TRANSPORTATION RESEARCH
CIRCULAR
Number E-C278 March 2022

Innovations in Freight Data Workshop

September 21–23, 2021

The National Academies of
SCIENCES · ENGINEERING · MEDICINE
TRB
TRANSPORTATION RESEARCH BOARD

TRANSPORTATION RESEARCH CIRCULAR E-C278

Innovations in Freight Data Workshop

September 21–23, 2021

Kathleen Hancock
Virginia Polytechnic Institute and State University
Rapporteur

Sponsored by

California Department of Transportation
Florida Department of Transportation
Iowa Department of Transportation
Texas Department of Transportation
Washington Department of Transportation
Wisconsin Department of Transportation
Federal Highway Administration Office of Freight Management and Operations

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

TRANSPORTATION RESEARCH CIRCULAR E-C278
ISSN 0097-8515

The **Transportation Research Board** is one of seven major programs of the National Academies of Sciences, Engineering, and Medicine. The mission of the Transportation Research Board is to provide leadership in transportation improvements and innovation through trusted, timely, impartial, and evidence-based information exchange, research, and advice regarding all modes of transportation

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.

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

Preface

The TRB Standing Freight Data Committee initiated the third workshop on innovations in freight data to bring together freight data users and decision-makers to learn about and share the latest applications that leverage new and maturing data sources for improving decision-making. This event brought together traditional freight planning stakeholders with data and technology innovators from related areas to explore opportunities to advance the state-of-the-practice, offering an interactive format to engage in productive dialogue.

During the virtual workshop, on September 21–23, 2021, 186 freight community members—including transportation professionals from transportation planning agencies, consultants, industry experts, and academic researchers—participated in the Third Innovations in Freight Data Workshop.

Alison Conway from City College of New York chaired the ad hoc planning committee for this workshop. The planning committee members were solely responsible for organizing the workshop, preparing the call for abstracts, reviewing the submitted abstracts, and developing topics for workshop sessions. Kathleen Hancock from Virginia Polytechnic Institute and State University served as the workshop rapporteur and prepared this document as a factual summary of the presentations and discussions at the workshop.

In 2017, the first Innovations in Freight Data Workshop highlighted “big data” as a potential new data resource for freight planning, freight operations and mobility, and visualizations for communicating freight issues and solutions for decision-makers. The event brought together traditional freight planning stakeholders with emerging technology innovators to look for opportunities to interact. Two years later, the 2019 event built upon these opportunities, bringing the stakeholders with recent experiences with big data for freight—and individuals, organizations, and agencies—to evaluate progress with new data sources and new methods and to discuss next steps to meet existing and future challenges. Now, the third in the series expanded the focus to highlight how freight data is used to make decisions.

This workshop report follows the program agenda and includes summaries of the presentations made by each panel member, augmented with audience questions, comments, reflections on advances in data sources and the decisions they support, along with new challenges. Topics covered in the various panels, interactive posters, and traditional posters included new uses of Global Positioning System (GPS) trace data, new machine learning techniques, new uses of existing freight data, and a variety of new approaches for collecting and analyzing freight data. A copy of the program with links to the presentations is available at https://trb.secure-platform.com/a/page/InnovationsFreightData/Virtual_Program.

Special acknowledgments go to the Transportation Research Board staff Tom Palmerlee and Gary Jenkins for their support and organizational expertise. Special thanks are extended to the Iowa Department of Transportation for its leadership of the pooled-fund study for this workshop and the five other state departments of transportation pooled-fund members; California, Florida, Texas, Washington and Wisconsin. Thanks also go to the Federal Highway Administration. Thank you as well to Brittney Gick and Donald Ludlow who served as peer reviewers of this document.

Contents

INTERACTIVE SHOWCASE

APPENDIXES

Opening Session

WELCOME AND CHARGE TO PARTICIPANTS

ALISON CONWAY, *City College of New York*

Building on the success of its predecessor workshops in 2017 and 2019, the 2021 Innovations in Freight Data Workshop brings together the community of freight data users and developers to share, discuss, and learn about recent advances in freight data with a focus on its use in making decisions. Panelists and presenters addressed topics that included:

- Emerging data needs for public and private-sector decision-making;
- Recent advances in the application of emerging data sources for broad multimodal freight applications;
- Experience and lessons learned from integrating new data sources in practice;
- Advances in data visualization, data collection technologies, artificial intelligence and machine learning, and data analytics methods with potential applications for freight; and
- Promising advances and future needs for continuing advancement in freight data research and practice.

Sessions are presented through a combination of live virtual formats to enable attendees to directly interact with presenters and to participate in discussions about important advances and the future of freight data.

KEYNOTE ADDRESS

JANNINE MILLER, *Georgia Department of Transportation*
KATHLEEN HANCOCK, *Virginia Tech, Transcriber*
(Note: Bold text indicates innovative data sources that were identified by the speaker.)

I am excited to participate in this event, the third tranche of diving in on sources of freight data and their uses, including sharing what is possible amongst practitioners and experts from every field. The importance of planning our infrastructure for the movement of freight is probably one of the most important things that we can do, as infrastructure owners and planners, for our economy, our prosperity, our quality of life, and our sustainability over the coming decades. To get to the best, most effective plan, we need the best, most-effective data.

