-recurring incidents, and establish variable speed limits. Archived historical speed or travel time data are used to develop delay and reliability performance measures, identify congested corridors, prioritize transportation improvement projects, conduct before-and-after studies to evaluate the effectiveness of major projects, and calibrate and validate statewide and MPO travel demand models.

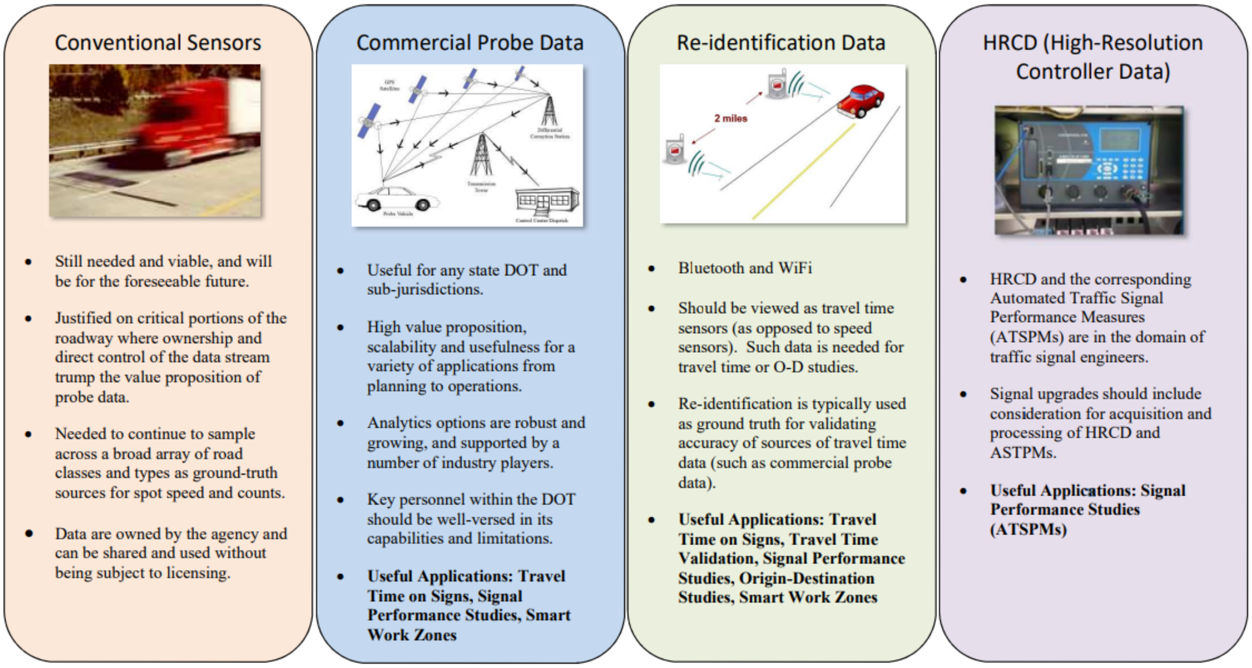

FIGURE 4 Description of different data collection methods (Source: I-95 Corridor Coalition 2019).

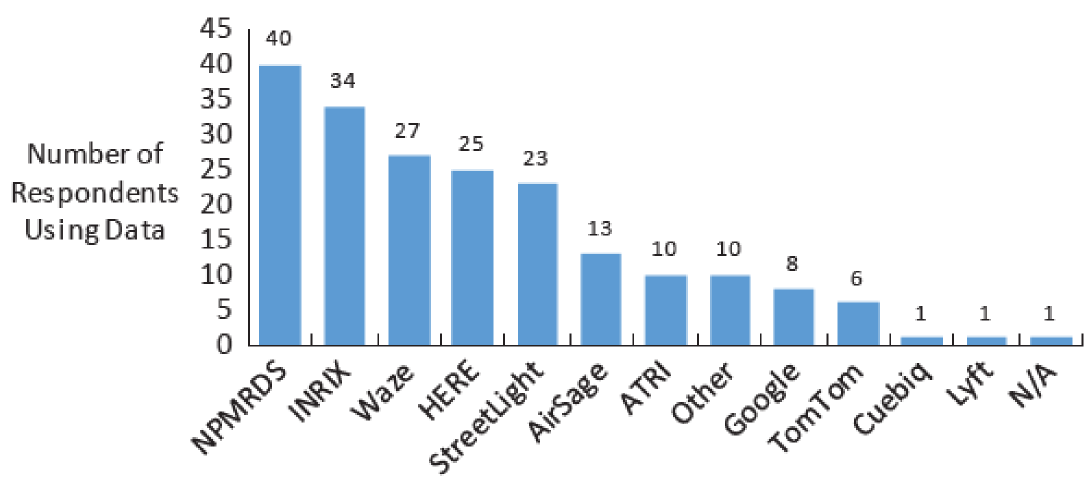

FIGURE 5 Histogram of probe data sources used by state DOTs (Source: *NCHRP Synthesis 561*).

There are several challenges and caveats associated with using probe speed data in practical applications. First, data quality remains a significant concern for agencies, ranking as the primary issue in both synthesis surveys (*6, 7*). Data coverage is generally sparser on arterials, collectors, and local roads due to the lower number of probe vehicles sampled compared to freeways (*8*). Although congestion is less likely to be an issue for these roads, the utility of probe vehicle data for systemically tracking their performance can be limited.

Second, data vendors often consider both their data sources and their methodologies for cleaning, processing, and aggregating the data to be proprietary. Moreover, they frequently alter their data provider mixes and data processing algorithms, which can result in performance trends that deviate from actual conditions (*9*). This lack of transparency is a key reason why agencies are hesitant to acquire and use probe data.

Third, the large quantities of data often encountered pose challenges to agencies' legacy data management systems. Integrating these data with existing agency data systems also requires staff members with the necessary skills. Some DOTs have chosen to use services provided by data analytics suites such as RITIS and iPeMS or to migrate data to cloud platforms such as Google and Amazon.

Another significant challenge is the use of proprietary data for travel monitoring, which is attached to a proprietary network that may differ from the networks maintained by state and local agencies (*10*). The integration process requires extensive knowledge of geographic information systems (GIS) and specialized tools, making it particularly challenging for agencies, cities, and municipalities with limited GIS resources and assets.

STATE OF THE ART

Trajectory Data

An increasing number of transportation agencies have begun to leverage raw GPS trajectory data in addition to origin and destination (OD) and link-level aggregated probe vehicle data. The trajectory data provides valuable information that was previously unavailable from other data types, such as specific routes and travel times taken by individual vehicles, variation in speeds, and travel times between trips on each route. Figure 6 illustrates individual trajectories as well as their aggregated travel times and speeds on US-31 in Holland, MI (*41*).

Because probe-based trajectory data and link-level aggregated probe speed data originate from the same data sources, the trajectory data has equivalent spatiotemporal coverage to link-level speed data. However, the trajectory data's size expands considerably due to the inclusion of waypoints for each individual vehicle trip. Consequently, processing such data demands intensive computation, posing a significant challenge in its use.

Two other notable obstacles to using trajectory data are the low penetration rate and infrequent updates (*12*). Several methodologies have been proposed to effectively leverage low-penetration probe data for travel time estimation. For example, a deep learning method, eq2seq graph convolutional neural networks (GCN) and long short-term memory (LSTM) was proposed by Abdelraouf et al. (*13*) to use connected vehicle data from a particular vendor with a penetration rate of 2%-5% to estimate traffic speed. Li and McDonald (*14*) proposed a fuzzy logic method that uses information such as instantaneous speed and acceleration provided by individual vehicles to estimate traffic speed more accurately. Various statistical and machine learning methods have also been proposed for obtaining travel time data from probe vehicle data with low update frequency. For example, a statistical model estimates travel time on links and delay at intersections using the probe trajectory data at a 2-minute update frequency (*15*). Additionally, an artificial neural network-based method estimates traffic speed using sparse probe vehicle data, with sensitivity analysis highlighting the critical role of positional information from probe vehicle data in model performance (*16*).

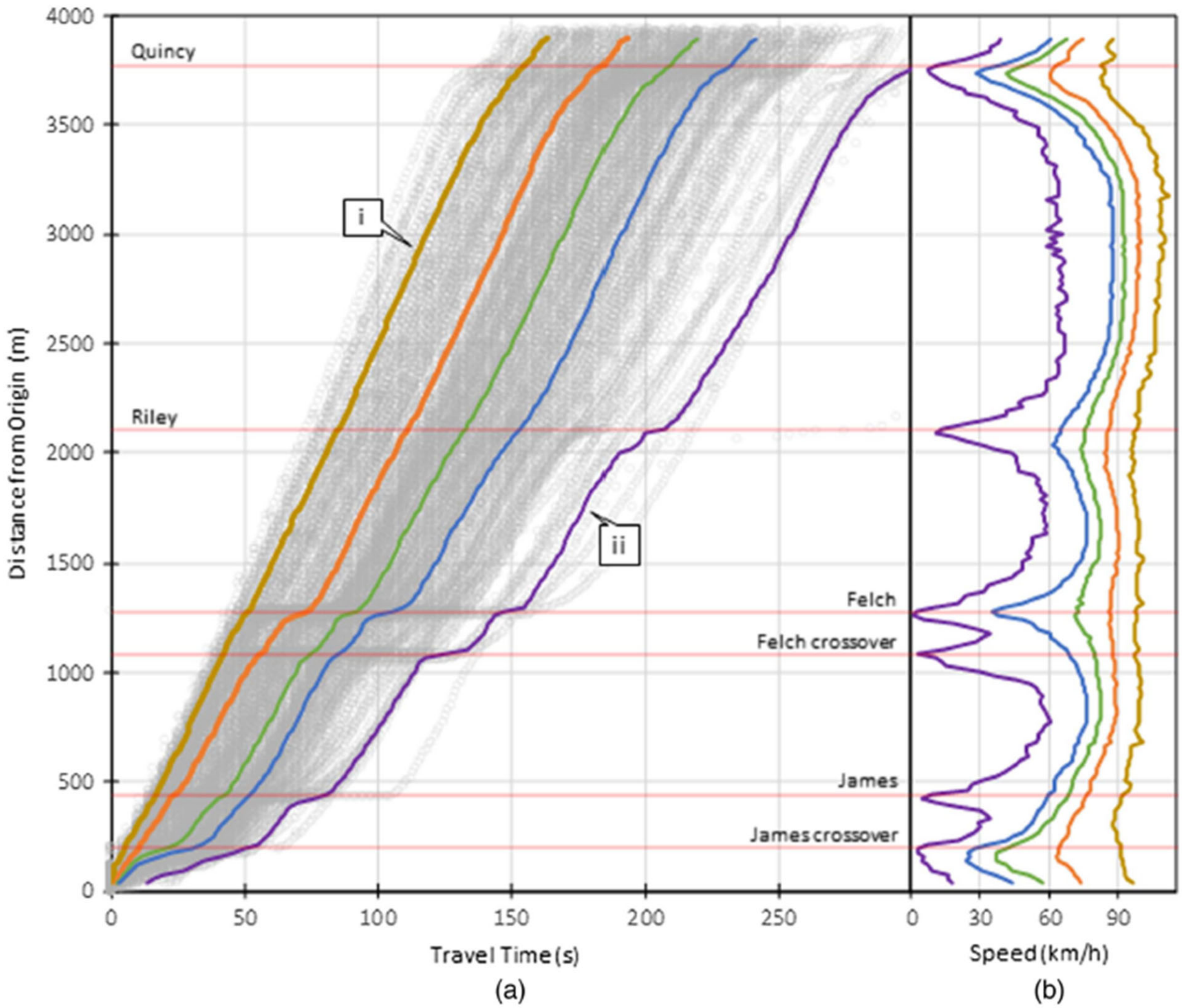

FIGURE 6 An example of GPS trajectory data on US-31 in Holland, MI (*41*).

Cellular Data

In addition to vehicles equipped with GPS devices, cellular phone data offers an alternative source of position data through cellular communication networks. However, cellular phone data presents a limitation in that it provides location area or cell information rather than precise positioning. To address this limitation, Ou et al. (*17*) proposed an algorithm based on a water container prototype that integrates cellular phone data and probe vehicle data for estimating freeway travel time. Sensitivity analysis revealed that the algorithm's accuracy is significantly influenced by the penetration rate of probe vehicles, though specific penetration rate thresholds were not provided in their research.

LiDAR Data

LiDAR sensors offer several advantages in traffic data collection, including their ability to operate irrespective of visible light and their resilience to environmental conditions, resulting in highly precise detection over an extended range. Consequently, numerous studies have explored the potential of this emerging technology for obtaining traffic data, including traffic speed and travel time. As the production of LiDAR sensors has scaled up, their costs have begun to decrease, attracting increased attention from transportation practitioners. Nevertheless, a limitation of using LiDAR sensors is the presence of noisy data, as these sensors not only collect vehicle data but also capture information detecting other objects in the environment. To extract valuable information, Petrovskaya and Thrun (*18*) proposed a complex model to track the speed of vehicles using multiple filter technologies. Furthermore, multiple methods (*19, 20*) have been proposed for detecting and tracking vehicles, enabling the provision of vehicle speed data.

EMERGING TRENDS AND DRIVERS OF CHANGE

The rapid advancement of Internet of Things (IoT) and Deep Learning/Artificial Intelligence (DL/AI) technologies has brought about a revolutionary shift in the research on travel time, speed, and reliability (TTSR). This transformative development has led to the emergence of two key focal points: innovative data sources and advanced analytics methods. These emerging trends are reshaping the way we understand and address TTSR-related challenges, opening up new possibilities for improving transportation systems.

One development in vehicle technology is the advent of advanced connected vehicles (CVs). CVs, equipped with multiple sensors, provide real-time information on vehicle location, status, and speed, making data accessible from prominent car manufacturers and third-party data vendors (*21, 22*). Such CV data has emerged as a crucial probe-based source for TTSR research. Notably, Purdue University and Indiana DOT conducted a case study using such data to enhance work zone safety and gain insights into traffic conditions (*23*). Furthermore, a comprehensive national dataset based on CV data was established for traffic analysis across all interstate routes in the United States (*24*).

An additional noteworthy trend pertains to the analytical methods employed. The progress in deep learning techniques and computational capabilities has enabled the analysis of large-scale datasets, encompassing probe data, CV data, and social media data, among others. Data-driven methodologies and deep learning approaches are used extensively for predicting travel speed on traffic networks. Table 9 provides an overview of important studies in this domain. Notably, deep learning methods such as convolutional networks, long short-term memory networks, and generative adversarial networks are gaining prominence, surpassing traditional data analytical techniques in performance.

TABLE 9 Representative Studies on Deep Learning Methods for Predicting Traffic Speed

Model	Data	Characteristics
Attention graph convolutional sequence-to-sequence model (AGC-Seq2Seq) (25)	Smartphone-based application data	Spatial and temporal dependencies are modeled through the Seq2Seq model and graph convolution network separately. The attention mechanism along with a newly designed training method is based on the Seq2Seq architecture to overcome the difficulty in multistep prediction and capture the temporal heterogeneity of traffic pattern.
Long short-term memory neural network (Bi-LSTM NN) (26)	Automatic vehicle identification (AVI) detectors	A path-based deep learning framework which can produce better traffic speed prediction at a city-wide scale. Furthermore, the model is both rational and interpretable in the context of urban transportation.
Wasserstein Generative Adversarial Nets (WGAN) (27)	AutoNavi, a third-party navigation system	The generative neural network models the road link features of the adjacent intersections and the control parameters of intersections using a hybrid graph block. In addition, the spatial-temporal relations are captured by stacking a GCN, a recurrent neural network (RNN), and an attention mechanism.
Bidirectional long short-term memory network (BiLSTM) and gated recurrent unit (GRU) (28)	Video images from FHWA's Next Generation Simulation (NGSIM) Program	A hybrid prediction model K-BiLSTM-GRU is proposed, which combines the adaptive ability of K-means to reasonably classify samples and the advantage of BiLSTM and GRU to solve long-range dependencies and reduce overfitting.

GAPS IN PRACTICE AND KNOWLEDGE

Based on the review of existing practice and research, the following gaps are identified.

Best Practices in Travel Time, Speed, and Reliability Data Processing, Integration, and Applications

The collection and use of travel time, speed, and reliability data presents significant challenges for transportation agencies, primarily due to the vast amount of data available in various formats and resolutions; this includes the private-sector probe data. The storage, processing, and management of such data require robust IT infrastructure and skilled staff resources within these agencies. According to *NCHRP Synthesis 561*, some agencies still encounter difficulties in acquiring or using the data due to quality and accuracy concerns, limited capacity to handle large datasets, or privacy concerns. Considering these challenges, the experiences and lessons learned by certain agencies can be valuable to others facing similar challenges.

Investigating Best Approaches to Working with Trajectory Data by Transportation Applications

As the use of GPS-based trajectory data in transportation applications continues to grow, there is a pressing need for comprehensive research to address key challenges faced by transportation agencies. While trajectory data offers detailed trip-level insights, its vast data size requires substantial computing resources and specialized expertise for data management and processing, which many agencies lack. In addition, the availability of trajectory data varies across time and locations, which can impact the representativeness of the data and accuracy of analyses. Considering data accuracy and resolution requirements of different transportation applications, it is desirable to develop effective methodologies and guidelines for working with trajectory data.