One purpose that brought me here is to set the stage and lay out some of the topics of the workshop which bring more detail and a deeper perspective to innovations in freight data. My goal is to provide some high-level aspects of this. Specifically, I want to focus on nontraditional data sources, related information, and perspectives that are atypical for infrastructure owners and planners.

Number one, I want to help bring the relevance to industry about the vitality and the criticality of the freight planning work that we do. We want to speak their language and understand the pressures felt by carriers, third-party logistics companies, trucking, rail, all modes, and the shippers who own the cargo and who depend on that cargo to arrive safely and reliably. If we can understand their perspective and their landscape, it will help to communicate why what we do is important. We want to convey to them that we are looking out for the reliability of their supply chain and that our infrastructure is being planned for redundancy and resiliency, thus addressing the pressures that are on them now and continue to be on them. The speed of supply chain and retail and manufacturing versus the speed of infrastructure development are worlds apart, so having that common language is going to be important.

Second, I am going to challenge the audience to think about the future of data availability. We are embarking on a new frontier in what is going to be possible as data availability increases for planning, operations, safety, and technology. I hope that, as we build out technology investments, we think about ways that we can mine the data that we are going to be collecting in ways that will advise infrastructure planning.

Finally, we will talk about COVID-19 and the state of logistics, e-commerce, freight infrastructure and the economy, and freight data for roadway infrastructure decision-making.

COVID-19 and the State of Logistics

We will begin by considering two perspectives of ongoing supply chain challenges: 1) worldwide motor vehicle production demonstrating the change in personal spending, and 2) housing supply versus price change.

The worldwide motor vehicle production increased from 55 million vehicles in 2001 to nearly 100 million vehicles in 2018, with a downturn of nearly 50% to less than 80 million vehicles in 2020.[1] Part of this is associated with the shortage in semiconductor chips necessary for the increasing digitization of our vehicle fleet. This shortage, according to Secretary of Commerce Gina Raimondo in June 2021, is going to be a daily problem for the next year or so. **From a data source perspective, consider the data that will be available from semiconductor chips that are required for our current and future vehicle fleets.**

We are also seeing ongoing supply chain challenges in housing. The housing supply is nowhere close to meeting the demand, including the goods that go into both the building and improving of these homes. The current supply of homes is below what was available 20 years ago, while the spike in demand is higher than it was at that time[2] and is the result of a COVID-aggravated "bullwhip effect" defined as a "massive ripple effect in the [supply] chain…like cracking a bullwhip". This is caused by "a change in demand at the retail level, disrupting all other partners, including distributors, wholesalers, manufacturers, and raw material suppliers, who are now also working to meet that new demand."[3] The bullwhip effect resulting from COVID-19 will "mean that consumers will face certain shortages of products in the near-term and will be paying more for those products in the mid-term."[4] We are now working to meet that new demand, so the bullwhip effect is in full play. Another way to think about it is that delay begets delay with a tremendous historic spike in demand compounded by a supply suppression.

The COVID-19 and logistics story is the story of a shift in consumer spending, sourcing that incites manufacturing adjustments, and supply disruptions. The best approach that we have identified in modern economy for managing the bullwhip effect is use of third-party logistics (3PL) companies. From 2019 to 2020, the 3PL market increased 32% in gross revenue to $281 billion.[5] Projections are for this to give a little, but continue on a growth pattern. 3PLs are the ones who shippers will look to for getting product where it needs to go at the lowest price at a specified level of service. This service potentially includes selection of mode and route. **Data**

from their quarterly or annual reports can provide insights into the pressures that the industry is experiencing.

A cautionary tale about nontraditional data sources—every year the Council of Supply Chain Management Professionals (CSCMP), has published an annual logistics report. They provided one this year in conjunction with A.T. Kearney and Penske Logistics that showed 2019 to 2020 year-over-year parcel costs increased 24%. It also showed the cost to ship over water declined by 28.6%.[6] This decrease may not be an accurate reflection of what is reported in the business news, which is reporting historic growth in the cost to ship a container, especially from Asia to the United States. The first caution, when considering nontraditional data sources, is to consider the reporting period, in this case, 2019 to 2020. Even though this is the most recent annual logistics report, it actually reflects a year ago, which reflects a different logistics environment, given the changes that occurred in response to the pandemic. The second caution is to consider what happened within the data itself. One of the main reasons the reported water transportation costs were down was because of a one-time reclassification in the underlying calculation methodology. Therefore, it is important to take a deeper look into the data to understand the context of what is being reported.

One data source, with potential to link public-sector operations and planning to the economy, is the Logistics Managers Index (LMI) provided by the Georgia Center of Innovation for Logistics. This provides a market snapshot of the same data every month, and is a useful resource for U.S. trade, retail, and e-commerce statistics and includes how the different modes are faring, whether it is rail, trucking, or ports, as well as warehousing and distribution. Specifically, LMI is a combination of multiple factors, including inventory levels and costs, warehousing capacity, utilization and prices, and transportation capacity. In July 2021, this index was 74.5, which tied for the third highest value in the history of the index and it has been above 70 for six consecutive months. Any value above 50 indicates increases in warehousing and transport costs. A 74.5 LMI is significant. With informed data processing, these trends could provide useful information for infrastructure operators and planners.

The emphasis on data-driven decisions has a lot of potential upside and profits for the private sector. DHL, an international shipper, wrote “More companies are shifting towards a data-driven mindset in their decision-making. [Internet of Things (IoT) technology is] forecasted to open up a $1.9 trillion opportunity in logistics.”

Covid has added to the pressures as well as the opportunities occurring in the current logistics environment. The initial thought is that Covid will pass, the bullwhip effect will fade, and we will return to normal. However, from a reader survey performed by *Logistics Management Magazine*, 72% of respondents indicated that the pandemic has had a major impact on their logistics and supply chain operations and 60% indicated that pandemic-related changes in supply chain could be permanent.[7] Once a supply chain shifts and suppliers of logistic services develop new relationships and processes, returning to pre-Covid operations is unlikely.

Growth in E-Commerce

E-commerce is a growing industry as a result of the shifts and changes in consumer demand and, more recently, the effects of Covid. From the Georgia Center of Innovation in Logistics and the U.S. Census Bureau, as of July 2021, e-commerce grew by 3.3% from the first to the second quarter of 2021, and 9.1% year-over-year from July 2020 to July 2021. This is incredible growth that is rippling through supply chain decisions throughout every industry. E-commerce sales accounted for 13.3% of total sales in the second quarter of 2021. At the same time, the U.S. industrial vacancy rate is down, tying the all-time-low rate of 4.5% last registered in the fourth quarter of 2018. This directly relates to our transportation infrastructure through the rapid expansion of e-commerce centers on the fringes of major metropolitan areas.

Georgia Department of Transportation (DOT) did an analysis in early 2021 that determined that e-commerce is going to have the largest impact on changes in our freight mix. E-commerce acceleration is going to drive a lot of changes in the way that logistics impact our infrastructure, as highlighted in Figure 1. Retailers and distributors are under increasing pressure to deliver consumer products and perishables into cities, often with very narrow timeframes, creating the need for more adequate delivery sites in and around dense urban areas. We have shifted from a single channel of regional-distribution-to-local-warehouse-to-retail-center to multiple channels moving all across the transportation network, resulting in transportation costs increasing between 10% to 15%. This has also resulted in last-mile transport shifting from personal chained trips by automobile to business delivery vehicles with trips that can be much longer than the traditional last mile. In most metro areas the last mile is between 5 and 9 mi. But once again, these costs are going up and shippers are relying on the infrastructure for timely delivery. This quote from United Parcel Service (UPS) emphasizes the impact of congestion: "If every UPS vehicle is five minutes late due to congestion per day, it costs the company $114 million annually."[8] This is a hit to their margin that they, as private industry, will look for any ways to reduce their supply chain and transportation costs and services. One way we are seeing that reduction occur is through automated delivery based on a recent conditional approval by U.S. DOT for the driverless Nuro vehicle in limited testing sites under strict safety protocols.

All of these provide options and potential data collection that might be done by private sector and that government entities, could in the future, procure so to understand this emerging demand.

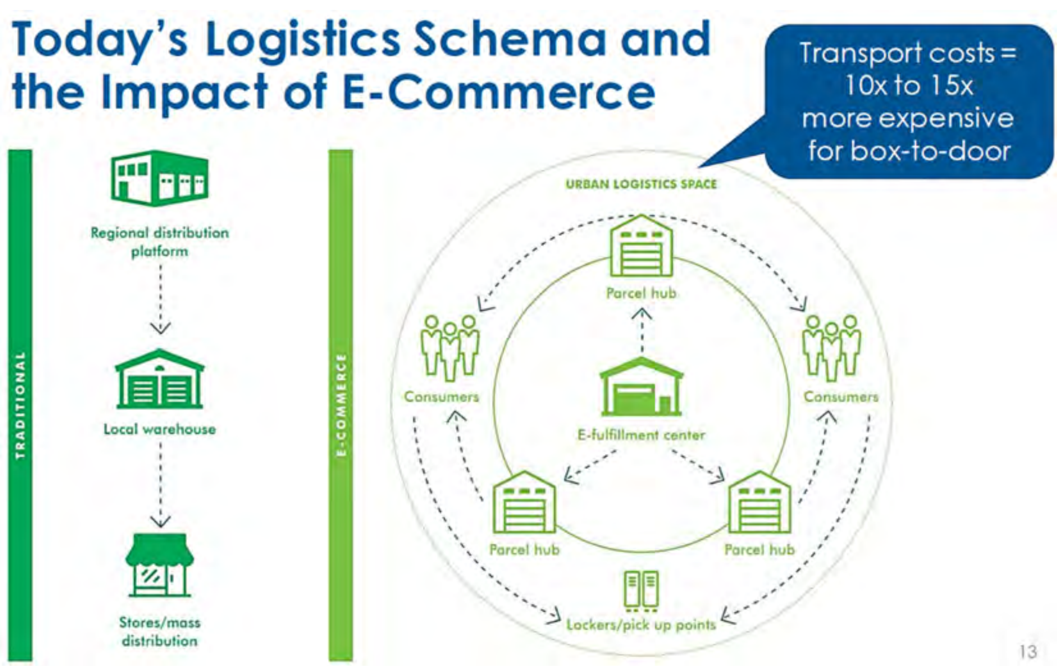


FIGURE 1 Impact of e-commerce on logistics.

Data and the Economy

The continued growth in the economy ensures the continued growth in goods, even with the blips that occurred due to the recession of 2008 and, more recently, COVID in 2020. According to Bloomberg, shipping rates are up more than 500% for a container moving from China to Europe. Likely trade growth in Georgia through 2050 is 77% for manufacturing and 43% for agriculture, forestry, and fishing. That is inbound and outbound logistics that are going to be putting demands on our freight infrastructure.