Use of Probe Vehicle Data for Longitudinal Performance Assessment

Probe data quality has significantly improved over the past decade such that it is increasingly supporting agency operations and planning efforts (*7, 29*). Currently, several agencies procure and use GPS-probe speed data for assessing delay and reliability performance metrics over time, for conducting before-and-after studies to ascertain the impact of decisions and investments, for project prioritization, and many other use cases. Meanwhile, data vendors use a combination of methods, technologies, or sources to generate the probe data (*30*). Given the importance of probe data to traffic management, operations, and investment decision-making today, it is imperative to understand how changes to the methods, technologies and data sources used to generate the probe data affect its use for tracking performance over time. Thus, research is needed to assess the impact of such changes on the use of probe data for longitudinal assessment of performance.

Large-Scale Network Conflation

Third-party data sources procured from commercial vendors are typically tied to geospatial networks with proprietary segmentations schemes and referencing systems different from the networks maintained by the transportation agencies that procure and use such data (*7*). As such, matching links and direction between private-vendor networks and agency networks remains a challenge. Some advancements have been made to overcome this challenge (*10*, *31–33*). Nonetheless, more research is needed to develop fast and automated algorithms to help agencies.

Need for Efficient Travel Time and Speed Distribution Models by Roadway Types for Travel Time Reliability Measurement

Accurate estimation of the travel time or speed distribution is crucial for attaining accurate reliability assessments. However, the heterogeneity in the statistical characteristics of the distribution poses a challenge in their characterization, as evidenced by the multitude of functional forms proposed in the literature for this purpose (*34–36*). This heterogeneity may strongly depend on factors such as data collection methods, as well as the geographical and temporal characteristics of the travel time dataset. Different routes may present different characteristics. As such, more research is needed to provide guidelines on the appropriate selection of distributions for different applications and on different roadway types.

Need for Standardized Travel Time Reliability Measures and Network Monitoring Metrics

Various reliability evaluation measures have been developed to assess travel time reliability. These include buffer time, 90th/95th percentile travel time, travel time budget, misery index, and many others (*34*). Most such measures are based on the statistical properties of the travel time distribution, but they have been shown to not behave consistently for the same assessment object (*37–39*). There is therefore the need for research to establish some criteria for selection of travel time reliability metrics for different practical applications and to unify and better understand the relationships between the existing pool of metrics. Also, reliability metrics for network-wide travel time reliability monitoring seem to be lacking. The current practice in literature is to calculate reliability measures based on the total travel time; this total time is obtained by multiplying the flow of each individual link by its corresponding travel time, representing a straightforward aggregation from the link level to the network level (*34*). Such aggregation is insensitive to disturbances in the network, highlighting the necessity for further research in developing metrics for comprehensive network-wide monitoring.

Investigating Probe or Connected Vehicle Data Latency Issues

Latency issues with probe data have also been reported. Sharma et al. (*40*) found that the average latency of probe data was approximately 5 minutes, in contrast to fixed-location sensor data, with latencies from traditional sensors varying across corridors. This finding has significant implications for time-sensitive applications, such as traffic-responsive ramp metering or queue warnings. It is important to research the latencies of prevalent data collection technologies and their impacts on real-time transportation applications.

CURRENT AND PROPOSED NATIONAL RESEARCH INITIATIVES

This section describes existing and proposed research to addresses gaps in practice and knowledge for pavement engineering applications related to traffic monitoring.

Current Research

Ongoing and proposed projects related to travel time, speed, and reliability data are listed below.

NCHRP Project 08-157, "Best Practices for Data Fusion of Probe and Point Detector Data"

Research is underway to identify the challenges, issues, and proven or potential practices for performing data fusion to measure or forecast travel time, speed, reliability, and other aspects of operational performance on roadway networks. The objectives of this research are as follows:

1. Develop a process to:
   a. Identify specific objectives for data fusion;
   b. Identify data sources available for fusion;
   c. Select the most suitable data for fusion; and
   d. Facilitate the fusion itself.
2. Develop guidelines for transportation agencies to facilitate data fusion, improve data reporting, and ultimately improve traffic management.

NCHRP Project 08-143, "Guide to the Application of Spatial Segmentation on Travel Time Reliability Measures"

At the minimum, the research shall: (1) identify different segmentation and aggregation methods for measuring travel time reliability in use by state DOTs, MPOs, and probe vehicle data providers across the nation; and (2) evaluate the variability in travel time reliability measures using different segmentation or aggregation methods appropriate for application in different roadway networks and contexts, taking into consideration roadway type, local density and environment, and temporal and geographic contexts, as appropriate.

NCHRP Project 08-119, "Data Integration, Sharing, and Management for Transportation Planning and Traffic Operations"

The objective of this research is to develop tools, methods, and guidance for improving data integration, sharing, and management practices to enable transportation agencies, in collaboration with private-sector and public-sector stakeholders, to make better planning and operations decisions. Secondary benefits will be increased uniformity of data across states and improved consistency of practice.

NCHRP Project 03-151, "Data Subsystems and Data Management Plans for Traffic Management Systems"

Agencies face challenges with systematically managing data as part of their operating a traffic management system (TMS). There are also limited resources for agencies to use to assist with managing data (e.g., archiving, use, configuration, and monitoring use), issues to consider with receiving and sharing or using data with third-party sources or within an agency (e.g., licenses, proprietary, and sensitive information). The objectives of this research are to review and compile information from available resources, offer insights, and synthesize current practices.

NCHRP Project 08-145, "Utilizing Cooperative Automated Transportation (CAT) Data to Enhance Freeway Operational Strategies"

The objective of this research is to assess operational scenarios and use cases where freeway operations strategies could be improved through the transmission of data between a TMS and the larger CAT system. This assessment should (1) spur development of enhanced and new operational strategies and (2) help agencies justify gaining access to additional CAT data.

NCHRP Synthesis 20-05/Topic 55-02, "Practices for Collecting, Managing, and Using Light Detection and Ranging (LiDAR) Data"

While collection and use of LiDAR data have become widespread, state DOTs often have questions on ways to improve their processes, especially as advances in data governance practices, analysis methods, tools, and technologies expand the potential benefits and challenges of using LiDAR data. The objective of this synthesis is to document state DOT practices related to collecting, managing, and using LiDAR data.

Assessment of Travel-Time Reliability and Operational Resilience of Metro Freeway Corridors https://rip.trb.org/view/2006249

This project will use and further enhance a software tool known as TeTRES, Travel-Time Reliability Estimation System. The TeTRES was developed to measure and analyze the travel-time reliability trends of the metro freeway corridors.

NCHRP Project 08-135, "Reliability and Quality of Service Evaluation Methods for Rural Highways"

Given that the *Highway Capacity Manual* (HCM) is accepted nationwide as the primary source on highway capacity and quality of service, the lack of a technical approach to address domestic rural highways is a major limitation. Another limitation of the current HCM methodology for rural highways is the analysis horizon, which is limited to a single study period. Use of a distribution of level of service values mimics the variability of traffic conditions on the facility and provides a better understanding of the quality of service across time. By having more appropriate performance measures for these types of facilities, state DOTs can better allocate their scarce resources.

NCHRP Synthesis 20-05/Topic 53-14, "Use of Probe Data for Freight Planning and Operations"

As the penetration of probe data has increased in the last decade, its use in transportation planning and operations at state DOTs has become common. However, the application of probe data in freight-related applications is vastly different from traditional transportation planning use. For instance, freight application tends to be specific to freight generators, freight attractors, and freight bottlenecks such as those near ports and borders; in these situations, the probe data may not be sufficient by itself. The objective of this project is to document current state DOT practices regarding the use of probe data in freight planning and modeling, and operations management applications.

Proposed Research

The following research needs are identified following the comprehensive literature review. The motivation for them is found in the Gaps in Practice and Knowledge section.

Best Practices in Travel Time, Speed, and Reliability Data Processing, Integration, and Applications

The collection and use of Travel Time, Speed, and Reliability (TTSR) data poses significant challenges for transportation agencies, mainly due to the sheer volume of data in diverse formats and resolutions. Managing such data demands strong IT infrastructure and skilled personnel. Additionally, some agencies struggle with data acquisition and use, often hindered by concerns over data quality and accuracy, limitations in processing large datasets, or privacy issues. This research would compile experiences and lessons learned by agencies that have successfully navigated these challenges can be highly beneficial to others facing similar situations.

Investigating Best Approaches to Working with Trajectory Data by Transportation Applications

The proposed research idea focuses on investigating the best approaches for using trajectory data in transportation applications. This data, while offering detailed insights at the trip level, poses challenges due to its large size and the substantial computing resources and specialized expertise needed for effective management and processing. The variability in the availability of trajectory data across different times and locations also affects their representativeness and the accuracy of analyses. Therefore, there is a need to develop effective methodologies and guidelines tailored to different transportation applications, considering the specific accuracy and resolution requirements of trajectory data.

Use of Probe Vehicle Data for Longitudinal Performance Assessment

The research idea revolves around the use of probe vehicle data, specifically GPS-probe speed data, for longitudinal performance assessment in traffic management. Many agencies rely on this data to evaluate delay and reliability metrics over time, to conduct before-and-after studies to determine the impact of decisions and investments, and for various other applications like project prioritization. Since data vendors employ a mix of methods, technologies, and sources to generate this probe data, it is crucial to understand how changes in these aspects affect its utility for tracking performance over time. Therefore, there is a need for research to assess the impact of such changes on the effectiveness of probe data in the longitudinal assessment of traffic performance.

Large-Scale Network Conflation

The proposed research idea focuses on large-scale network conflation. It addresses the challenge that third-party data are often tied to proprietary geospatial networks, which have different segmentation schemes and referencing systems than those maintained by transportation agencies. Matching links and directions between these third-party data networks and agency networks is a significant challenge. Therefore, there is a need for research to develop fast and automated algorithms to help agencies efficiently conflate these different network data sources, enabling more effective use of third-party geospatial data in transportation planning and analysis.

Need for Efficient Travel Time/Speed Distribution Models by Roadway Types for Travel Time Reliability Measurement

The research idea focuses on developing efficient travel time/speed distribution models tailored to different types of roadways for the purpose of travel time reliability measurement. Accurate estimation of travel time or speed distribution is vital for reliable assessments and is crucial in performance measurement, project programming, and benefit-cost analysis. The characterization of these distributions is influenced by various factors, including data collection methods and the geographical and temporal characteristics of the travel time dataset. Additionally, different road types, such as interstates and urban arterials, exhibit distinct characteristics. Hence, there is a need for more research to establish guidelines on selecting appropriate distribution models for different applications across various roadway types, ensuring more accurate and relevant travel time reliability assessments.

Need for Standardized Travel Time Reliability Measures and Network Monitoring Metrics

The research idea emphasizes the need for standardized travel time reliability measures and network monitoring metrics. Despite the development of various measures to assess travel time reliability, most are based on the statistical properties of travel time distribution and have shown inconsistent behavior for the same assessment object. This inconsistency highlights the necessity for research aimed at establishing criteria for selecting travel time reliability metrics suitable for different practical applications. Additionally, there is a notable gap in reliability metrics for network-wide travel time reliability monitoring, indicating a need for further research in this area to develop comprehensive and consistent measurement tools.

Investigating Probe or Connected Vehicle Data Latency Issues

The research idea focuses on investigating latency issues in probe or connected vehicle data. It has been reported that probe data often experiences latency, with one study revealing an average delay of about 5 minutes. This is a contrast to the quicker response of fixed-location sensor data, and these latencies vary across different corridors. This latency is particularly consequential for time-sensitive traffic applications like traffic-responsive ramp metering or queue warnings. Thus, there is a significant need to research the latency characteristics of current data collection technologies and understand their impact on real-time transportation applications, aiming to optimize these systems for more immediate and effective responses.

REFERENCES AND OTHER RESOURCES

1. List, G. F., B. Williams, N. Rouphail, R. Hranac, T. Barkley, E. Mai, A. Ciccarelli, L. Rodegerdts, K. Pincus, B. Nevers, A. F. Karr, X. Zhou, J. Wojtowicz, J. Schofer, and A. Khattak. *SHRP 2 Report S2-L02-RR-1: Establishing Monitoring Programs for Travel Time Reliability*. Transportation Research Board of the National Academies, Washington, DC, 2014.

2. Cambridge Systematics, Inc., Texas A&M Transportation Institute, University of Washington, Dowling Associates, Street Smarts, H. Levinson, and H. Rakha. *SHRP 2 Report S2-L03-RR-1: Analytical Procedures for Determining the Impacts of Reliability Mitigation Strategies*. Transportation Research Board of the National Academies, Washington, DC, 2012.
3. Bauer, J. K., Margiotta, R., and Pack, M. *Applying Archived Operations Data in Transportation Planning – A Primer* (Internet). 2016. Report No.: FHWA-HOP-16-082.
4. Antoniou C., Balakrishna R., and Koutsopoulos, H.N. A Synthesis of Emerging Data Collection Technologies and Their Impact on Traffic Management Applications. *Eur. Transp. Res. Rev*. 2011 Nov; 3(3):139–48.
5. *Transportation Research Circular E-C227: Advancing Highway Traffic Monitoring Through Strategic Research*, 2017, Transportation Research Board, Washington, DC.
6. Zhang, X., C. Van Dyke, G. Erhardt, and M. Chen. *NCHRP Synthesis 541: Practices on Acquiring Proprietary Data for Transportation Applications*. Transportation Research Board, Washington, DC, 2019.
7. Pack, M. L., and N. Ivanov. *NCHRP Synthesis 561: Use of Vehicle Probe and Cellular GPS Data by State Departments of Transportation*. Transportation Research Board, Washington, DC, 2021. Available from: https://www.nap.edu/catalog/26094.
8. Chen, M., Zhang, X., Rahman, F., Brashear, J., and Souleyrette, R. *Measuring Congestion for Strategic Highway Investment for Tomorrow (SHIFT) Implementation (PL-32)* (Internet). Kentucky Transportation Center Research Report, 2019. Report No. 1682.
9. Congestion Management Program 2021 (Internet). San Francisco County Transportation Authority, 2021. Available from: https://www.sfcta.org/sites/default/files/2022-05/Congestion_Management_Program_Report_220517_FINAL.pdf.
10. Kaushik, K., E. Wood, and J. Gonder. Coupled Approximation of U.S. Driving Speed and Volume Statistics Using Spatial Conflation and Temporal Disaggregation. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2672, pp. 1–11, 2018.
11. Mathew J. K., Li H., Desai J., Suryakant Sakhare R., Saldivar-Carranza E., Hunter M., et al. Integration of Probe Data Tools into TMC Operations (Internet). Purdue University; 2022 (cited 2023 Sep 6). Available from: https://docs.lib.purdue.edu/jtrp/1781/.
12. Cheu, L. R., Xie, C., and Lee, D. H. Probe Vehicle Population and Sample Size for Arterial Speed Estimation. *Comp-aided Civil Eng*. 2002 Jan; 17(1):53–60.
13. Abdelraouf, A., Abdel-Aty, M., and Mahmoud, N. Sequence-to-Sequence Recurrent Graph Convolutional Networks for Traffic Estimation and Prediction Using Connected Probe Vehicle Data. *IEEE Transactions on Intelligent Transportation Systems*. 2022; 1–11.
14. Li, Y., and M. McDonald. Link Travel Time Estimation Using Single GPS Equipped Probe Vehicle. In: *Proceedings The IEEE 5th International Conference on Intelligent Transportation Systems*. 2002. p. 932–7.
15. Jenelius, E. and H. N. Koutsopoulos. Travel Time Estimation for Urban Road Networks Using Low Frequency Probe Vehicle Data. *Transportation Research Part B: Methodological*. 2013 Jul 1; 53:64–81.
16. Zheng, F., and H. Van Zuylen. Urban link Travel Time Estimation Based on Sparse Probe Vehicle Data. *Transportation Research Part C: Emerging Technologies*. 2013 Jun 1; 31:145–57.
17. Ou, Q., Bertini, R. L., van Lint, J. W. C., and Hoogendoorn, S. P. A Theoretical Framework for Traffic Speed Estimation by Fusing Low-Resolution Probe Vehicle Data. *IEEE Transactions on Intelligent Transportation Systems*. 2011 Sep; 12(3):747–56.
18. Petrovskaya, A., and Thrun, S. Model Based Vehicle Detection and Tracking for Autonomous Urban Driving. *Auton Robot.* 2009 Apr 1; 26(2):123–39.

19. Zhang, Z., J. Zheng, H. Xu, and X. Wang. Vehicle Detection and Tracking in Complex Traffic Circumstances with Roadside LiDAR. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2673, 2019, pp. 62–71.
20. Sun Y., H. Xu, J. Wu, J. Zheng, and K. M. Dietrich. 3-D Data Processing to Extract Vehicle Trajectories from Roadside LiDAR Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2672, 2018, pp. 14–22.
21. Wejo. Real-Time Traffic Intelligence (Internet). Available from: https://www.wejo.com/products/real-time-traffic-intelligence.
22. Otonomo. Traffic Management Data - Use Cases (Internet). Available from: https://otonomo.io/use-cases/traffic-management-data/.
23. Wejo. Connecting the Dots: Connected Vehicle Data and Work Zone Safety (Internet). Available from: https://www.wejo.com/resources/connecting-the-dots-connected-vehicle-data-work-zone-safety.
24. Desai J., Mathew J. K., Li H., Sakhare R., Horton D., and Bullock D. M. National Mobility Analysis for All Interstate Routes in the United States (Internet). Purdue University; 2023 (cited 2023 Jun 1). Available from: https://docs.lib.purdue.edu/imr/8/.
25. Zhang Z., Li M., Lin X., Wang Y., and He F. Multistep Speed Prediction on Traffic Networks: A Deep Learning Approach Considering Spatio-Temporal Dependencies. *Transportation Research Part C: Emerging Technologies*. 2019 Aug; 105:297–322.
26. Wang J., Chen R., and He Z. Traffic Speed Prediction for Urban Transportation Network: A Path Based Deep Learning Approach. *Transportation Research Part C: Emerging Technologies*. 2019 Mar; 100:372–85.
27. Jin J., Rong D., Zhang T., Ji Q., Guo H., Lv Y., et al. A GAN-Based Short-Term Link Traffic Prediction Approach for Urban Road Networks Under a Parallel Learning Framework. *Trans Intell Transport Sys*. 2022 Sep; 23(9):16185–96.
28. Li Q., Cheng R., and Ge H. Short-term Vehicle Speed Prediction Based on BiLSTM-GRU Model Considering Driver Heterogeneity. *Physica A: Statistical Mechanics and its Applications*. 2023 Jan; 610:128410.
29. Pack M. L., Ivanov N., Bauer J. K., Birriel E. Considerations of Current and Emerging Transportation Management Center Data (Internet). 2019. Report No.: FHWA-HOP-18-084. Available from: https://ops.fhwa.dot.gov/publications/fhwahop18084/index.htm.
30. Center for Advanced Transportation Technology, University of Maryland, KMJ Consulting, Inc. Vehicle Probe Project: Data Use and Application Guide. I-95 Corridor Coalition. (Internet). 2011. Available from: https://tetcoalition.org/wp-content/uploads/2015/03/008-7G_VPP_DATA_USE_Report_Final_April_2011.pdf.
31. Daneshgar F., K. F. Sadabadi, and A. Haghani. A Conflation Methodology for Two GIS Roadway Networks and Its Application in Performance Measurements. T*ransportation Research Record: Journal of the Transportation Research Board*, No. 2672, pp. 284–293, 2018.
32. Pandit D. M., K. Kaushik, and C. Cirillo. Coupling National Performance Management Research Data Set and the Highway Performance Monitoring System Datasets on a Geospatial Level. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2673, pp. 583–592, 2019.
33. Zhang X., and M. Chen. Methodology for Conflating Large-Scale Roadway Networks. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2677, pp.189–202, 2023.
34. Zang Z., Xu X., Qu K., Chen R., and Chen A. Travel Time Reliability in Transportation Networks: A Review of Methodological Developments. T*ransportation Research Part C: Emerging Technologies*. 2022 Oct; 143:103866.
35. Guo F., H. Rakha, and S. Park. Multistate Model for Travel Time Reliability. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2188, pp. 46–54, 2010.

36. Tufuor E. Development of an Improved Arterial Roadway Performance Reliability Methodology (Internet) (Doctor of Philosophy). (Lincoln): University of Nebraska; 2020. Available from: https://digitalcommons.unl.edu/civilengdiss/171/.
37. Pu W. Analytic Relationships Between Travel Time Reliability Measures. *Transportation Research Record: Journal of the Transportation Research Board*. No. 2254, pp. 122–130, 2011.
38. Van Loon R., Rietveld P., and Brons M. Travel-time reliability impacts on railway passenger demand: a revealed preference analysis. J*ournal of Transport Geography*. 2011 Jul; 19(4):917–25.
39. Wakabayashi H., and Matsumoto Y. Comparative study on Travel Time Reliability Indexes for Highway Users and Operators: Comparative Study on Travel Time Reliability Indexes. *J Adv Transp*. 2012 Oct; 46(4):318–39.
40. Sharma A., Ahsani V., and Rawat S. Evaluation of Opportunities and Challenges of Using INRIX Data for Real-Time Performance Monitoring and Historical Trend Assessment (Internet). Nebraska Department of Transportation; 2017.
41. Waddell J., Remias, S., and Kirsch J. Characterizing Traffic-Signal Performance and Corridor Reliability Using Crowd-Sourced Probe Vehicle Trajectories. *Journal of Transportation Engineering, Part A: Systems*. 2020 Jul 1; 146(7):04020053.

Bicycle and Pedestrian Traffic Monitoring

SIRISHA KOTHURI
Portland State University

FRANK PROULX
Frank Proulx Consulting

GREG LINDSEY
University of Minnesota

ELIZABETH STOLZ
Marlin Engineering

KRISTA NORDBACK
University of North Carolina at Chapel Hill

TONY HULL
Independent Consultant

JULIA GRISWOLD
University of California, Berkeley

PETER OHLMS
Virginia Department of Transportation

JOSH ROLL
Oregon Department of Transportation

PATRICK SINGLETON
Utah State University

INTRODUCTION

Bicycling and walking are fundamental transportation modes. Many benefits are associated with them, including lower congestion and emission levels and improvements in personal health. Many state, regional, and municipal agencies have established counting programs to monitor bicycle and pedestrian traffic with the objective of characterizing flows on major streets in networks and intersections that integrate with vehicular monitoring networks. As with vehicular traffic, bicycle and pedestrian count data provide the foundation for decision-making and are useful for measuring trends, developing performance indicators, and planning and designing new infrastructure. They also are used as the exposure metric in safety, equity, and health analyses.

Bicycle and pedestrian count data are used in the same ways as motorized vehicular count data and similar principles and protocols are followed in monitoring. However, specialized guidance for bicycle and pedestrian traffic monitoring is required because of inherent differences in travel patterns, particularly the ability of bicyclists and pedestrians to vary directions and cross roadways regardless of roadway elements designed to direct flows. Thus, this chapter includes sections that mirror preceding chapters in this E-Circular but focus on details characteristic of bicycle and pedestrian traffic monitoring. Specifically, this chapter elaborates on general traffic monitoring principles introduced in previous sections (Continuous Counts, Short-Duration Counts, Managing Databases, Performance Measures, Data Quality, and Data Integration). This chapter does not include sections related to chapters that focus on topics in monitoring related primarily to vehicular traffic (e.g., pavement engineering applications).

In general, methods to monitor bicycle and pedestrian traffic have lagged motorized monitoring methods, but there have been tremendous advances over the past two decades along several avenues. Research has led to advances in monitoring equipment for continuous and short-duration counts; guidance on the design of monitoring programs, including optimal length of short-duration counts; new factoring methods for expanding short-duration counts to annual average daily traffic estimates; development of archives or warehouses to store bicycle and pedestrian counts; and data quality checks to improve accuracy of counts. The recent emergence of third-party data vendors has allowed exploration of data fusion techniques that can integrate traditional count data with new types of counts and with modeled estimates of bicycle and pedestrian flows. The past decade has also seen the emergence of micromobility devices such as shared (e-) scooters, which pose new challenges and opportunities for data capture and classification.

In 2004, reflecting the growing popularity of bicycling and walking and the need for better data, the Institute of Transportation Engineers and Alta Planning + Design launched the National Bicycle and Pedestrian Documentation Project. This was an initiative that established procedures for 2-hour peak hour counts and was often undertaken by volunteers working with staff (*1*). The number of count programs grew, along with recognition that continuous count programs analogous to vehicular monitoring programs were needed to provide more valid data. Since then, a number of guidance documents have been developed to assist agencies with establishing active travel count programs. FHWA's *Traffic Monitoring Guide* (TMG) was first published in 1985; however, that edition did not cover nonmotorized count data (*2*). The 2013 update to the TMG was the first edition to provide guidance on collecting nonmotorized count data (*3*), while the 2016 TMG provided updates to the format (*4*). Other guidance documents include the Guidebook for Developing Pedestrian and Bicycle Performance Measures (*5*); Coding Nonmotorized Station Location Information in the 2016 Traffic Monitoring Guide Format (*6*); Exploring Pedestrian Counting Procedures (*7*); and the Bicycle-Pedestrian Count Technology Pilot Project (*8*). *NCHRP Report 797: Guidebook on Pedestrian and Bicycle Volume Data Collection* and its two associated web-only documents provide guidance on all facets of count programs (*9, 10,11*).

The 2022 TMG broadened the focus from bicycle and pedestrian traffic to include all types of micromobility modes. The document also integrates guidance on micromobility monitoring with parallel guidance on vehicular monitoring (*12*). The growth of automated monitoring has led to reduced emphasis on manual monitoring, with a few cities and states (e.g., City of Minneapolis) discontinuing their manual count programs.

The objective of this chapter is to document recent advancements in bicycle and pedestrian traffic monitoring, to characterize state of the practice and state of the art, and to highlight knowledge gaps where research is required. As noted, this chapter also addresses data integration with estimated traffic volumes modeled using location-based data by third-party vendors. While volume, behavioral, network, and sociodemographic data are all necessary for understanding active travel modes, this chapter focuses on traffic monitoring and the count data produced by monitoring. Mode share as a performance measure is not addressed because it cannot be accurately computed from count data. The remainder of this chapter reviews continuous counts, short-duration counts, data management, data quality and equipment calibration, performance measures, and data integration. The summary section contains a list of important research needs that will be useful to researchers and practitioners. This list reflects the subjective assessment of the authors and by no means is exhaustive of the many priorities to be addressed if bicycle and pedestrian traffic monitoring is to realize its potential.

CONTINUOUS COUNTS

As noted in the first chapter, continuous count programs are the foundation of traffic monitoring programs and generate the counts of traffic needed to characterize traffic volumes on transportation networks. Counting bicycling and walking using continuous counting equipment is common in North America; it is made possible by new technological improvements in sensors that can achieve over 90% accuracy for bicycles and over 80% accuracy for pedestrians (*9,10,11*). However, data standards and quality are variable, partly because these data are usually not required by federal or state agencies, and funding for equipment maintenance and validation and data management is often insufficient. As a result, bicycle and pedestrian traffic monitoring has yet to realize its potential.

State of the Practice

Continuous bicycle counts have been conducted since at least the 1980s using inductive loops cut into the pavement of paths and roadways. Continuous pedestrian counts have been conducted using infrared sensors in North America since the 2000s if not before. Other technologies are also used for continuous counts as documented in *NCHRP Report 797: Guidebook on Pedestrian and Bicycle Volume Data Collection* (*9, 10,11*) and the *Traffic Monitoring Guide* (*3,4,12*). These technologies include pneumatic tubes, piezoelectric strips, radar, and active infrared counting devices. Technologies for counting bicycles are sometimes different than those for counting pedestrians, and accurate methods for differentiating micromobility devices are still being developed. Small diameter pneumatic tubes (minitubes) are usually used for short-duration counts. However, they have been used successfully and accurately for many years as bicyclist permanent continuous counters in areas which are not snow plowed and where staff are available to replace tubes multiple times per year. More recently, some agencies have deployed video camera counting with either manual or automated post-processing to produce counts. These approaches remain costly and generally have not produced counts to characterize flows across entire networks.

Most continuous count sites have been installed either on paths or in bicycle-and-pedestrian-only facilities because most current technologies perform better at these locations. However, with improvements in loop technology and pneumatic tubes that better distinguish among vehicles and bicycles, deployment in roadways is increasing. Continuous count programs for pedestrians seem to have lagged bicycle count programs because of the complexity of pedestrian flows on sidewalks and the technological challenges to capturing different movements.

Most of the current data collection efforts are driven by specific data needs (safety analysis, project justification, before/after implementation evaluation), which limits the scope and effectiveness of emerging count programs. Many agencies commonly use these data for specific and limited applications such as validating facility use and generating support for future funding requests. Occasionally, data are also used to prove the need for maintenance and winter snow clearing. Comparatively few state and local government transportation or public works agencies have comprehensive monitoring networks that enable them to characterize bicycle or pedestrian volumes and flows on along entire networks.

State of the Art

New technologies are constantly being developed and traditional technologies continue to improve. Among the most recent developments are video cameras using AI and machine learning to detect, classify, and track the movement and speed of motorized and nonmotorized road users, including at intersections. However, the accuracy and reliability of this approach has yet to be sufficiently documented, and costs remain high.

As experts seek better data, researchers are exploring how other types of counts or measures can be used to estimate or model data. For example, push button calls by pedestrians at signalized intersections can be used for studying pedestrian activity (*13, 14*). Although these are not pedestrian volumes specifically, studies documenting the relationship of observed pedestrian crossing volumes to pedestrian pushbutton data show that, despite obvious limitations (e.g., only capturing pedestrians who actually use the call button and at what frequency this occurs), this readily available source of data is correlated with pedestrian volumes. More recent research has used pushbutton actuations to estimate pedestrian crossing volumes in Utah (*15*).

Emerging Trends

Integrating continuous count data with third-party data such as data from smartphone apps (e.g., Strava, CycleTracks), bikeshare, and other sources is an emerging method for estimating volumes across a network and has been employed for safety studies (*16*). However, the reliability of this approach is still an area of research, as the penetration and availability of app data varies by location and year. In addition, the continuous count data currently available for calibrating third-party data are insufficient; reliable count locations are too few and not in a sufficient number of representative locations. Specifically, continuous counters at low-volume and rural locations are often very few or non-existent.

Other approaches for measuring active transportation have been or are being explored. These include WIFI or Bluetooth detection; low-cost indoor-style infrared counting devices; observations from probe vehicles (motorized or nonmotorized); bicycle parking surveys; and observations of Google Earth, satellite, or other static continuous photo data collections. Many of these data sources, which often are provided by private vendors using proprietary protocols, provide indices of volumes but not actual counts and do not consistently report measures of validation.

Micromobility devices have created a need to better classify types of users in counts. While micromobility devices are new, growing, and evolving in North America, locations with e-scooter sharing programs increase their use considerably by providing access to the e-scooters instead of relying on road users to own a personal e-scooter. Agencies increasingly want to count these users separately from pedestrians and cyclists to better study their safety, and current technologies are still developing to meet this need. Electric bicycles (e-bikes) are becoming more common but are difficult to differentiate from standard bicycles either with the human eye or via automated counting technology. Nonetheless there is interest in tracking their speed and behavior, in part to assess safety on shared use paths and other infrastructure not originally designed to serve bicycles that can travel at higher speeds.

Knowledge Gaps

Many knowledge gaps affect current practice and deployment of state-of-the-art practices. These gaps include understanding of the validity of counts and modeled estimates provided by emerging technologies and the validity of many different design parameters for comprehensive, network-wide monitoring programs.

The validity and accuracy of counts provided by emerging technologies need to be evaluated. This is an ongoing research need; technologies are constantly developing, and unverified manufacturer claims may lead jurisdictions to invest in unproven technologies. Automated counter data validation also is needed. Suitable data checks differ from those used for motor vehicles. Due to the variability of cycling and walking in North America, the data checks also vary from jurisdiction to jurisdiction. In most areas in North America, low-volume sites (<100 people walking and biking per day) are common, and data from these sites are the most challenging to check for potential problems.

With respect to the design of monitoring programs, knowledge gaps include:

- How to design a statistically based continuous count program to minimize error in estimates of annual average daily nonmotorized traffic (AADNT);
- How to select continuous count sites in conjunction with short-duration count sites to optimize the cost and accuracy of estimates of AADNT;
- How to best integrate continuous count data with emerging data sources, including statistically robust methods for placing continuous count sites to properly inform the emerging data; and
- How to document case studies of the use of continuous count data, including typologies and priorities for uses and needs.

SHORT-DURATION COUNTS

As noted in the second chapter, short-duration count programs are a critical component of traffic monitoring programs. However, states and most local jurisdictions lack resources and capacity to implement state-of-the-art practices required to provide network-level estimates of bicycle and pedestrian volumes on roadway segments and at intersections. Emerging trends include exploration of third-party data sources to augment limited count programs through data fusion procedures. For the most part, knowledge gaps identified in early versions of the *TMG* persist, and research is needed to address them (*3,4,12)*. These research needs include topics related to each of the key steps and issues in short-duration programs:

1. Selection of count locations,
2. Integration of segment and intersection count programs,
3. The duration of short-duration counts,
4. Quality assurance/quality control,
5. Identification of traffic patterns and factor groups, and
6. Procedures for factoring short-duration counts in estimates of average annual daily bicyclists (AADB) or pedestrians (AADP).

Data fusion initiatives may provide a framework for the design and integration of short-duration and permanent count data along with other data, but they are not a panacea. Long-term investments in count programs are essential to ensuring high data quality.