To begin to address the latter, **the U.S. Department of Agriculture (USDA) has a new data source and open data platform** (https://agtransport.usda.gov), which provides information for four agricultural transport modes: truck, rail, barge, and ocean. It also includes the online version of the *2021 Compendium of Agricultural Transportation Research*. It includes a playbook for how to use this data for creating an infrastructure framework based on the recent report, *The Importance of Highways to U.S. Agriculture*, which was produced in conjunction with USDA, U.S. DOT, and the Volpe Center.

The proposed Infrastructure Investment and Jobs Act includes an expansion of the Bureau of Transportation Statistics (BTS) data analysis tools to spur economic development through infrastructure development, connecting supply chains for freight planning.

Data for Freight Infrastructure Investment Decisions

As mentioned earlier, we are experiencing a surge in demand and a shortage of supply of logistics and goods that need to be moved during COVID. That cost of delay, something that we, as infrastructure managers in the roadway sector, are quite familiar with, happens every day as a result of incidents on our highways, in our growing metro areas, and along our Interstates. If we have 8,000 vehicles moving on a four-lane road segment, and a crash occurs, we obviously increase delay. With 7,600 cars at $20/h and 400 trucks at $70/h, a single hour of delay is roughly $180,000. This is a data-driven component of the decision-making process.

In Georgia, the DOT Connected Infrastructure Initiative provides a wide spectrum of opportunities to address delay and other transportation issues. These include:

- Signal priority and preemption for transit and emergency vehicles;
- I-95 information technology service (ITS) for faster incident response times;
- In-cab messaging where volunteer trucking fleets at the Port of Savannah receive select signal prioritization and onboard traffic alerts; and
- Traffic Vision for immediate detection and verification of slow-downs and incidents in combination with faster response from Georgia DOT.

All of this is going to generate a lot of data that Georgia DOT will be able to harvest and incorporate into bigger datasets as we plan for a safer and more effective, efficient network throughout Georgia. Figure 2 provides a table of challenges faced by transportation asset managers unique to freight with corresponding ITS solutions and example implementations.

Georgia is piloting in-cab travel alerts with DriveWyze, which let truckers know that they are entering a high-crash or congested section. The results to-date include nearly 20% reduction in hard braking, which results in fewer crashes. **This data will help us understand where to improve infrastructure in the future.**

Freight Traffic's Unique Challenges to Transportation Asset Managers

Industry Freight Challenges	ITS Solutions	Example Implementation
Truck Parking	Truck Parking Availability Systems "TPAS"	FDOT's Truck Parking Availability System
Truck size & weight limits	Weigh-in-motion systems	GDOT's Weigh-In-Motion Program *(nationally recognized)*
Freight Corridor Bottlenecks	Truck Platooning, Dynamic/Dedicated Lanes, and Traffic Signal Preemption for Freight	**Peloton** Truck Platooning Technology, GDOT I-75 CVL Express Lanes *(under design)* GDOT's SR 21 / SR 307 Freight Signal ITS
Internet of Things "IoT" Availability (tracking & tracing)	Right-of-way fiber cabling	GDOT Fiber Network
Safety	Connected Vehicle Roadside Units, Workzone Safety Notifications	GDOT Roadside Safety Units, **iCone** (IoR – "Internet of Roadwork")
"Last-Mile" delivery & Distribution	Delivery-space booking systems, Autonomous Delivery Vehicles	**curbFlow** pilot in Columbus, OH; **Amazon** Scout / Starship / Nuro *(robotic delivery devices)*
Environment: Air Quality Considerations	Truck Electrification	**Volvo / Telsa / Scania**

FIGURE 2 Challenges to transportation asset managers unique to freight traffic.

Georgia also has the Ray, which is a multifaceted suite of technology applications on I-85 in central Georgia near the Alabama border and near the KIA auto plant. This is a test bed for experiments and includes solar arrays, a vehicle-to-everything (V2X) ecosystem, and a data management platform. Testing has included traffic and work zone warnings, freight signal priority, and roadway maintenance notification. **As resulting data become available, the goal is to incorporate it into the freight plan.**

For connected autonomous vehicles, it is unknown what the data opportunities will be. These technology deployments may lead infrastructure owner operators to partnerships and stronger data linkages with automakers to provide them with the data they need for autonomous vehicles. An example is the cross-border ITS implementation, Nordic Way. Based in Europe, Nordic Way connects traffic management centers and technology platforms with a common architecture across Scandinavia to more effectively move goods in the challenging Nordic environment. The partnership includes the transportation agencies of Sweden, Norway, Denmark, and Finland and private partners such as Ericsson, Volvo, IBM, and Kapsch and will test traffic management strategies, in-vehicle notifications, and automated driving in arctic conditions. The prospects for mining data are exciting.

Georgia DOT plans to initiate their freight logistics update: which they are calling a business plan for freight infrastructure. This will include private-sector involvement and direct collaboration with Atlanta's metropolitan planning organization (MPO). Much data will be available for the process. The approach will include a supply chain focus with freight fluidity analyses similar to the recent Federal Highway Administration (FHWA) deployments. Including the private companies who fund our infrastructure and who rely on it for their businesses will ensure a robust plan. For example, one key issue in Georgia is truck parking supply. We are currently using National Performance Management Research Dataset (NPMRDS) data to understand where trucks are stopping and to get a sense of where they are parking. We plan to expand this understanding through procuring and integrating private-sector Geotab data, which has telematic-based fleet data that includes fleet size and composition, geography and regional insights, and driving patterns.

From the freight fluidity multimodal perspective, we will be incorporating the Geotab data and private supply chain data from DAT Freight and Analytics (previously Freight Market Intelligence Consortium from Chainalytics). As an example, using the North American Industry Classification System (NAICS) code for food products manufacturing, which is critical to Georgia, we can identify the origin and destination points and understand the demand, as well as the volume and timing of that demand on our infrastructure. We can then plan where we might need to add capacity, improve connectivity, and expand capabilities of our rail network and intermodal terminals.

We are very excited about what expanding data sources are going to allow us to do. Thank you to FHWA for leading that path and to all the support that we have received from our State Transportation Board and the governor’s office for this exciting work.

NOTES

1. *Statista*, April 2021.
2. Federal Reserve, July 2021.
3. Redwood Logistics.
4. Insidelogistics.com, August 2021.
5. Armstrong & Associates, Inc.
6. Logistics Management, July 2021, courtesy of A.T. Kearney, CSCMP, and Penske Logistics.
7. Logistics Management, Reader Survey June 17, 2021.
8. Tom Jensen, Vice President of Transportation Policy, UPS, InfrastructureUSA.org, Jan 2020.

PANEL

Informing Decision-Making with Freight Data

ERIC CARIS
Port of Los Angeles

HANNAH WALTER
California Transportation Commission

In this session, two experts from leading transportation agencies discussed how their organizations use freight data for strategic and operational decision-making.

INTRODUCTION

Caroline Mays opened the session by laying the groundwork for the workshop and the sessions from a state agency perspective. The panel members then introduced themselves, described their agencies, and presented how their agencies use freight data for informing decisions.

Freight is continually transforming, and state and local agencies are faced with emerging issues such as shifting supply chains, increasing automation, new and expanding technologies, and innovations in almost every aspect of goods delivery. Traditional concerns, such as congestion and safety, are expanding to include things like truck parking. Mays stated that freight professionals face challenges which require data and the resulting information and knowledge that it generates, and which require effectively communicating to agency leaders and decision-makers. She specifically identified the following areas:

- Justifying freight: why freight is important.
- Planning: what data can be analyzed and transformed to support defining priorities in planning decisions on funding, programs and projects.
- Policies: what types of data and how is it analyzed and transformed into information to identify policies related to freight movement.
- Operations and engineering.
- Funding is a big component of decision-making. Transportation leaders are making funding decisions from how many staff are needed for freight planning to justifying freight projects.

Her final point was defining who the decision-makers are. In Texas, several levels of agency decision-makers report to the executive leadership which reports to a commission, and the commission reports to the governor. There are also decision-makers in the private sector. Decision-makers exist at the local, state, and federal level beyond those in transportation. How do freight data users transform data to information so that they can make informed decisions?

THE PORT OF LOS ANGELES

Eric Caris is the Director of Marketing for the Port of Los Angeles. He is responsible for marketing, sales, and promotional programs designed to attract and retain shipping business to the port. A significant activity involves the negotiation of container terminal leases in collaboration with the Cargo and Industrial Real Estate Division. On the efficiency side, Cargo Marketing plays an integral part of implementing the Port Optimizer program, which provides supply chain stakeholders with added visibility to the supply chain. The Port of Los Angeles is the busiest seaport in the western hemisphere. Located in San Pedro Bay, it encompasses 7,500 acres of land and water along 43 mi of waterfront and includes both passenger and cargo terminals, including cruise, container, automobile, breakbulk, dry and liquid bulk, and warehouse facilities that manage billions of dollars' worth of cargo each year. The port is also focused on new technologies to enhance digital information flow throughout the supply chain.

After providing statistics about the current operations of the port during and post COVID, Caris described the Port Optimizer developed for the port by the Wabtec Corporation. It is a cloud-based software solution and data repository that provides real time (updated every 30 min) information and analytics to the port's supply chain partners. Data sources include:

- Ocean carrier data: manifest data, booking information, stowage, destination terminal and rail–truck destination information;
- Marine terminal data: movement events, gate moves, yard location, last free day, and holds;
- U.S. Customs and Border Protection data;
- Southern California Marine Exchange Vessel Traffic Service data; and
- GeoStamp data.

Components of Port Optimizer include:

- Track & Trace for container visibility and planning. Its objective is to provide visibility for operational decisions of intermodal containers.
- The Port of Los Angeles signal for high-level visibility and data metrics. Its objective is to provide 3-week forward projections of import volumes for chassis, truck and rail stakeholders.
- The Port of Los Angeles return signal for improving empty returns. Its objective is to provide empty return information by size, carrier, and type for all terminals across time in a searchable format.
- Control tower for reporting portwide performance indicators including truck turn times, cargo dwell times, import–export volumes, rail metrics, and financial metrics. The Horizon view provides 6-month predictions by 20-ft equivalent units (TEU) or container.

These components support decisions to optimize the role of the port in the supply chain. Knowing warehouse capacity and use of containers as storage; truck chassis turnover and availability; street and rail dwell times are all important for short-term decisions, such as routing cargo and long-term investments, to maximize land use.