State of the Practice

States and most local jurisdictions are unable to follow the approach used in vehicular traffic monitoring and have not been able to implement multiyear sampling programs to generate estimates of AADB or AADP. Instead, most jurisdictions have purposeful or ad hoc short-duration count programs designed to inform ongoing management activities such as the timing of traffic signals or the need for safety countermeasures. Over time, these purposeful, need-based short-duration count programs can provide insight into bicycle and pedestrian traffic patterns and volumes within networks, but they are insufficient to fully characterize network flows. With respect to key steps in comprehensive monitoring programs:

- Most count locations have been chosen purposefully for need-specific reasons rather than as strategic part of a comprehensive monitoring program that includes high, medium, and low volumes sites. The locations often reflect other characteristics that may influence factors used to extrapolate short-duration counts.
- Counts may include both manual sample counts and automated counts for durations ranging from as little as 2 hours (e.g., peak hour counts) to 2 weeks or more.
- Counts typically are undertaken during months when conditions are good for cycling or walking (e.g., year-round in jurisdictions with mild weather; May through October in jurisdictions with cooler climates).
- QA/QC may not be performed or may be limited to visual screening.
- Both conventional and day-of-year factoring approaches may be used to estimate AADB or AADP from the short-duration counts.
- Estimates of AADB and AADP extrapolated from counts may be available online in reports or spreadsheets, but online and interactive maps with historical monitoring results may not be available.

State of the Art

A state-of-the-art bicycle and pedestrian traffic short-duration count program should follow procedures analogous to those used in vehicular traffic monitoring, with modifications made to address distinctive characteristics of bicycle and pedestrian traffic (e.g., greater variation in demand due to weather or higher traffic on weekend than weekdays in some locations). In a state-of-the-art program:

- Count locations would be determined in a strategic monitoring plan designed to produce key performance measures (e.g., AADB, AADP) for relevant segments and intersections of a network within a specified, multiyear time frame.
- Counts would be collected through automated sensors and taken consistently for periods of 10 days to 2 weeks to ensure adequate inclusion of day-of-week variation.
- Short-duration counts would be taken during months of the year when extrapolations using preferred factoring methods produce the lowest maximum absolute percentage error.
- Statistical QA/QC checks would be automated and applied to all counts.
- Factoring procedures would be programmed and automated to produce key performance metrics specified in the monitoring plan.
- Results would be integrated into online, interactive maps on agency websites that enable users to download both key performance metrics and historical data, either for individual sites or for the entire network (*3, 4, 5*).

Key performance metrics would be available for bicyclists, pedestrians, and other micromobility device users for the entire network. While some jurisdictions have achieved some of these program elements, few if any can provide all relevant metrics for all elements of a network, primarily because of resource and capacity constraints.

Emerging Trends

Growing numbers of state and local transportation, public works, and recreation agencies have implemented monitoring programs that involve estimation of key performance metrics from short-duration counts using patterns, factors, and other information from permanent counters. The capital, labor, maintenance, analytic, and communication costs associated with automated count programs have limited the implementation of comprehensive state-of-the-art programs. The costs are also a primary factor fueling interest in estimates derived from third-party data sources of bicycle and pedestrian traffic volumes. Some agencies are using available crowdsourced data or data purchased from third-party data firms who use location-based service (LBS) data. However, the need for field counts remains as an essential component of local transportation planning efforts to address limitations of third-party data sources and to validate estimates produced from these third-party data firms (16). Some public agencies that historically conducted 2-hour, peak hour short-duration counts have discontinued their programs, focusing instead on automated short-duration counts or data from third-party sources.

Knowledge Gaps

Local and state “data wranglers”—professionals who have responsibility for the complex tasks of implementing and managing bicycle and pedestrian traffic monitoring programs—have gained considerable insight into the technological, analytic, programmatic, and administrative requirements for effective short-duration count programs over the past decade. However, gaps in knowledge remain, especially with respect to strategies for maximizing the validity of key performance metrics produced from short-duration counts. These programs are underfunded, and local agencies are facing growing financial constraints. Gaps in knowledge can be identified in each of key elements of a comprehensive program. While specific elements may vary by micromobility modes (e.g., bicycles vs. pedestrians vs. scooters) these gaps include:

- How to segment networks and optimally place short-duration sensors to maximize validity of measures while minimizing costs of field work (e.g., *17*).
- Tradeoffs in validity of estimates of key metrics, given short-duration counts of different lengths for different modes and factor groups at different times of the year (e.g., how can fixed numbers of counters be deployed to optimize estimates of flows on networks under different assumptions about segment length, choices of monitoring technologies, and costs of labor for deployment?) (e.g., *18*).
- “Minimum acceptable” QA/QC procedures for hourly and daily short-duration counts, and procedures for integrating and automating QA/QC. (e.g., *19*).
- The level of error in key metrics that is acceptable for decision-making in different contexts.
- Efficient methods for distribution of count data and key metrics to improve transportation decision-making.

DATA MANAGEMENT

Traffic monitoring programs have long stressed the importance of effective data management in producing valid estimates of traffic from continuous and short-duration counts. Data management practices vary widely across the federal government, the states, and local jurisdictions. Some states purchase data software programs and others develop their own. Data management programs originally developed for vehicular traffic data now are being redesigned or augmented to interface or integrate with bicycle and pedestrian traffic data. At the local level, many data management programs have been built from the ground-up using combinations of spreadsheets, GIS software, and other programs. Data publication and distribution programs often are ad hoc, constrained by limited resources and competing demands on staff.

State of the Practice

Local, regional, and state transportation agencies collect both continuous and short-duration active transportation count data. As noted above, most continuous count sites are not selected to optimize the validity of estimates of counts across networks. Most short-duration count studies are conducted ad hoc for periods from 2 hours up to 1 week before and/or after infrastructure projects, although some agencies conduct annual counts at a selection of sites.

These data are generally stored in spreadsheet or PDF format and may include diagrams to describe the different pedestrian or bicyclist movements collected, particularly for intersection count studies. Agencies generally do not have a central database for storing these data, so a count study may languish in a project folder on an agency server. Even within a given agency, departments may use different technologies to collect similar types of data, and one department may not be aware of the data collected by another department. Although some states and local agencies have developed online data dashboards to share bicycle and pedestrian traffic data, and others post data in spreadsheets online, most agencies do not have well-developed and documented programs for sharing bicycle and pedestrian traffic data and performance measures.

The *Traffic Monitoring Guide* produced by the FHWA provides guidance for producing counts that can be uploaded and archived in the Traveling Monitoring and Analysis System (TMAS) (*12*). The FHWA has also produced guidance of use of TMAS and has encouraged state and local agencies to upload data. A challenge inhibiting the use of TMAS as a repository and for archival purposes has been that many state and local agencies lack the capacity to reformat data to meet TMAS requirements documented in the TMG.

In contrast, most vendors of automated counters manage online portals for storing, visualizing, and downloading their data. Only some of these systems, however, provide easy access to bulk download data, and that option is generally only available to the agency that owns the counter equipment or that ordered the count study. These practices make it more difficult for researchers or other interested parties to access the data.

State of the Art

A state-of-the-art data management program would support the sharing of count data through an internal or public database so the data can be more fully used. Examples of state-of-the-art systems in the United States include several regional databases and dashboards for bicyclist and/or pedestrian count data managed by agencies in the Philadelphia, PA; Arlington, VA; and Eugene, OR areas that provide map interfaces and basic visualization (*20, 21, 22*). The Bike Ped Portal is a national database developed at Portland State University that contains both short-duration and continuous segment or screenline counts from throughout the United States (*23*). The Texas Bicycle and Pedestrian Count Exchange (PB|CX) includes a map view for navigation, data visualization for each count station, and an admin portal that allows data owners to quality review their data before making it public (*24*). The Florida DOT has also implemented an online dashboard (*25*).

Emerging Trends

As investments in active transportation infrastructure increase throughout the United States, more agencies are conducting active transportation count studies or installing continuous counters to measure project performance and justify further funding. There is increasing attention to the need to calibrate and validate continuous counts so that only good quality data are made available for use. The California DOT, for example, is currently developing a statewide active transportation database that will incorporate calibration and validation procedures as well as expansion methods for developing annual estimates or AADNT estimates from short-term counts. Researchers are also looking at how data fusion techniques can incorporate third-party data products with continuous counts to develop network-wide volume estimates.

Knowledge Gaps

Data management involves a complex set of activities, from data collection and cleaning to quality assurance to publishing and disseminating. The state of practice currently lags the state of the art due to knowledge gaps as well as resource and capacity constraints. The knowledge gaps include effective practices for:

- Integration of data from new technologies such as video analytics continues to raise questions for travel monitoring. These technologies allow agencies to gather counts from existing equipment, such as traffic cameras at intersections that have traditionally been used for detection but not counting. Data from these sources may require different relational database designs to capture the areas in which the cameras collect user activities. Additionally, implementation of database systems that are user-friendly for traffic monitoring practitioners will be necessary to ensure access to the data. Key features include systems for integrating quality checks and events known to data users into the database as well as for automating useful reports and data visualizations.
- Integration of traffic volume data both for motorized and for bicycle and pedestrian traffic data into the same systems, dashboard, or maps. Important transportation initiatives such as Safe Routes to Schools and Toward Zero Deaths (or Vision Zero) require information about all modes of traffic. Analysts now must consult multiple databases to assemble the data they need to work on these initiatives. Integration of all types of count data into common databases would facilitate this work.

DATA QUALITY AND EQUIPMENT CALIBRATION

The chapter on Traffic Monitoring Statistics describes how traffic monitoring experts worked for decades to establish programs ensuring quality of motorized traffic data. Elements of effective data quality programs begin with the purchase of equipment. They include procedures for installation of monitors, equipment calibration, data cleaning, and quality assurance to identify and to either eliminate or correct data errors; they also include equipment maintenance and replacement procedures. Similar to motorized data, accurate counting of bicycle and pedestrian traffic modes is technically feasible. Quality assurance protocols and procedures have been developed to ensure the validity of counts from both continuous and short-duration counts. Progress includes development of new automated checks for data validity and better procedures for reducing error in estimates of AADNT. However, ensuring and maintaining the accuracy of bicycle and pedestrian traffic counts is an ongoing issue for data managers and users.

State of the Practice

Data quality programs often are ad hoc and underfunded, limiting the implementation of quality assurance programs and the validity of count data. Funding is often available to purchase equipment, but calibration at first installation may not be done and continued maintenance and validation often is not budgeted or is not possible for overcommitted staff. The lack of resources and capacity leads to many sites with unusable data. Examples of causes of poor data quality include improper installation of equipment (e.g., infrared counters pointed toward motor vehicle traffic or loops that were improperly connected to the data logger), failure to maintain counters (e.g., insects or vandalism may obscure infrared sensors or cause overcounts), failure to replace batteries, and damages from repaving or infrastructure repairs that may cause data gaps or dramatic undercounts for loops. For most programs, increased investment in each step of the quality assurance process would improve data quality.

Another issue is that even when there are known problems with automated counters, lack of infrastructure to properly document these erroneous data leads to use of the problematic data. Lack of automated data checks leads to obvious errors such as either long periods of zeros or dramatic overcounts being used in reported averages, leading to unreliable results. Manual data cleaning is often employed as automated checks either are not implemented or are not able to flag all the data problems visible to the human eye. In many analyses, known biases (e.g., undercounts due to occlusion) are not accounted for.

State of the Art

State-of-the-art data quality programs should include readily available quality assurance procedures at each step of the traffic monitoring process. For example, after equipment acquisition, managers should pay careful attention to the installation of sensors. If sensors are used to count in locations where cycling and walking is mixed with motor vehicles, errors may occur if motor vehicles are incorrectly classified as cyclists or pedestrians. Even a small percentage of misclassified motor vehicles can obscure bicycle and pedestrian travel patterns because the volumes of motor vehicles are often large relative to volumes of bicyclists and pedestrians. For example, if 1% of motor vehicles are counted as bicyclists and there are 5,000 vehicles per day on a lane shared with 50 bicyclists per day, the 50 motorists incorrectly counted as bicyclists would erroneously double the daily bicycle total. This example indicates the need for counter calibration and validation. This is especially true in traffic mixed with motor vehicles, where nonmotorized and micromobility travelers may travel in unexpected ways relative to those of motor vehicle traffic. Similarly, a state-of-the-art program would calibrate equipment and correct data for known or probable biases, such as occlusion. For infrared counters, calibration and correction are essential to obtain valid volume estimates. Sources of data problems are numerous: insect infestations; people gathering in front of counters; heat from vehicles; and even radiant heat reflected off buildings, water, or vegetation can result in erroneous detection. Undercounts are common and caused by occlusion when people walk side by side or pass one another. This is an issue that is more pronounced in locations with higher walking density.

Emerging Trends

Both researchers and practitioners have recognized the need for better quality assurance programs and for published or reported useful protocols and procedures. As universities and data management firms expand their data management capabilities, automated data checks are becoming more common.

Knowledge Gaps

Knowledge gaps in the field continue to include how to set up a low-cost system to regularly calibrate and validate equipment in the field. Current standards documented in *NCHRP Report 797* require many hours of manual video watching to gather ground truth (*9, 10, 11*). Video counts generated by AI have not yet been proven reliable for this purpose.

Research that includes application of automated checks to a robust, sufficient sample of continuous and short-duration count data is needed. As protocols and procedures are developed and implemented, case studies to assess how these systems work in the field are needed to advance practice.

PERFORMANCE MEASURES

As noted in the fourth chapter, a principal use of monitoring data is to produce performance measures such as AADT. In the case of bicycle and pedestrian traffic monitoring, analogous measures include AADNT, AADB, and AADP. These measures, in turn, are used as inputs to develop additional measures such as bicycle miles traveled (BMT) (*26*). While these measures are analogous to those used routinely in motorized traffic monitoring (e.g., to vehicular miles traveled), the distinctive characteristics of bicycle and pedestrian traffic travel necessitate different procedures for production of performance measures. For example, bicycle and pedestrian traffic modes of travel are seasonal. This introduces additional considerations into factoring methods used to estimate measures like AADB from short-duration counts and leads to the use of metrics such as Seasonal Average Daily Traffic (SADT) (*27*). Methods for developing performance measures for bicycle and pedestrian traffic data still are evolving. As noted earlier, mode share, and thus mode shift, cannot be accurately computed from count data and therefore are not addressed in this circular.

State of the Practice

The most recent version of the *Traffic Monitoring Guide* (TMG 2022) includes Annual Average Daily Bicycle Traffic (AADBT), Annual Average Daily Pedestrian Traffic (AADPT), and Annual Average Daily Nonmotorized Traffic (AADNT) in its glossary (*12*). The inclusion of these concepts highlights progress in FHWA's institutionalization of monitoring these modes of travel. Additionally, the Travel Monitoring Analysis System (TMAS) now accepts nonmotorized data in a standardized format.

The state of the practice to generate bicycle and pedestrian volumes includes two key processes:

- Temporal: “Annualizing” short-duration counts using continuous automated count data.
- Spatial: Cross-sectional estimation of bicycle and pedestrian traffic volumes in one of multiple possible modeling frameworks.

The process of annualizing short-duration counts refers to estimating common performance metrics such as AADBT and AADPT (see Traffic Monitoring and Performance Measures) (*9,10,11*). This analysis typically involves grouping continuous counters into factor groups based on the observed temporal variation in traffic volumes, calculating expansion factors for each group, assigning short-duration counts to factor groups, and using the expansion factors to annualize the short-duration counts. This process may involve use of data from several different agencies to build the short-duration count database and to derive the expansion factors from continuous count datasets. In addition, although researchers have (a) illustrated how hour-of-day and weekend-weekday bicycle traffic patterns can classify monitoring locations into factor groups (*28*) and (b) provided guidance on the number of sites needed in factor groups to reduce error in estimates of annual average daily traffic, these procedures have not been standardized (*18*).

State of the Art

There have been some singular attempts by state DOTs and MPOs to use nonmotorized traffic count data to estimate BMT and pedestrian miles traveled (PMT). However, most measures of these quantities use household travel surveys or some third-party data products (*29, 30*). Using traffic counts in an analytic approach called direct-demand modeling or data fusion modeling, researchers have highlighted (*29, 31, 32, 33*) the promise of merging multiple data sets to produce measures of network-wide activity. However, no agency has implemented these methods in an ongoing monitoring program. Questions on their viability remain, such as how precise these processes are and how that precision would impact longitudinal reporting.

Methods have been developed to estimate AADB and AADP when traffic counts are incomplete, either because a permanent traffic counter had technical problems or permanent counters are not available for factoring. Methods have been developed that use different procedures, including simplistic factoring versus more sophisticated statistical and machine learning approaches that impute missing data (*34, 35*). These methods can improve the data completeness for permanent counters and produce a more valid and reliable AADB or AADP.

Emerging Trends

Many transportation agencies lack the dedicated staff and funding resources to have dedicated bicycle and pedestrian count programs that use the approaches common in motorized traffic count programs. One potential cost-effective solution may be to use third-party data products derived from mobile sources. These would be combined with observed bicycle and pedestrian counts, network and land use data, infrastructure data, and sociodemographic data to estimate nonmotorized travel activity. Research is still underway on the coverage, completeness, representativeness, and accuracy of the third-party data sources in bicycle and pedestrian volume estimation.