THE CALIFORNIA TRANSPORTATION COMMISSION

Hannah Walter is an Associate Deputy Director of the California Transportation Commission (CTC), which is responsible for programming and allocating funds for the construction of highway, passenger rail, transit and active transportation improvements throughout California. She is responsible for the Trade Corridor Enhancement Program programming, which determines where to allocate funding for freight infrastructure projects.

Projects submitted for consideration must include performance metrics for three scenarios—no build, build, change—and include the following:

- Change in daily vehicle and truck hours of delay;
- Change in truck volume;
- Change in rail volume;
- Velocity freight;
- Truck Travel Time Reliability Index (TTRI);
- Air quality;
- Cost effectiveness;
- Vehicle crash safety; and
- Jobs created.

Table 1 summarizes the datasets used to generate these metrics along with their sources. The application process is well-defined and allows the CTC to equitably evaluate proposals. However, Walter stressed the importance of communicating this information effectively by understanding how people make decisions and who the audience is. In particular, she emphasized presenting data effectively to decision-makers so that they can answer questions like: 1) What is a “good” project; 2) What are the benefits of the project?; and 3) Why is this important now?

Walter then discussed data gaps, including more disaggregate freight flow data, understanding where trucks are, and maritime and rail information. CTC is working with the U.S. Army Corps of Engineers (USACE) to develop imputation techniques and predictive highway speed and multi-aspect traffic models. They are also developing a “Forcing Model” relating port, terminal, and traffic congestion to current and predicted economic demand using

TABLE 1 CTC Regional Funding Target Datasets

Dataset	Source
Value by customs district	USA trade: imports and exports
Annual truck vehicle miles traveled	CSTDM or regional data
Annual average daily truck traffic	Caltrans Traffic Ops
Daily vehicle hours of delay	Caltrans Performance Measurement System
Emissions: tons per day	EMFAC/OFFROAD
Maritime TEU	USACE and port data
Population	Department of Finance
Disadvantaged communities	CalEnviroScreen
Border freight	Regional data, northbound land POE tonnage

econometric indicators. Other tasks include developing an analytics toolkit for evaluating policy interventions and a model to forecast future consumer demand. Her final point was the need for a method to predict the impact of a project on future reliability.

In conclusion, Walter reiterated the need for a better model of statewide freight movement to inform where system improvements are most needed and beneficial. With better predictive models, future freight project benefits can be better articulated to support improved decisions.

PARTICIPANT DISCUSSION

The panelists were asked to provide specific examples of how data was used successfully to inform a decision. Walter stated that the CTC process is built on evaluation of data and that the projects that are selected are typically the ones that provide a detailed and easily understood summary of the supporting data and information. Caris stated that the Port Optimizer is currently focused on cargo moving through the port system and that the data is used to optimize port operations. Port operators expect that it will be used by other stakeholders as it is more integrated across operations.

Panelists were asked about the role that visualization plays for communicating resiliency measures to decision-makers. Caris responded that a primary role of the Port Optimizer is to communicate exactly what is happening where through graphs and tables. Walter reinforced the importance of interactive geographic information system (GIS) mapping systems and dashboards to present information to decision-makers.

KEY TAKEAWAYS

- The future of using data to inform decisions for freight planning will continue expanding as tools, like the Port Optimizer, and the requirements used by CTC become imbedded in agencies.
- Including stakeholders in both providing data and using the resulting information has the potential to improve decisions across the supply chain.
- Effectively communicating information is critical to the link between data and decisions.

BREAKOUT SESSION

Estimating E-Commerce

Understanding Parcel Activity

ALISON CONWAY
City College of New York, Moderator

CARLA TEJADA
City College of New York, Recorder

This session focused on recent advances in measuring e-commerce and residential delivery activity. Presenters discussed new data sources, data integration, and machine learning methods for demand analysis.

- **Modeling Household Online Shopping Demand in the United States: A Machine Learning Approach and Comparative Investigation Between 2009 and 2017,** Limon Barua, *University of Illinois at Chicago*.
- **Residential Loading Berth and Delivery Activity Study**, Tinotenda Jonga, Seth Contreras, and Fatemeh Ranaiefar, *Fehr and Peers*.
- **Approach to Estimate a Last-Mile E-Commerce Delivery Trip Matrix**, Christopher Lamm, Salah Momtaz, Brent Selby, and Amit Mondal, *Cambridge Systematics, Inc.*; Jakub Rowinski, *North Jersey Transportation Planning Authority*.

INTRODUCTION

Limon Barua, Tinotenda Jonga, and Christopher Lamm addressed parcel delivery and demand using different approaches and data that ranged from public to observational to private datasets. The first presentation used machine-learning methods to show the effect that demographic, economic, and land use variables have on e-commerce demand. This study is done at a national level using the U.S. National Household Travel Survey (NHTS) dataset. The second presenter showed a new approach to delivery data and how it can inform curb space sharing decisions to make the final 50-ft more efficient. Finally, the last presentation used private parcel data from the New Jersey area to develop a last-mile e-commerce delivery trip matrix that would help to forecast e-commerce demand, analyze demand predictors at the zip-code level, and analyze carrier activities.