Knowledge Gaps

Even though support from federal agencies has grown for collecting nonmotorized count data, no state transportation agency has a program that continually reports on AADBT or AADPT in an ongoing fashion. Nor are there any state agencies reporting on BMT or PMT on an ongoing basis. The Minnesota Department of Natural Resources estimates and reports trail miles traveled on its 640 miles of multiuse bicycle trails following procures outlined in the *TMG* (*36*). The importance of tracking BMT and PMT cannot be overstated; these quantities are key inputs into tracking progress on meeting active transportation policy and program goals as well as better understanding changes in traffic injury outcomes. Some cities and metro regions, such as the City of Minneapolis (*37*), Delaware Valley Regional Planning Commission (*20*), and Central Lane MPO (*22*), do report on AADBT and AADPT. Some states like Texas (*24*) and Washington have made great strides in centralizing their bicycle and pedestrian traffic count data, some with help from universities such as Portland State University, Texas A&M University, and the University of North Carolina. Centralizing these count data will be key to measuring activity across the state or within regions.

DATA INTEGRATION

An emerging area of practice in active transportation count data (noted in Integrating Traffic Monitoring with Connected Vehicle Data) is the practice of data integration or estimating overall pedestrian and bicycle traffic volumes based on a combination of data sources (*16, 30*). Various data sources used in data integration include:

- Manual counts
- Automated counts
- Crowdsourced activity data
- Passively collected activity data
- Bikeshare and e-scooter fleet usage data
- Regional travel demand models
- Contextual GIS data, such as land use, transportation network characteristics, and socio-demographics of surrounding areas.

In addition to these types of data, data integration also includes collection and use of bicycle and pedestrian traffic data from state, regional, and local metropolitan planning organizations; natural resource agencies; parks and recreation agencies; and transportation and public works agencies and departments.

State of the Practice

Manual and automated counts are used together to estimate AADPT and AADBT, as described previously. Once AADPT/AADBT numbers have been produced for each segment or intersection count site, models can be developed to predict traffic volumes at locations where counts have not been performed. This often takes the form of “direct-demand modeling,” where traffic counts are predicted as a function of surrounding land uses, transportation network characteristics, socio-demographics of the population living around the count site, and other factors (*38*). In other cases, active transportation volumes are estimated through a regional travel demand model, particularly in activity-based models (*39*).

State of the Art

Incorporating user data from mobile phone applications represents the state of the art in bicycle and pedestrian data collection and integration (*40*). These data sources are typically classified into two types: passively collected user data and crowdsourced data. For passively collected data, mobile phone apps running in the background produce a series of location pings for users. The pings are aggregated and analyzed to estimate the volume of people, by mode, passing a given location. Passive data, as the name suggests, do not require users to manually report their trips. Crowdsourced data, on the other hand, is collected by users logging their trips, typically broken down by mode. In the most pervasive crowdsourced data apps, users report their trips to log their activity, such as for keeping track of their exercise. As a consequence, many researchers have noted bias in crowdsourced data (*41*).

However, despite the noted biases, both crowdsourced and passively collected activity data have proven useful in predicting active transportation volumes, particularly bicycle volumes. Multiple studies have explored integrating user activity data in direct-demand models (*33*).

Emerging Trends

Researchers and practitioners continue to advance the field. They are improving methods in annualizing short-duration counts and experimenting with integration of different types of counts from different agencies and new types of data made available by third-party vendors. When annualizing short-duration counts, the question of which factor group a short-duration count site belongs to is not immediately apparent. Some research has explored identifying this group membership as a function of adjacent land uses (*42*).

Systematic implementation of data integration algorithms on an ongoing basis is an area of growing need. As agencies adopt increasingly sophisticated data systems, these data integration processes will become more automated and routine.

Knowledge Gaps

Knowledge gaps include the lack of evidence to determine effective practices both for annualizing short-duration counts and for integrating traditional counts with third-party data. Standard procedures for classifying short-duration counts into factor groups do not exist. When annualizing short-duration count data, the underlying temporal distribution of traffic must be assumed or inferred. Initial research has explored linking this to surrounding land uses, but additional investigation into what governs these patterns and how to accurately predict them is warranted.

Standard procedures for integrating different types of data also have not been developed. Researchers are exploring alternative approaches and methods for combining and reconciling multiple user activity datasets for the purpose of predicting nonmotorized traffic volumes. Given the propriety nature of some third-party data sources, the validity of some volume estimates remains unknown, posing challenges for integration. In addition, procedures for archiving counts or volume estimates from different data sources have not been standardized.

CURRENT AND PROPOSED NATIONAL RESEARCH INITATIVES

This chapter on nonmotorized traffic monitoring described the state of practice, state of the art, emerging trends, and knowledge gaps for each of the key programmatic elements in a comprehensive monitoring program: continuous counts, short-duration counts, data management, data quality and equipment calibration, performance measures, and data integration. For each of these programmatic elements, the state of practice and the state of the art have advanced over the past 20 to 25 years, resulting in new knowledge gaps and corresponding research needs. While many detailed research needs are identified in the preceding subsections of this chapter, broad categories of research needs are relevant to all of the key programmatic elements. These needs include: the accuracy and usability of emerging technologies; the multiple different uses for bicycle and pedestrian count data; enhancing and expanding short-duration count programs; creating and integrating nonmotorized datasets; integrating emerging data with volume data; validation and calibration procedures; and automated data quality checks. Each is described further in the next section.

Proposed Research

Assessment of the Accuracy and Usability of Emerging Technologies for Bicycle and Pedestrian Detection and Counting

The use of AI and machine learning to count cycling and walking from video is a promising area of research. Although individual sites have been studied with good results, studies of the accuracy of counting for the wide-scale deployment of these types of technologies have not been published. Because new technologies (e.g., LiDAR) or different approaches to traditional technologies (radar, infrared, and video) are constantly emerging, continued research in this area is essential.

Documentation of State and Local Use of Micromobility Count Data

Increasing numbers of state, regional, and local jurisdictions are collecting micromobility count data and have expressed interest in using these data for many different purposes. However, few published documents show how agencies are using the data to inform programmatic decision-making. Surveys of agencies mention the use of counts to track trends over time, to demonstrate need when writing funding proposals for infrastructure such as new pedestrian and bicyclist trails, and to develop exposure metrics in safety studies. Anecdotal accounts and case studies of the use of counts have been presented at conferences, but there is a lack of systematic study of how agencies are using these data. Documentation of the multiple uses of micromobility count data is needed to increase understanding of the value of investments in monitoring.

Designing a Statistically Based Continuous Count Program

Agencies have typically placed continuous counters at high-volume locations to encourage more people to walk and bike and to show stakeholders the value of these installations. However, placing these counters primarily at high-volume locations can lead to gaps in monitoring the network, especially at low-volume locations. Therefore, a statistical design of continuous counter placement is needed to minimize error when estimating AADNT. Other considerations for selecting continuous count sites include both selecting them in conjunction with short-duration count sites to optimize cost and accuracy of AADNT estimates and placing them appropriately to validate the emerging crowdsourced data.

Enhancement and Expansion of Short-Duration Count Programs

Key research needs related to implementation and management of short-duration count programs include: (a) strategies for optimally locating sensors for short-duration counts in networks; (b) better procedures for annualizing short-duration counts; and (c) strategies for data fusion and integrating data from multiple sources. In addition, to the extent that limitations in funding for monitoring programs may be linked to the lack of understanding of the value of count data, case studies of the uses and value of short-duration counts in transportation planning and engineering may be useful as a means of supporting investment in monitoring.

With respect to better procedures for annualizing short-duration counts, the current state of practice involves grouping permanent counters into factor groups, calculating expansion factors within these groups, assigning short-duration sites to factor groups, and applying the expansion factors from the associated permanent counters to the short-duration counts. Research is needed into alternative approaches to this problem, including procedures for directly predicting “peak indices” for count sites (e.g., ratio of average PM peak to average midday traffic, or the ratio of average weekday to average weekend traffic) and using these to assess expansion factors. Empirically testing alternative approaches against the AASHTO methodology to assess whether additional accuracy of AADBT/AADPT could be accomplished with short-duration (e.g., 2 hour) counts.

Creation and Integration of Relevant Nonmotorized Datasets

This need is long-standing and is described in a research need statement published on the TRB website (https://www.mytrb.org/RNS/Details/137). This statement observes:

> Practitioners have access to an increasing amount of nonmotorized data from varied datasets. These can be broadly divided into three tiers: travel monitoring data (volumes), travel behavior data (surveys) and other (infrastructure and crash data). However, the lack of standards for formatting such data make data integration challenging. Practitioners are left to their own devices to consolidate and analyze the data, resulting in multiple formats and strategies that cannot be easily shared. Chapters 7.9 and 9.10 of FHWA’s Traffic Monitoring Guide provide a standard data format for nonmotorized travel monitoring data, but this format is not universally adopted and does not seamlessly

integrate with other data sets such as crash data or facility information such as sidewalk inventories. The lack of standards results in clear limitations for current analyses dependent upon nonmotorized data.

Additional research on challenges in developing and maintaining data warehouses is essential to maximize the potential of bicycle and pedestrian data.

Integration of Emerging Data (Including Trajectory Data from AI, GPS Traces from Smart Phones, Micromobility/Bikeshare Fleet Data, App Data, and Other Sources) with Volume Data

This need is related to the integration of counts into data warehouses. However, these data types are often streams of data which are more extensive and thus require substantially more storage space and different architecture from the more traditional, smaller data sources (counts, demographics, infrastructure, crash data).

With respect to integration of datasets to produce better performance measures, multiple studies have assessed the improvements in accuracy that can be achieved by incorporating user activity datasets in direct-demand models. However, few have experimented with modeling approaches for combining these datasets to account for underlying biases. This type of research would gather count data, contextual data, and multiple user activity datasets for a diversity of regions and test for optimal approaches to combining data sources to produce low-variance predictions.

Validation and Calibration Procedures

NCHRP Report 797 (*9*) and associated web-only documents (*10,11*) provide detailed procedures for counter calibration. The North Carolina Department of Transportation has implemented these procedures for validating continuous counts, but this process is time consuming and may be cost-prohibitive for jurisdictions with limited resources. Most other jurisdictions use a much more abbreviated approach to validation, if they validate equipment at all. Additional research is needed to establish reasonable standard procedures for counter validation and calibration. For example, if an agency determines counts should be within 10% of actual values (>90% accuracy), what validation procedures are required to achieve this standard, and is the standard practicable? Other related questions include operational considerations such as the frequency with which validation needs to be done after initial installation or equipment maintenance.

Automates Data Quality Checks

The lack of standard quality control checks indicates the need for further research to optimize automated quality checks and the quality checking process. Researchers have reported quality assurance checks for a variety of data problems including data gaps, consecutive zero counts, consecutive non-zero counts, asymmetries in directional flow, deviations from expected counts, and suspicious or high-volume outliers. Most of these checks have been applied to daily counts, but similar work on hourly counts has been reported. Research into the types of quality assurance most appropriate for different types of counts in different contexts is needed.

REFERENCES AND OTHER RESOURCES

1. Alta Planning and Design. National Bicycle and Pedestrian Documentation Project. Accessed at https://www.bikepeddocumentation.org/.
2. *Traffic Monitoring Guide*, Federal Highway Administration. 1985.
3. *Traffic Monitoring Guide*, Federal Highway Administration. 2013.
4. *Traffic Monitoring Guide*, Federal Highway Administration. 2016.
5. Semler, C., Vest, A., Kingsley, K., Mah, S., Kittelson, W., Sundstrom, C., and Brookshire, K. *Guidebook for Developing Pedestrian and Bicycle Performance Measures*, Federal Highway Administration, 2016.
6. Laustsen, K., Mah, S., Semler, C., Nordback, K., Sandt, L., Sundstrom C., Raw, J., and Jessberger, S. *Coding Nonmotorized Station Location Information in the 2016 Traffic Monitoring Guide Format*, Federal Highway Administration, 2016.
7. Nordback, K., Kothuri, S., Petritsch, T., McLeod, P., Rose, E., and Twadell, H. *Exploring Pedestrian Counting Procedures*, Federal Highway Administration, 2016.
8. Baas, J., Galton, R., and Biton, A. FHWA Bicycle-Pedestrian Count Technology Pilot Project, 2016.
9. Ryus, P., E. Ferguson, K. M. Laustsen, R. J. Schneider, F. R. Proulx, T. Hull, and L. Miranda-Moreno. *NCHRP Report 797: Guidebook on Pedestrian and Bicycle Volume Data Collection*. Transportation Research Board of the National Academies, Washington, DC, 2014.
10. Ryus, P., E. Ferguson, K. M. Laustsen, R. J. Schneider, F. R. Proulx, T. Hull, and L. Miranda-Moreno. *NCHRP Web-Only Document 205: Methods and Technologies for Pedestrian and Bicycle Volume Data Collection*. Transportation Research Board, Washington, DC, 2014.
11. Ryus, P., A. Butsick, F. R. Proulx, R. J. Schneider, and T. Hull. *NCHRP Web-Only Document 229: Methods and Technologies for Pedestrian and Bicycle Volume Data Collection: Phase 2*. Transportation Research Board, Washington, DC, 2017.
12. Federal Highway Administration. *Traffic Monitoring Guide*, 2022. Retrieved from https://www.fhwa.dot.gov/policyinformation/tmguide/.
13. Kothuri S., K. Nordbac, A. Schrope, T. Phillips, and M. Figliozzi. Bicycle and Pedestrian Counts at Signalized Intersections Using Existing Infrastructure: Opportunities and Challenges. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2644, pp. 11–18, 2017.
14. Day C. M., H. Premachandra, and D. M. Bullock. Rate of Pedestrian Signal Phase Actuation as a Proxy Measurement of Pedestrian Demand. Presented at 90th Annual Meeting of the Transportation Research Board, Washington, DC, 2011.
15. Singleton, P. A., and F. Runa. Pedestrian Traffic Signal Data Accurately Estimates Pedestrian Crossing Volumes. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2675, pp. 429–440, 2021.
16. Broach, J., S. Kothuri, M. M. Miah, N. McNeil, K. Hyun, S. Mattingly, K. Nordback, and F. Proulx. Evaluating the Potential of Crowdsourced Data to Estimate Network-Wide Bicycle Volumes. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2678, pp. 573–589, 2023. https://doi.org/10.1177/0361198123118238.
17. Hankey, S., G. Lindsey, and J. Marshall. Day-of-Year Scaling Factors and Design Considerations for Non-Motorized Traffic Monitoring Programs. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2468, pp. 64–73, 2014. DOI: 10.3141/2468-08.
18. Nordback, K., S. Kothuri, D, Johnstone, D., G. Lindsey, S. Ryan, and J. Raw. Minimizing Annual Average Daily Nonmotorized Traffic Estimation Errors: How Many Counters Are Needed per Factor Group? *Transportation Research Record: Journal of the Transportation Research Board*, No. 2673, pp. 295–310, 2019.

19. Lindsey, G., S. Coll, and G. Stewart. Quality Assurance Methods for Hourly Nonmotorized Traffic Counts. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2678, pp. 723–742, 2023.
20. DVRPC Travel Monitoring Counting a Region in Motion https://www.dvrpc.org/webmaps/trafficcounts/.
21. Bike Arlington Counter Dashboard. https://counters.bikearlington.com/.
22. Central Lane Bike Counting, https://storymaps.arcgis.com/stories/241cfe53fdc54602b313eeb299729031.
23. Portland State University, BikePed Portal https://bikeped.trec.pdx.edu/.
24. Bicycle and Pedestrian Count Resources. Accessed 2023. https://www.txdot.gov/data-maps/bicycle-pedestrian-count-program/bicycle-ped-count-resources.html.
25. Florida Department of Transportation. Statewide Non-Motorized Traffic Monitoring Program, https://fdot.maps.arcgis.com/apps/webappviewer/index.html?id=df6696c128514bb6b0c6710758fd050b.
26. Nordback, K., Sellinger, M., and Phillips, T. Estimating Walking and Bicycling at the State Level. (NITC), 2017. http://ppms.trec.pdx.edu/media/project_files/NITC_708_Washington_State_Pedestrian_and_Bicycle_Miles_Traveled.pdf.
27. Minge, E., Falero, C., Lindsey, G., Petesch, M., and Vorvick, T. Bicycle and Pedestrian Data Collection Manual. Minnesota Department of Transportation. https://hdl.handle.net/11299/188996.
28. Miranda-Moreno, L. F., Nosal, T., Schneider, R. J., and Proulx, F. Classification of Bicycle Traffic Patterns in Five North American Cities. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2339, pp. 68-79, 2013. https://doi.org/10.3141/2339-08.
29. Roll, J.F. Nonmotorized Traffic Monitoring and Crash Analysis. Oregon Department of Transportation, 2021. https://www.oregon.gov/odot/Programs/ResearchDocuments/SPR_813Final-Nonmotorized.pdf.
30. Proulx, F.R., Pozdnukhov, A. Bicycle Traffic Volume Estimation using Geographically Weighted Data Fusion, Unpublished Manuscript, 2017.
31. Munira S., Sener, I.N. A geographically weighted regression model to examine the spatial variation of the socioeconomic and land-use factors associated with Strava bike activity in Austin, Texas. *Journal of Transport Geography* 88, 2020.
32. Roll, J.F. Bicycle Count Data: What Is It Good For? A Study of Bicycle Travel Activity in Central Lane Metropolitan Planning Organization. Oregon Department of Transportation, 2018. https://www.oregon.gov/ODOT/Programs/ResearchDocuments/304-761%20Bicycle%20Counts%20Travel%20Safety%20Health.pdf.
33. Kothuri, S., Broach, J., McNeil, N., Hyun, K., Mattingly, S., Miah, M.M., Nordback, K., and Proulx, F. Exploring Data Fusion Techniques to Estimate Network-Wide Bicycle Volumes. NITC-RR-1269. Portland, OR: Transportation Research and Education Center (TREC), 2022. https://doi.org/10.15760/trec.273.
34. Beitel, D., S. McNee, F. McLaughlin, and L. F. Miranda-Moreno. Automated Validation and Interpolation of Long-Duration Bicycle Counting Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2672, pp. 75–86, 2018. https://doi.org/10.1177/0361198118783123.
35. Roll, J. Daily Traffic Count Imputation for Bicycle and Pedestrian Traffic: Comparing Existing Methods with Machine Learning Approaches. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2675, pp. 1428–1440, 2021. https://doi.org/10.1177/03611981211027161.
36. Robinson, M. Wascalus, J., and Gozali-Lee (2020). 2019 Minnesota State Trail Visitor Study. Minnesota Department of Natural Resources. https://files.dnr.state.mn.us/aboutdnr/reports/recreation/dnr_state_trail_visitor_study_2019_final_report.pdf.

37. City of Minneapolis Pedestrian and Bicycle Traffic Counts. https://cityoflakes.maps.arcgis.com/apps/webappviewer/index.html?id=11f21f912eef40d8bb32fb4fe94ac31b.
38. Munira, Sirajum, and Ipek N. Sener. Use of the Direct-Demand Modeling in Estimating Nonmotorized Activity: A Meta-Analysis. Safety through Disruption (Safe-D) National University Transportation Center (UTC) Program, 2017.
39. Shen, Q., Chen, P., Schmiedeskamp, P., Bassok, A., and Childress, S. Bicycle Route Choice: GPS Data Collection and Travel Model Development – Year 1 (2012-13). Final Project Report. University of Washington, 2014. Accessed at http://depts.washington.edu/pactrans/wp-content/uploads/2012/12/PacTrans-19-625083-Shen-Qing-Small-Project.pdf.
40. Nelson, T., Ferster, C., Laberee, K., Fuller, D., and Winters, M. Crowdsourced data for bicycling research and practice. Transport Reviews, 41(1), 2021, pp. 97-114. Accessed at https://www.tandfonline.com/doi/full/10.1080/01441647.2020.1806943.
41. Camacho-Torregrosa, F., Llopis-Castello, D., Lòpez-Maldonado, G., and Garcia, A. An Examination of the Strava Usage Rate - A Parameter to Estimate Average Annual Daily Bicycle Volumes on Rural Roadways. Safety 7(1), 8, 2021.
42. Griswold, J. B., A. Medury, R. J. Schneider, and O. Grembek. Comparison of Pedestrian Count Expansion Methods: Land Use Groups versus Empirical Clusters. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2672, pp. 87–97, 2018. https://doi.org/10.1177/0361198118793006.
43. Nordback, K., Kothuri, S., and Sanders, R. Creating and Integrating Relevant Nonmotorized Datasets. Research Needs Statement. https://www.mytrb.org/RNS/Details/137.

Probe Data for Traffic Volume Estimation

JOSH ROLL
Oregon Department of Transportation

MARK HALLENBECK
University of Washington

JOSEPH FISH
National Renewable Energy Laboratory

VINAY AMATYA
Pacific Northwest National Laboratory

STAN YOUNG
National Renewable Energy Laboratory

STATE OF THE PRACTICE

This chapter discusses how state transportation agencies are currently using and interacting with vehicle probe-based data for volume estimation.

Data Sources and Methods

Probe data describe the location of specific persons or vehicles in time and space. Placed in time sequence, the data become a "trace" that describes the movements of those individuals. Monitoring road user traces in this way offers great potential for traffic monitoring. However, it has significant challenges, including protecting people's privacy and developing a public agency workforce able to manage and analyze very large data sets and to work with private vendors selling products derived from these probe data. Numerous different types of vehicle probe data sources exist. Among the most common are the following:

- GPS data points reported as part of fleet tracking and management systems,
- Connected vehicle time and location reports as collected by vehicle manufacturers and aggregated by data vendors,

- Location-based services (LBS) time and location reports collected by cell phone applications,
- Cell phone location reports from cell-tower triangulation, and
- Bluetooth and WiFi device time and location reports from Bluetooth/WiFi detection hardware.

Most state transportation agencies that obtain traffic volume estimates from probe datasets get those estimates from private companies. These companies collect probe datasets from one or more sources and use various analytical techniques to estimate traffic volumes for the roadways on which those probes travel. A common method to convert raw vehicle probe data to traffic volume estimates includes applying machine learning techniques. Work done by Streetlight Data for an FHWA Pooled Fund study (*1*) described the process as shown in Figure 7. A variety of different mathematical models are leveraged in machine learning tools, including multiple types of regression (linear, nonlinear, multi-variate), artificial neural networks (*2*), and numerous decision tree style models such as Random Forest and Gradient Boosting techniques.

In addition to basic probe data, a variety of other data types are also incorporated as input to the model building and execution tasks. These datasets typically include permanent count data that provide ground truth estimates against which to calibrate or train the AADT prediction models. They also include data that provide the context necessary to convert probe data observations into traffic volume estimates, such as census, infrastructure, built environment, and even weather and holiday information.

Because probe data monitors an individual's movements in both time and space, individuals making these movements can be readily identified, as shown by Montjoye, et al., in their seminal paper, Unique in the Crowd (*3*). By purchasing traffic volume data from private vendors, agencies are relieved of not only the technical tasks of building, calibrating, and maintaining complex models, but also the data management tasks associated with the large input datasets

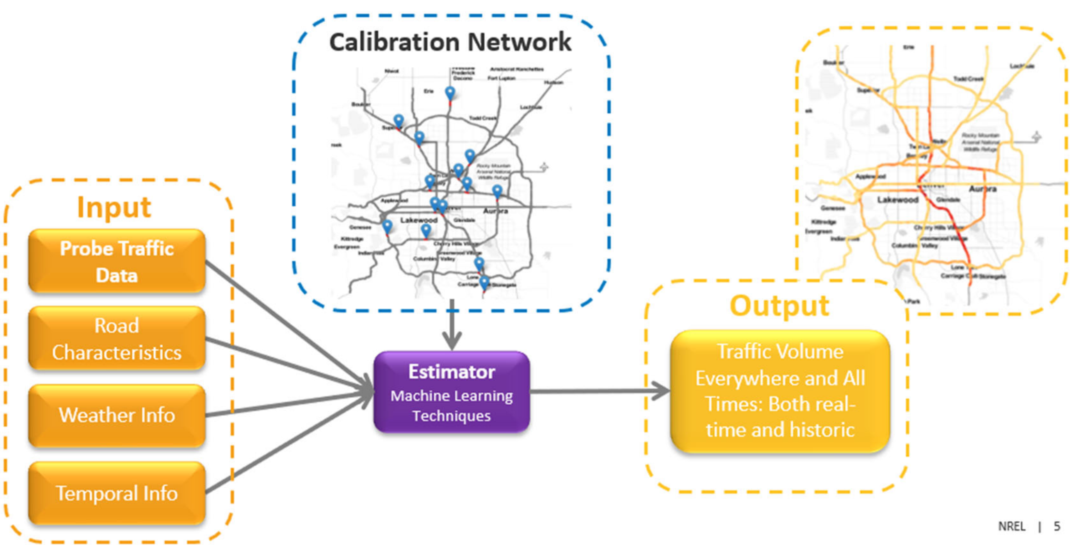

FIGURE 7 Conceptual model of use of probe data in traffic volume estimation.

required to perform those predictions. This relief includes not having to deal with the personally identifiable information (PII) associated with the probe data. Therefore, when external private companies handle these data, an agency's exposure to public records requests and the significant data security requirements associated with protecting PII are reduced.

Typically, a private vendor supplies the public agency only with finished products, such as annual average daily traffic (AADT) statistics by roadway segment, which do not have PII issues. However, even if the agency does not directly handle the raw probe data and compute traffic volume estimates, it still needs to perform several important data management tasks. These management tasks are discussed in the Data Management subsection below.

Understanding the Raw Probe Data

Even if the state transportation agency does not perform the mathematics that convert vehicle probe data to traffic volume estimates, it is important that agency staff understand the data and the process that a vendor uses to transform the raw data into traffic volume estimates. This allows the agency to be confident that its private-sector vendors are able to address potential limitations and biases in the data they are using as the basis for their traffic volume estimates. In general, the quality of vehicle probe-based traffic volume estimates is a function of the following:

- Nature of the vehicle probe data upon which the estimates are based,
- Whether the nature of those data creates biases in the traffic volume estimates,
- Number of probes reporting position data,
- Percentage of vehicles traveling on the roads in question that are included in the vehicle probe data,
- Frequency with which those probes report their position, and
- The effectiveness of the mathematical approach to convert those data to traffic volumes.

Traffic volume estimates resulting from those data become more reliable with larger numbers of probe vehicles that report data, more frequent data point reporting, higher percentages of vehicles that report their position, and probes that are more representative of the entire vehicle population.

The first two of these (large numbers of probe vehicles and more frequent reporting) have the advantage that the entire data source consists of vehicles that are being monitored. That is, every data point represents a vehicle location in time and space. Their disadvantage is that they represent a limited number of vehicles (i.e., a modest fraction of the overall vehicle fleet); the vehicles in the data set may also represent a biased sample relative to the total vehicle fleet. For example, datasets from fleet tracking systems typically contain few passenger vehicles, and those passenger vehicles typically present a specific type of fleet or vocation (e.g., taxis). As a result, fleet data generally consist of a limited subset of the trips taken by passenger cars and may cover a limited geographic area. This means that traffic volumes computed with these data need to compensate for the biases in these datasets. However, fleet tracking data can be very helpful when truck volumes are estimated within the larger traffic stream. Connected vehicle data may be biased toward certain brands as well as newer vehicles, with less expensive or older vehicles less represented.

The next two data sources (percent of vehicles reporting and representative probes) collect cell phone locations. The advantage of data from cell phones is that a large percentage of people have cell phones so data from vehicles of all types are collected. This reduces the kind of bias found in data from fleet management datasets. However, cell phones are also carried by people who use other modes of travel, including transit vehicles, bicycles, and walking. Therefore, the first task when these data sources are used is to examine the traces that result from the time and location data for each device. This allows a data analyst to estimate which mode of travel was employed at that time by that device user in order to remove any data points that are not associated with traffic volume. Only after that step has been taken can the remaining data be used to estimate traffic volumes. Another potential source of data aberration in such data is possible duplication, as cell phone data from more than one occupant in a vehicle could have been collected. The data cleaning process needs to ensure such duplicated data are properly addressed.

The final data source are approaches that collect data from cell phones but also data from other types of Bluetooth and WiFi devices. Fixed detectors observe all passing Bluetooth and WiFi devices. These passing devices are primarily, but not exclusively, cell phones but include other devices like in-vehicle Bluetooth-enabled devices. As with the cell phone-based data, the first task is to identify which devices were carried in cars, trucks, or buses and remove all other data from devices that do not represent traffic volumes.

Once probe data associated with only traffic volumes are determined, the size of the data sample can be discussed. Sample sizes for all of the above raw data sources are affected by a number of factors. For example, cell phone location data based on cell tower-to-phone communications are collected by all cellular phone companies. But these data are available only to the cell company. While they can be purchased for specific uses, in most cases, only one company's data is purchased. This limits the fraction of phones available from which to estimate traffic volume. That fraction will change from one geographic area to another, based on the market penetration of that company relative to other cell phone companies.

Changing market penetration from one part of a state to another (or from one state to another) is also an issue for LBS data. Location data is collected by numerous applications for a variety of business purposes; the most common use is to select advertisements to send to that phone. Those location data are also sold to companies interested in estimating traffic volumes. Individual cell phone applications have different geographic market penetration, and LBS data are obtained from each of those applications whenever those applications collect data (*4*). For example, applications commonly used in dense urban areas for shopping or social interaction may be far less commonly used in rural areas. Therefore, penetration rates for LBS data change from one application to another. The amount of data that certain LBS applications provide also changes from location to location across the country. Companies looking to estimate traffic volumes purchase LBS data from large numbers of applications to reduce the bias associated with any one application. However, geographic bias is still an issue with LBS data and needs to be addressed as part of the analytical process for estimating traffic volume.

Another set of issues that arise with LBS data are the policies, regulations, and practices associated with data sensitive to PII. In the 24 months coming out of the pandemic, cell phone operating system manufacturers changed their policy on what data could be gathered by applications hosted on their smart phones. This was in response to general privacy concerns (and possibly to preempt government regulation). These data policy changes resulted in a reduction in available LBS data—perhaps as much as 90%. This further resulted not only in a reduction in sampling percentage and also disrupted the continuity of the data supply chain for estimating volumes. A change in volume may be a result of policy change of smart phone manufacturer rather than real changes on the roadway.

Data Management

As noted above, most state transportation agencies purchase probe-based traffic volume data from private vendors. Consequently, they typically do not deal directly with the raw probe data that are the basis for those traffic volume estimates; therefore, managing the raw probe data is not covered in this section. However, agencies still need to manage the data they purchase and use. Topics that the agency needs to understand as it compares alternative private-sector proposals and then works with the resulting data include the following:

1. Data rights. Who has ownership, usage, publication, and distribution rights to the purchased data, and how long do those rights exist once the contract ends? If a state DOT purchases a license to the data, can they share it with local jurisdictions and MPOs?
2. Access to the data. Is the agency provided with a copy of the data it can keep, or is access to the data on a case-by-case basis via a web-platform maintained by the vendor? If a copy of the data is provided to the agency,
    a. what file format (e.g., csv, shapefile, database file) is used to transfer those data?
    b. what location referencing system is used for the traffic volume data? (Are volumes estimated for roads at specific points in space defined by X/Y coordinates? Are volumes provided for defined roadway segments, and how are those segments defined?)
3. Archiving the data. In the event that a copy of the data is not provided to the agency, how will the purchased data be accessed if the web-platform usage right is terminated at some point in the future (*5*)?
4. Data integration. How does the agency integrate the provided traffic data into the agency's current internal software systems? Most data vendors are using the Open Street Maps (OSM) as their base map.
5. Data Integration likely includes the need to develop and apply a conflation (*6*) process to match the location referencing system the vendor uses to supply data to the location referencing system used by the agency's software systems (e.g., the vendor may supply data with the Traffic Message Channel (TMC) (*7*) segmentation format, which must be conflated to the agency's HPMS (*8*), All Road Network of Linear Referenced Data (ARNOLD) (*9*), or linear referencing system).
    a. Version control. Because the models that estimate volume from probe data change over time, each dataset needs to have a model version and date of extraction so that the agency can properly reference changes in the future.

When agencies purchase relatively small amounts of traffic volume data, such as for a corridor project for which short count data are not available, the private sector-supplied data are often extracted as needed from a vendor's data portal. The extracted data are then entered manually by engineers and planners into the software used for the corridor analysis.

However, when larger traffic volume data purchases occur, such as when an agency plans to use the private-sector data to meet federal HPMS reporting requirements, the agency typically downloads a copy of the data for the entire state in one or more major downloads. Agency IT staff need to work with the vendor to understand how and when this transfer will occur and how the georeferencing of the volume data supplied by the vendor can be matched against the georeferencing system that the agency uses. The responsibility for developing the conflation tables necessary to describe the correlation between the vendor and agency data needs to be determined as part of the purchase agreement for the data, as this task can be substantial. There are typically three ways in which large volumes of network data are transferred between data systems:

- The agency delivers a base map to the vendor when the contract is signed, and the vendor populates that map with volume data.
- The vendor provides the agency with a standard base map (e.g., TMC or OSM [*10*]), and the agency must translate or conflate the data from that mapping system into the referencing system desired by the agency.
- The agency can provide specific geographic locations (X/Y points) and travel directions (heading) on roadways for which volume data are needed, the vendor supplies data on those roadways at those points, and the agency assigns those geolocated data points to the appropriate roadway location (e.g., road segment or route and milepost).