SUMMARY OF INDIVIDUAL PRESENTATIONS

Barua focused on understanding e-commerce from the demand point of view. As stated by the presenter, the demand for e-commerce has quickly grown in the last 10 years. From 2018 to 2019, online transactions in the United States increased from $524 to $600 billion. Further growth was present during 2020 due to the COVID-19 pandemic. The way people order online may impact the distribution of personal travel trips and truck–van traffic expected for these deliveries. Hence, understanding e-commerce demand should help maintain and enhance freight transportation systems' performance and understand what drives e-commerce demand. A valuable input could be to inform transportation and urban freight planning and policymaking.

This research uses the 2009 and 2017 NHTS to develop machine-learning models to help predict e-commerce demand at a household level. The inputs for the model include socioeconomic, land use, and Internet-use variables that feed into a Gradient Boosting Machine (GBM) model. The GBM model works well with high-dimensional, mixed-type inputs of numerical and categorical variables. The model is invariant to transformations of input variables and insensitive to outliers, and can handle complex correlated variables. Furthermore, the authors used Shapley values (from game theory and machine learning) to quantify the importance of the features in 2009 and 2017, which ranks these features and allows for importance comparison on a year-over-year basis. Finally, they measured the existing relationships among all inputs using accumulated local effects (ALE).

Important results are the graphic representation of the relations between online shopping and several household characteristics. One of these graphics shows a positive relationship between the number of online orders and household income. This pattern depicts that a higher income equals more online shopping opportunities. One of the explanations the researchers give to this is that people with higher income could attribute more value to time and are insensitive to shipping costs. Another graphic exhibits a negative relation among online shopping and the average household age. Younger people seem to be less averse to online shopping, perhaps because they are more familiar with the technology. The team also indicated that older people may perceive more risks to online shopping, which are not experienced by younger generations.

Further analysis shows a difference in the impact of gas prices in 2009 and 2017. In 2009, the rise in gas prices negatively affects the number of online orders, while in 2017, the rise in gas prices has almost no effect. Finally, the results show that the number of online orders tends to be more significant in places where population density is either very low or very high.

Jonga gave initial findings linked to a loading and unloading berth analysis of residential sites in Seattle. This study was performed to provide information to the Seattle Department for Construction and Inspection (SDCI) on curb loading and unloading zones and alley activities. It also provided input to regulatory recommendations on land use code updates, mainly in the downtown and South Lake Union areas. The motivation for this study was a result of the dense use and increasing demand for curb space in the city, as well as the change that is occurring to Seattle's built environment as a result of e-commerce and other economic changes.

Researchers reached out to stakeholders that would be affected by or would benefit from potential changes that might result from the study and identified the following factors:

- Residents did not seem to have issues with alley use for loading and unloading activities. They felt that ongoing scheduling systems provided the correct performance and use of loading berths.
- Residential building managers did not agree and pointed out that use of loading berths for moving was difficult to predict since tenant move-ins or move-outs do not follow a regular pattern.
- Delivery companies did not find major operational issues related to double parking and traffic blocking. These companies affirmed that most goods and parcel deliveries are

conducted using on-street curb space in front of residential buildings, not in alleys.

- City planners, architects, and other city staff noted that an important design consideration is implementation of off-street building access for garbage trucks.

Based on this information, the team analyzed several alleys and loading docks for a 12-h window in the significant study areas. Findings from the study included the following:

- Approximately 70% of the vehicles in alleys were private cars.
- Only a few of all vehicles accessed the loading berths.
- Most commercial vehicle activity blocks the alley.
- In general, alley use ranged from 1 to 60 min.
- Primary alley uses included garbage collection and parcel delivery.
- More curb activity occurred downtown than in the South Union City neighborhood.
- Most curb activity is linked to private cars, mainly for food delivery.

Recommendations included constructing at least one loading berth for each residential building, providing an extra loading berth for buildings with more than 300 units, improving waste management operations, and incorporating dedicated parking within alleys.

Lamm provided the results of a study of freight forecasting and e-commerce delivery trip estimation for the Northern New Jersey area. This study was completed in 2020 using data from before the pandemic. The researchers used the New Jersey Transportation Planning Authority (NJTPA) Freight Forecasting Tool that processes commodity flow data to generate alternative forecasts. The primary inputs are (*a*) commodity flow data, (*b*) economic forecasts, (*c*) make-use tables, and (*d*) "what-if" scenarios. The primary outputs are (*a*) commodity flow forecasts, and (*b*) commodity trip tables for assignment on the model network. The tool does not include e-commerce specific data. As a result, the researchers quantified the e-commerce market demand and identified delivery vehicle patterns to create an e-commerce delivery trip table. The primary data source for determining the e-commerce demand was proprietary data from Rakuten Intelligence from 2018 to 2019. The team also analyzed the e-commerce percentage of retail sales trend in the area, corresponding demographics, and carrier information for the top four carriers in the region. A trip assignment table was created, adding 12 e-commerce categories to the original NJTPA Freight Forecasting Tool, informing new plans related to e-commerce.

Lamm pointed out that further research opportunities are dependent on data availability. Some research opportunities include:

- Expanding the market research to validate or enhance the vendor-source data;
- Tracking carrier data to confirm the validity of results;
- Incorporating information from more carriers; and
- Working with additional local and demographic data to increase the robustness of the model.