Once the conflation task has been completed, the vendor's data can be transferred into the agency's traffic data warehouse. This task is likely to require additional IT resources, especially the first time it is conducted. This is because the traffic data warehouse was likely designed around the ability to process and store short-duration counts and then adjust those counts to represent AADT values using factors developed from the agency's permanent counters. This process is very different from accepting large files of many AADT values for different locations around the state. The need to change the traffic data ingestion process also raises the question of whether the vehicle-probe volume data will contain other data typically provided by traditional short-duration counts. For example, will the probe dataset also provide either truck volumes or hourly or other time-of-day volumes? If these are provided, then they too need to be incorporated into the traffic data ingestion process. If they are not provided, then the agency needs to determine how these statistics will be provided to its engineers and planners when they are required.

Finally, there is a need for the agency to develop and implement data quality checks for the vendor-supplied data that are entered into the corporate data system. This will likely be a joint task of the central traffic office, which will oversee developing the acceptance testing rules, and the IT staff, who need to code those rules into the data acceptance and ingestion processes.

FHWA's published guidance for purchasing vehicle probe-based traffic volumes (*11*) identifies three ways to perform quality assurance testing:

- Use independent, third-party certification if such entities exist (such as the Eastern Transportation Coalition's validation program within the Traffic Data Marketplace).
- Validate the vendor's data accuracy by using either agency staff or outside assistance (e.g., a consulting firm/university) to compare data submitted by the vendor with a validation dataset).
- Trust the vendor's data quality report.

FHWA's guide also provides recommended data accuracy targets that can be adopted as part of the purchase specification.

In addition to these overall quality assurance tests, the agency must be aware that, as with short-duration counts, vehicle probe-based traffic volumes can change from year to year for reasons associated with statistical variation, changes in the vehicle-probe fleet, and changes in the machine learning model formulation over time (*12*). Any of these changes can cause unusual year-over year trends to appear in the data. These same volume trend discontinuities can occur when short-duration counts are used to estimate AADT; short counts can be affected by unusual local traffic conditions on the days when the counts occur (e.g., construction on a nearby road sends additional traffic over the count location during the week that count occurred). These unusual trends are typically identified as part of the short count review process upon ingestion of those data into the data system. Where necessary, additional short counts are then conducted to confirm or refine the AADT values for those locations.

With vehicle-probe-based AADT estimates, a process is needed to identify when AADT predictions identify unusual year-over-year patterns. Then the agency will need a new process (e.g., sending out a few short-duration count requests) to confirm when those significant changes in volume are due to actual changes in traffic patterns. It will also need a process to refine those values when the changes are not actually occurring but are the result of limitations in the vehicle-probe data or statistical variation in the machine learning process. As privacy concerns have grown with smart phone users and legislation has forced technology companies to more openly present device users with opt-out options for LBS, some data streams have changed significantly. This presents risk to agencies looking to rely on these data sources.

Data Uses

Currently, agencies are interested in vehicle-probe-based traffic volume estimation for a variety of uses, including the following:

- Address a lack of personnel to collect traditional short-duration counts
- Reduce safety concerns associated with field crews working in the roadway right-of-way to collect short-duration counts
- Reduce the costs of collecting traffic counts at very large numbers of locations spread over large geographic areas (as required by recent U.S. DOT mandates to report AADT across the entire roadway network, including low-volume roadways which traditionally have not been counted)

- Obtain other valuable traffic statistics beyond volume data, such as origin and destination patterns through intersections, that are not readily collected through short-duration counts.

Considerable interest in vehicle-probe-based traffic volume data has come from state agencies looking for less expensive ways to provide AADT statistics for all roadway segments as part of addressing changed federal reporting requirements. These same AADT statistics can be key inputs for a variety of other uses, including the following:

- HPMS and HPMS-based statewide analyses such as the Highway Economics Requirements System (HERS) model.
- Matching traffic volume data with roadway performance (speed and delay) data from the National Performance Monitoring Research Dataset (NPMRDS). This allows for a significant improvement in operational performance reporting and trend analysis by providing insight into the number of travelers affected by the delays reported by the NPMRDS.
- National Environmental Protection Act (NEPA) and State Environmental Protection Act (SEPA) analyses, which require traffic volumes as input to environmental analyses on topics such as air pollution, noise pollution, and water runoff estimation.
- Safety analyses in which traffic volumes are key measures of exposure.
- A variety of planning and operations studies in which traffic volumes describe the use of facilities being studied, serve as ground truth for calibrating models, and describe trends occurring in areas being studied.

To date, most use of probe-based traffic volume data has been for planning purposes where aggregation of volume over time and space (such as AADT and AHDT) are sufficient. As traffic volume estimation base data and techniques advance, additional operational uses may become viable. Already such volume estimates are being applied to traffic signal control applications to update timing plans for anticipated approach volumes. The Eastern Transportation Coalition ran a real-time proof of concept for monitoring hurricane evacuation in 2019, concluding that data is sufficient for such an application (though management of big data sets was still an issue at the time).

Background Reading

For readers interested in learning more about vehicle-probe-based traffic volume estimation, the following references are useful.

FHWA Guidance

- Schewel, Laura, et. al. Non-Traditional Methods to Obtain Average Annual Daily Traffic (AADT), FHWA-PL-21-030, September 2021.
- Hallenbeck, M., Schewel, C., and Wergin. Guidelines for Obtaining AADT Estimates from Non-Traditional Sources, Publication No. FHWA-PL-21- 031, September 2021.

- Tsapakis, W., Holik, S., Das, E., Kraus, and P. Anderson. Informational Guide on Data Collection and Annual Average Daily Traffic (AADT) Estimation for Non-Federal Aid System (NFAS) Roads, FHWA-SA-20-06.

Papers Describing Specific Techniques

- Hou, Y., Young, S. E., Dimri, A., and Cohn, N. Network Scale Ubiquitous Volume Estimation Using Tree-Based Ensemble Learning Methods (No. NREL/CP-5400-70896). National Renewable Energy Lab. (NREL), Golden, CO (United States), 2018.

Case Studies and Evaluations

- J. Roll. Evaluating Streetlight Estimates of Annual Average Daily Traffic in Oregon, OR-RD- 19-11, June 2019.
- Sekuła, P., Marković, N., Laan, Z. V., and Sadabadi, K. F. Estimating Historical Hourly Traffic Volumes via Machine Learning and Vehicle Probe Data: A Maryland Case Study, 2017. arXiv preprint arXiv:1711.00721.
- S. Turner. Evaluation of StreetLight Data's Traffic Count Estimates from Mobile Device Data, MN 2020-30, November 2020.
- TETC Validation Team. Traffic Volume Validation - Literature Review and Recommendations, TDM-VAL-1, November 2022.
- Tsapakis, I. "Yianni," S. Turner, P. Koneman, and P. Anderson. Independent Evaluation of a Probe-Based Method to Estimate Annual Average Daily Traffic Volume, FHWA-PL-21-032, September 2021.
- J. Roll. Evaluating Third-Party Traffic Volume Data: A Case Study and Proposal for a Data Quality Evaluation Clearinghouse. Conference Paper for Transportation Research Board Annual Meeting, 2023. https://www.researchgate.net/publication/366896831_Evaluating_Third-Party_Traffic_Volume_Data_A_Case_Study_and_Proposal_for_a_Data_Quality_Evaluation_Clearinghouse.

General Papers

- Chang, H. H., and Cheon, S. H. The Potential Use of Big Vehicle GPS Data for Estimations of Annual Average Daily Traffic for Unmeasured Road Segments. *Transportation*, 46(3), 1011-1032, 2019.
- Exploring Non-Traditional Methods to Obtain Vehicle Volume and Class Data, Transportation Pooled Fund Study TPF-5(384). Accessed: July 07, 2022. (Online). Available: https://www.pooledfund.org/Details/Study/636.
- Young, Stan, K. Sadabadi, and D. Markow. Estimating Highway Volumes Using Vehicle Probe Data – Proof of Concept, National Renewable Energy Laboratory, NREL/CP-5400-70938, 2018.
- Young, S. Real-time Volume and Turning Movements from Probe Data: Cooperative Research and Development Final Report, CRADA Number CRD-16-614, NREL/TP-540-77458, 2020.

- Zhang, X., and M. Chen. Enhancing Statewide Annual Average Daily Traffic Estimation with Ubiquitous Probe Vehicle Data, *Transportation Research Record: Journal of the Transportation Research Board*, No. 2674, pp. 649– 660, September 2020.

STATE OF THE ART

The state of the art, with respect to volume estimation from probe data sources, includes novel and emerging use cases, innovative handling of privacy issues, and thoughtful approaches to uncertainty characterization and validation.

A number of use cases for probe data have been proposed or tested beyond AADT estimation. Many of these examples are based on traditional traffic analyses, but leveraging probe data opens up new possibilities. By comparison to their conventional counterparts, probe data offer the benefit of ubiquitous coverage throughout the network as well as ongoing monitoring. On the other hand, probe data currently reflect a relatively small proportion of travel and thus special care is needed to interpret results derived from these data. More work is needed to determine the feasibility of these use cases and develop guidance for using probe data for these purposes.

Turning Movements

Turning movements are a natural candidate for application of probe data. Turning movement studies are conducted on a routine basis and, like other short-duration count efforts, are resource-intensive. Given that turning volumes represent only a fraction of traffic at any given location, turning volume estimates derived from probe data are subject to volume-related accuracy concerns noted elsewhere in this report. For instance, one study found that turning movement counts derived from probe data deviated from conventional methods by 8% and 14% for AM and PM periods, respectively (*13*). This concern may be offset by the ability to collect data over a much longer time horizon than is traditionally performed for a turning movement study.

Equity-Focused Analyses

People from communities where English is not the native language of most residents, from lower-income communities, and from minority communities may be less likely to participate in conventional travel survey data collection efforts for a variety of reasons (*14*). Probe data features the potential to fill this important gap by providing information about trip ends and travel behavior, which is critical to understanding the mobility needs of underserved populations. Probe data have been used to support development of detailed mobility profiles of underserved populations that would not have been possible otherwise (*15*).

Transit Network Redesign

Probe data have been used in conjunction with other data sources to support the redesign of transit networks to better meet community needs (*16, 17*). Importantly, the use of probe data provides insight into trips that are not made by transit, and revised transit networks can be designed to accommodate these trips. This approach holds promise, but given recent declines in transit ridership, its effectiveness remains to be seen.

Emergency Evacuation Planning and Monitoring

Use of probe data for evacuation planning and monitoring is a growing area of interest and research. Past efforts point to the value that probe data can provide to support real-time situational monitoring during emergency evacuations (*18, 19*). These include better understanding and quantification of pre-storm preparation activities, sheltering-in-place, evacuation patterns, and bottlenecks. Real-time data integration and visualization (particularly across state lines) presents a significant challenge for evacuation monitoring.

Mileage-Based User Fees

As states seek alternatives to gas taxes for transportation infrastructure funding, mileage-based user fees (also referred to as vehicle miles traveled fees) have emerged as a potential funding mechanism. Probe data has been identified as one of several options for collecting the data needed to monitor use (*20*). Smartphones, automaker telematics, and other onboard devices have been proposed as potential solutions (*21*). Privacy, equity, public acceptance, system integration, and institutional change are among the challenges facing mileage-based user fee programs. However, if these hurdles can be addressed, mileage-based user fees that are assessed through probe data collection systems could become a valuable data source for volume estimation.

Addressing Privacy Concerns

As mentioned in other sections of this document, privacy needs to be considered when working with data harvested from people’s movements. The collection and use of probe data, including raw data and products derived from probe data, raise concerns about the privacy of subjects or study participants. “Disclosure risk” is central to the discussion around privacy as it relates to probe data and can be defined as “the degree of risk that a data record from a study could be linked to a specific person or organization, thereby revealing information that otherwise would not be known or known with as much certainty (*22*).” Location and movement data have been shown to create potential for disclosure risk, particularly when subjects are tracked over time and when movements can be paired with additional personal information, such as gender, race, age, or behavioral patterns (*23*).

Probe data vendors address privacy through a variety of methods. In general, these methods result in some loss of information through anonymization, aggregation, obfuscation, perturbation, or other means. Development of synthesized datasets that seek to match the properties of the original dataset without revealing individual trips offers yet another approach

(*23*). While there is no single correct way to address privacy concerns, end users of probe data should be aware of how privacy was handled throughout the data lifecycle and how any privacy-preserving techniques may have influenced the data or limited its usability. Maintaining consumer privacy throughout the lifecycle of these probe data can help ensure that major legislative changes governing the use of these data are not introduced that then limit the utility of these data.

Uncertainty Characterization and Validation

Assessing the accuracy and representativeness of probe data sources poses another challenge for public agencies. Probe data providers should validate the accuracy of their data on an ongoing basis and share the results with potential end users. The results may include detailed information about model training and validation datasets used, relevant quality control checks that were performed (and which may influence the findings), metrics used for evaluation, and any limitations of the results. Model accuracy and reliability should be reported for different contexts, including roadway volume ranges, modes, urban and rural areas, and on the basis of other factors which may influence results.

A key consideration with respect to modeled data (such as volume estimates generated from probe data) is whether an evaluation of the model was performed against a test dataset that was never used in the model validation process. Accuracy statistics based on cross-validation alone are likely to overestimate accuracy relative to what would be observed when predictions are compared to truly independent test data (*24*). This is important because the value of products derived from probe data is in providing volume estimates where they do not already exist (i.e., where permanent counters do not exist).

To address the issues of model accuracy and validation and to confirm vendor-reported accuracy metrics, independent validation is needed. Several efforts to perform such validation have occurred or are ongoing, but these efforts are resource-intensive (*1, 11, 25, 26*). Many agencies are not in a position to perform their own robust validations and thus must rely on others to provide accuracy estimates and hope they are transferable to their jurisdiction.

Lack of independent data presents a key challenge for entities seeking to perform robust validation, as probe data vendors generally incorporate any available public data into their model development process. Agencies should explore whether data obtained from ITS sensors can be used as a potential testing data source. Additionally, high-quality verifiable short-duration counts (such as from videos or other well-calibrated counters) could serve as a testing data source. More research is needed to determine whether these opportunities are viable.

A final consideration for state-of-the-art handling of probe data quality is to account for the full range of uncertainty within the traffic data collection and volume estimation process. The same scrutiny that is given to probe data sources should be applied to traditional data collection methods and volume estimation processes. This will level the playing field for evaluating the accuracy of probe data sources relative to traditional methods.

EMERGING TRENDS AND DRIVERS OF CHANGE

The use of probe data for volume estimation offers clear benefits to transportation agencies but comes with some inherent risks. These risks generally relate to the evolving nature of probe data, both in terms of the data itself and of the market.

Evolving and Fragmented Privacy Landscape

The emergence and widespread collection and use of digital location data has elevated privacy concerns among regulatory bodies, data vendors, data users, and the public at large. At present, the United States does not have comprehensive privacy regulation similar to the EU's General Data Protection Regulation (GDPR) (*27*). However, a comprehensive bill titled "The American Data Privacy and Protection Act" was introduced in the U.S. House of Representatives in 2022 and could serve as a blueprint for such regulation going forward. This bill would have increased the transparency of data collection and transfer practices, limited the amount of data collected, provided individuals with greater control over their data (including the ability to delete their data), and implemented civil rights protections (*28*). Meanwhile, a few states have begun to address the data privacy issue by giving greater control to consumers as to how their data are collected and used (*29–32*).

While it remains unclear whether these specific laws will make an appreciable impact, the evolving patchwork of regulations and the possibility of new legislation could hinder efforts to expand data collection or develop new products. The evolving privacy landscape could also affect the usefulness of products developed with probe data, as vendors may have to change their offerings in response to new laws. Similarly, consumers may be more or less likely to opt-in to passive data collection as regulations change over time, additional information comes to light, or new methods to exercise control over personal data are implemented. How companies respond to these changes is another point of uncertainty. For example, when Google and Apple made it easier for users to opt out of location data tracking, the amount of data collected dropped substantially (*33*).

Unstable Input Data

Whether due to privacy laws or other factors, the consistency of probe data collected by a given vendor may change over time as the underlying data changes. As different apps come and go, levels of use vary, or relationships between app providers and data aggregators change, vendor products are likely to be impacted. In practical terms, these variations impact probe penetration levels (the percent of travel accounted for in probe data), which form the basis for volume estimation. The variations may also impact the representativeness of data with respect to certain population groups or types of travelers (e.g., specific demographic groups that may favor certain apps, transit users, bicyclists). While traffic data vendors should routinely calibrate and validate their volume estimation models, there is nonetheless a risk that a given product may be impacted by changes in underlying data sources from year to year or even more frequently.

Market Uncertainty

Over the past several years, numerous companies have entered the traffic data market, offering products based at least in part on probe data. While the expansion of the traffic data market has enabled new analysis capabilities, there are some risks inherent to participating in this rapidly evolving market. Some companies in the industry have been propped up by venture capital and have yet to prove their long-term financial viability. For instance, at least one prominent connected car data provider declared bankruptcy in 2023, despite its reputation as a high-quality data provider (*34*). Another similar company reported losses of $17M in 2022 (*35*). Competition among data providers is another potentially destabilizing force in the market. Agencies that come to rely on a given vendor to fulfill a core agency function would be especially impacted by future bankruptcies, consolidations, or other service disruptions that reduce their ability to fulfill their mission.

Bias

The accuracy and representativeness of data are important considerations for transportation agencies. These data influence infrastructure project development, prioritization, and performance reporting, among other agency activities. Several types of bias could impact the accuracy or representativeness of probe data, including bias introduced by differing levels of cell phone or app use by population subgroup, varying cell phone signal strength, and apps with varying levels of use by region. Temporal variation in app use introduces another level of bias. Some of these sources of bias may be accounted for through robust calibration and validation processes, but others may be more difficult or impossible to address. A recent study looking at bias in mobile location data found that minority groups, low-income households, and individuals with lower levels of education were underrepresented in mobile location data from one vendor (*36*). Bias may increase if more users opt out of providing location data in the future.

CURRENT AND PROPOSED NATIONAL RESEARCH INITIATIVES

This section describes existing and proposed research to addresses gaps in practice and knowledge for pavement engineering applications related to traffic monitoring.

Proposed Research

Determining the Technical and Fiscal Feasibility of Developing Volume Estimation Models for State DOTs

Probe data sources offer tremendous opportunity for improving how transportation authorities conduct travel monitoring. However, most of the tools developed to harness these data for volume estimation are controlled by private firms. The firms vary in transparency of their data and methods but are generally black boxes that public agencies have little depth in understanding. In addition, purchasing their products can be very expensive. Lastly, backend algorithms and data used to produce metrics like traffic volume continue to evolve, leaving past evaluation nearly obsolete. This reinserts uncertainty about the quality of these products.

One potential solution to the issues of data quality, methods, and cost could be state DOTs developing their own data fusion models using probe data available for purchase from data brokers. These probe data could be combined with freely available and ubiquitous data sources, such as measures of accessibility, network centrality, and other information from Census and the Longitudinal Employment Household Dynamics data. This combination could likely yield estimates of traffic volume with quality comparable to that from third-party data firms.

In addition to testing the technical feasibility of developing a data fusion model, this research would determine the fiscal feasibility of either a single state DOT or a collaboration of DOTs developing and maintaining such a model system. DOTs are increasingly struggling with funding data and analysis processes so understanding a sustainable fiscal plan for maintaining this capability is necessary.

Creation of a Clearinghouse for Monitoring Quality of Third-Party Data Products

An increasing number of private companies offer transportation agencies an expanding list of travel activity metrics, including estimated traffic volumes. Some of these firms aim to document and make available their methods and measures of quality. Nonetheless, much uncertainty exists in these products' accuracy, completeness, and validity, among other commonly used data quality metrics. Many agencies need to understand the quality of a data product to feel comfortable purchasing it. To ensure quality, agencies have two options: doing a primary analysis to evaluate the product or searching for past evaluations of the same product. Primary analysis and evaluation are very expensive. On the other hand, finding all past evaluations can be difficult, especially if the evaluation was done without naming the data vendor.

This research concept proposes the creation of a repository for independent evaluations of third-party data products, specifically traffic volume estimates, to ensure easy access to these evaluations by travel monitoring practitioners. It may be similar to the FHWA Crash Modification Factor Clearinghouse, in which studies of traffic safety interventions are collated and rated for quality. This third-party traffic data product clearinghouse would house and rank evaluations so that practitioners could easily understand how a product meets given data quality performance measures.

Understanding Data Quality Across Geographies, Populations, Time and Changing Data Inputs

Independent researchers have frequently evaluated and documented the quality of probe-based traffic volume estimates from private vendors. However, these evaluations are often limited to a single data provider and performed at a relatively high level with only limited amount of disaggregate reporting performed, typically at the volume bin level.

This research concept is designed to enhance our understanding of the data quality of probe-based traffic volume estimates. It involves a comparative analysis of estimates provided by various third-party firms, focusing on data quality metrics such as accuracy, completeness, and validity. The study will cover a range of geographies, including rural and urban areas, as well as large versus small metropolitan regions.

Additionally, the research aims to assess data quality for different populations, particularly in areas with higher concentrations of low-income roadway users. It will also explore data quality across various time periods. This includes examining disaggregated temporal resolutions—such as daily and monthly variations—as well as longitudinal analyses across years. The goal is to determine how differences in data inputs might affect the quality of the outputs.

By independently evaluating multiple probe-based traffic volume estimates across these diverse dimensions, the research seeks to aid transportation authorities. This information will be crucial for those interested in procuring these data products for their travel monitoring programs.

REFERENCES AND OTHER RESOURCES

1. Schewel, Laura, et. al., Non-Traditional Methods to Obtain Average Annual Daily Traffic (AADT), FHWA-PL-21-030, September 2021.
2. Young, Stan, K. Sadabadi, and D. Markow, Estimating Highway Volumes Using Vehicle Probe Data – Proof of Concept, National Renewable Energy Laboratory, NREL/CP-5400-70938, 2018.
3. de Montjoye, YA., Hidalgo, C., Verleysen, M. et al. Unique in the Crowd: The privacy Bounds of Human Mobility. Sci Rep 3, 1376 (2013). https://doi.org/10.1038/srep01376.
4. Most cell phone applications collect more data when the application is active rather than when it is open but inactive, and even less—if any—data when the application is closed but still available on the phone.
5. It is recommended that the agency ensure that it gets to maintain, in perpetuity, access to any data used for federal reporting purposes to meet various Freedom of Information Act or open government regulations.
6. Conflation describes how segments and points in one network representation can be identified in a second network representation, allowing for transfer of data contained in a system that uses one network representation into a data system that uses the second network representation.
7. TMC is the Traffic Messaging Channel format developed to allow for standardized reporting of incident information.
8. HPMS: Highway Performance Monitoring System.
9. ARNOLD: FHWA's All Road Network of Linear Referenced Data GIS network.
10. OSM: OpenStreetMap.
11. Hallenbeck, M., Schewel, Co, and Wergin, Guidelines for Obtaining AADT Estimates from Non-Traditional Sources, Publication No. FHWA-PL-21- 031, September 2021.
12. Machine learning models must be retrained periodically as the underlying factors change. In the case of vehicle-probes, the penetration rates of probes captured in the raw data stream will change over time as the usage rate of LBS data providing apps change, as well as when the apps providing data to the process change over time.
13. Barriso, J. A., and Casburn, R. *Estimating Turning Movement Counts from Probe Data*. Kittleson & Associates, Inc., 2019. https://www.kittelson.com/wp-content/uploads/2019/11/Estimating-Turning-Movement-Counts-from-Probe-Data_Kittelson.pdf.
14. Lubitow, A., McGee, J., Liévanos, R., and Carpenter, E. *Developing Data, Models, and Tools to Enhance Transportation Equity*. National Institute for Transportation and Communities (NITC), 2019. https://trec.pdx.edu/news/advancing-transportation-equity-through-inclusive-travel-survey-data-methods.

15. City and County of Denver. Let's Get Moving: Denver Moves Everyone 2050 | State of the System, 2022. https://denvermoveseveryone.com/wp-content/uploads/2023/04/DenverMovesEveryone-2022-state-of-the-system.pdf.
16. Reed, V. LA's Plan to Reboot Its Bus System—Using Cell Phone Data. *WIRED*, April 22, 2019. https://www.wired.com/story/future-of-transportation-los-angeles-bus-cell-phone-data/.
17. Massachusetts Department of Transportation and Massachusetts Bay Transportation Authority. MBTA Bus Network Redesign: Final Report, 2023. https://cdn.mbta.com/sites/default/files/2023-04/2023-04-18-bnrd-final-report-accessible.pdf.
18. The Eastern Transportation Coalition. Hurricane Proof of Concept Results: States' Experience with Real-time Connected Vehicle Data, 2021. https://tetcoalition.org/wp-content/uploads/2021/03/MT2008_Wejo_HurricaneReport_2021.pdf.
19. The College of New Jersey. Incorporating Probe Vehicle Data to Analyze Evacuation Route Resiliency: Final Report. University Transportation Research Center - Region 2, 2018. http://www.utrc2.org/sites/default/files/Final-Report-Incorporating-Probe-Vehicle-Data.pdf.
20. Minott, O. Mileage-Based User Fee Pilot Programs and the IIJA | Bipartisan Policy Center. *Bipartisan Policy Center*, February 11, 2022. https://bipartisanpolicy.org/blog/mileage-based-user-fee-pilot-programs-and-the-iija/.
21. National Conference of State Legislatures. *NCSL Road Usage Charges Summit Report*. October 3, 2022. https://www.ncsl.org/transportation/ncsl-road-usage-charges-summit-report.
22. ICPSR. (n.d.). Data Confidentiality. Retrieved August 3, 2023, from https://www.icpsr.umich.edu/web/pages/datamanagement/confidentiality/index.html.
23. Macha, M., Foutz, N. Z., Li, B., and Ghose, A. Personalized Privacy Preservation in Consumer Mobile Trajectories, 2023. *Information Systems Research*. https://doi.org/10.1287/isre.2023.1227.
24. Cross-validation: Evaluating estimator performance. (n.d.). Scikit-Learn. Retrieved August 3, 2023, from https://scikit-learn.org/stable/modules/cross_validation.html.
25. Chang, H., and Cheon, S. The potential Use of Big Vehicle GPS Data for Estimations of Annual Average Daily Traffic for Unmeasured Road Segments. *Transportation*, *46*(3), 1011–1032, 2019. https://doi.org/10.1007/s11116-018-9903-6.
26. Hou, Y., S. E. Young, A. Dimri, and N. Cohn. Network Scale Ubiquitous Volume Estimation Using Tree-Based Ensemble Learning Methods. Transportation Research Board 97th Annual Meeting, Washington, DC, 2018. https://trid.trb.org/view/1494958.
27. General Data Protection Regulation (GDPR)—Official Legal Text. (n.d.). General Data Protection Regulation (GDPR). Retrieved August 3, 2023, from https://gdpr-info.eu/.
28. American Data Privacy and Protection Act Draft Legislation: Section by Section Summary. (n.d.). U.S. Senate Committee on Commerce, Science, and Transportation. https://www.commerce.senate.gov/services/files/9BA7EF5C-7554-4DF2-AD05-AD940E2B3E50
29. California Consumer Privacy Act (CCPA). State of California, Department of Justice, Office of the Attorney General, October 15, 2018. https://oag.ca.gov/privacy/ccpa.
30. Protect Personal Data Privacy, SB21-190, Colorado General Assembly, 2021 Regular Session.
31. Virginia Office of the Attorney General. Virginia Consumer Data Protection Act. 2023. https://www.oag.state.va.us/consumer-protection/files/tips-and-info/Virginia-Consumer-Data-Protection-Act-Summary-2-2-23.pdf.
32. Plumb, T. Data Collection and Privacy: Understanding the Legal Limits. *VentureBeat*, June 15, 2023. https://venturebeat.com/data-infrastructure/the-legal-limits-of-data-collection/.

33. Newman, J. *Apple and Google's Tough New Location Privacy Controls Are Working*, Fast Company, January 23, 2020. https://www.fastcompany.com/90454921/apple-and-googles-tough-new-location-privacy-controls-are-working.
34. Taylor, M. WeJo runs out of money—Intends to file for administration | TheBusinessDesk.com. *TheBusinessDesk.Com*, May 31, 2023. https://www.thebusinessdesk.com/northwest/news/2115376-wejo-runs-out-of-money-files-for-administration.
35. Otonomo Announces Fourth Quarter and Full Year 2022 Financial Results. February 15, 2023. Otonomo. https://otonomo.io/press-releases/fourth-quarter-and-2022-results/.
36. Li, Z., Ning, H., Jing, F., and Lessani, M. N. Understanding the Bias of Mobile Location Data Across Spatial Scales and Over Time: A Comprehensive Analysis of SafeGraph Data in the United States, 2023 (SSRN Scholarly Paper 4383333). https://doi.org/10.2139/ssrn.4383333.

Appendix A: Research Ideas Summary and Scoring

INTRODUCTION

One of the purposes of compiling the information in this E-Circular is to inform the Transportation Research Board's Highway Traffic Monitoring Committee (ACP70) on worthy research ideas to consider into the next five years. The development of this E-Circular led to the creation of 45 research ideas; all but the chapter Managing Large Traffic Datasets include at least one such idea. Each proposed research idea (described in detail within the chapters above) was introduced and discussed at the 103rd Annual Meeting of the Transportation Research Board during the Highway Traffic Monitoring Committee Meeting held on January 8, 2024, in Washington, DC. Following a discussion led by each of the respective chapter leaders, committee members and friends scored each idea between 1 and 5, with higher values being given to research ideas with more merit. Nearly 50 committee members and friends, representing state and federal agencies, universities, and private firms, participated in this scoring exercise. The respondent types and their respective institutions are summarized in Table A-1.

TABLE A-1 Summary of Agency Type of Research Idea Scoring Participants

Institution Type	Count
Local	0
State	14
Federal	5
Consultant	5
University/Research	9
Vendor	3
Other	2
No Response	9
Total	**47**

The average scores are presented in Table A-2 which shows research idea titles, the total number of votes, and the average score. The scores are intended to help the Highway Traffic Monitoring Committee determine which research ideas to develop into problem statements for funding consideration.

TABLE A-2 Average Scores and Number of Participants for Each Research Idea

Chapter	Research Idea Title	Total Votes	Average Score
1	Ground Truth Method and Tools for Evaluating Accuracy, Precision, and Bias of Counting Equipment	35	3.9
2	Determine Accuracy and Bias of Portable Technologies for Obtaining Short-Term Traffic Volumes	32	3.8
	Interagency Coordination to Increase Number of Counts and Share Data	30	3.4
	Impact of Unusual Travel on Properly Trending ADT and AADT Values	32	3.0
4	Enhancing Traffic Estimation on Unmonitored Roads Using Machine Learning Techniques and Probe Data	34	4.1
	Leverage Emerging Data Sources to Enhance Traffic Monitoring and Performance Measures	33	3.8
	Congestion Management and Incident Detection Tools	31	3.5
	Best Practices in Effective Using of Performance Measures	33	3.3
	Re-examine FHWA Definition of Peak Hours Post-COVID	33	2.7
5	Advanced Methods for Project-Level Traffic Loading Estimation	24	3.7
	Improving Accuracy of Traffic Inputs	25	3.6
	Enabling Detailed Traffic Loading Data Collection	23	3.4
	Freight and the Urban Environment	23	3.1
6	Data Collection and Funding Methods to Obtain AADT for Lower Functional Class Roadways	33	3.9
	Determine Accuracy of Different Types of Traffic Counts for Diverse Applications	33	3.6
	Automated Site Calibration: Use of Probe Based and Roadside Readers to Verify Classification Site Accuracy	32	3.2
	Develop Data Imputation Methods for Missing Traffic Count Data	35	2.8
7	Integration of CV Data into Traffic Monitoring Programs	33	3.7
	Maximizing Benefits and Value of CV Data	31	3.3
	Privacy Issues Related to CV Data	31	2.9
	Impact of Big Data Analytics on Extracting Value from CV Data and Risks of Compromising Privacy	29	2.9
	Develop Guidance and Requirements for Advanced Messaging for Connected Vehicles	28	2.9
8	Portable WIM Solutions	23	3.8
	WIM Data for Strength and Safety Assessment of Existing Bridges	21	3.7

Chapter	Research Idea Title	Total Votes	Average Score
	Remote WIM Sensor Calibration Using CVs, Onboard WIM Systems, and Smart Road/WIM Infrastructure	26	3.5
	WIM to Support Autonomous Truck Weight and Size Enforcement for CVs	26	3.3
	Road Safety Improvement Using CVs and WIM with Tire Anomaly Detection Sensors	23	3.1
9	Large-Scale Network Conflation	23	3.9
	Best Practices in Travel Time, Speed, and Reliability Data Processing, Integration, and Applications	25	3.9
	Investigating Best Approaches to Working with Trajectory Data by Transportation Applications	25	3.8
	Use of Probe Vehicle Data for Longitudinal Performance Assessment	25	3.7
	Need for Standardized Travel Time Reliability Measures and Network Monitoring Metrics	29	3.4
	Efficient Travel Time/Speed Distribution Models by Roadway Types for Travel Time Reliability Measurement	25	3.4
	Investigating Probe or Connected Vehicle Data Latency Issues	28	3.1
10	Accuracy and Usability of Emerging Technologies for Bicycle and Pedestrian Detection and Counting	27	4.3
	Designing a Statistically Based Continuous Count Program	24	4.2
	Enhancement and Expansion of Short-Duration Count Programs	25	3.7
	Validation and Calibration Procedures	26	3.7
	Documentation of State and Local Use of Micromobility Count Data	25	3.6
	Automated Data Quality Checks	27	3.6
	Integration of Emerging Data with Volume Data	26	3.6
	Creation and Integration of Relevant Nonmotorized Datasets	24	3.5
11	Understanding Data Quality Across Geographies, Populations, Time and Changing Data Inputs	26	3.9
	Determining the Technical and Fiscal Feasibility of Developing Volume Estimation Models for State DOTs	25	3.6
	Creation of a Clearinghouse for Monitoring Quality of Third-Party Data Products	25	3.6