SUMMARY OF PARTICIPANT DISCUSSION

During the discussion, several participants pointed out the difficulty of working with nonrepresentative and incomplete datasets. A lot of the public data is aggregated at a national or state level and does not allow for analysis at the local level. For privately owned datasets, the challenge is to find a robust sample size that is representative of the area being studied and that efforts to validate the data typically require combining datasets.

One of the presenters emphasized the importance of outreach to stakeholders. Although this would involve substantial work, it is essential to confirm the sample size and the obtained results even when stakeholders have different views and competing activities.

Pre-processing e-commerce data is necessary and requires substantial time and resources because some e-commerce observations were not explicit, such as food delivery made by private automobiles.

Participants also pointed out the importance of localized analysis for e-commerce since current trends show a difference between urban and rural areas. It was noted that this type of analysis can only be performed with more comprehensive data. This is also necessary for considering product categories since they have different growth characteristics in e-commerce.

KEY TAKEAWAYS

- Use of e-commerce demand data could inform transportation and urban freight planning and policy making.
- Graphic representations of the relations between online shopping and household characteristics provides important insight into e-commerce demand.
- Useful e-commerce curb use information is now available from the results of the activity study performed in Seattle.
- E-commerce information required for regional freight planning is not readily accessible and requires pre-processing using data sources, such as retail sales and market characteristics.

BREAKOUT SESSION

Measuring Freight Flows for Roadway Infrastructure Decision-Making

NIKOLA IVANOV
Center for Advanced Transportation Technology, University of Maryland, Moderator

RYAN GRUBE
United States Department of Transportation, Recorder

In this session, presenters discussed freight data needs, sources, and analysis approaches for highway planning applications. The following presentations addressed performance measurement, resiliency planning, truck routing, and lane prioritization.

- **Using GPS Waypoint Data to Identify Through Urban Truck Movements for Consideration of Truck-Bypass Alternatives,** Allan Rutter, *Texas A&M Transportation Institute.*
- **Beyond Roadside Traffic Counters: Breaking Big Data's AADTT Barrier,** Christy Willoughby, Louis Yan, and Nancy Clarke, *StreetLight Data.*

INTRODUCTION

The purpose of this session was to hear from freight modelers and practitioners on measuring freight flows for roadway infrastructure decision-making. Allan Rutter discussed the value of GPS waypoint data in identifying truck congestion. Christy Willoughby discussed the value of using big data in detecting truck traffic patterns.

SUMMARIES OF INDIVIDUAL PRESENTATIONS

Building on a prior study conducted by the Texas Department of Transportation on mitigating truck congestion in the Texas Triangle (Dallas, Austin, San Antonio, and Houston), Rutter presented on research done by the Texas A&M Transportation Institute (TTI) on a follow-up study that specifically targeted truck congestion in the San Antonio area. In the earlier study, I-35 through downtown San Antonio was found to be the 20th most congested highway corridor in Texas. Researchers evaluated this highway with the goal of identifying alternative, ancillary highways in the San Antonio area that could be used instead. This led to the research question of "How much of the downtown truck congestion included through truck trips?"

Using GPS waypoint data purchased from INRIX, researchers were able to obtain the trip trajectories by plotting waypoints on a map. The analysis period was from February to May 2018, and contained spatial and temporal information. From this dataset, four groups of trip paths – including both commercial and passenger vehicles – were identified:

- External–External (EE): paths that begin and end outside the region (4%);
- Internal–External: paths that begin in the region and end outside (13%);
- External–Internal (EI): paths that begin outside the region and end inside (13%); and
- Internal–Internal (II): paths that begin and end inside the region (70%).

TTI researchers isolated this trip path data to the I-35 segment of downtown San Antonio (I-35D). After viewing the directionality (north or south bound) and time of day (morning versus afternoon), it was found that 17% to 19% of trips are through trips (i.e., an EE trip).

Observations include:

- Truck rerouting could have a positive effect on I-35D congestion. Modeling the impact of removing these through trips from I-35D is reserved for future research.
- Similar studies have occurred in other major metropolitan areas.

Next steps include:

- Establishing working groups to examine truck routing alternatives and implications.
- Working with local officials in the San Antonio area.

Willoughby demonstrated the importance of using big data to uncover truck traffic patterns, especially as it relates to annual average daily traffic (AADT). The goal of the StreetLight study was to create a model for 2019 average daily traffic for passenger vehicles (FHWA Class 1-3), single-unit trucks (FHWA Class 4-7), and combination-unit trucks (FHWA Class 8-13). The study measured data from across the spatiotemporal network by using data collected from permanent traffic sites (stationary sensors) and location-based sensors (GPS). Data included over 100 billion anonymized location records from smart phones and GPS navigation devices in cars and trucks, and data collected from over 1,800 permanent-count stations. Using RouteScience Technologies, the data points were transformed into aggregated travel patterns which were then overlaid onto the OpenStreetMap road network.

The model used machine learning, with the best results seen from the extreme gradient boosting method. The predictive features varied based on the type of AADT model:

- The passenger vehicle AADT model relied exclusively on location-based sensor features.
- The single-unit truck AADT model relied on location-based sensors and GPS trip counts, latitude and longitude, weather, trip length, and duration.
- The combination-unit truck AADT model relied mostly on tagged heavy commercial GPS trips.

Results of the model varied by type of vehicle. The combination-unit truck had the highest AADT at 12,000 while the other categories varied due to higher error and low counts.

Future work by the researchers involves expanding the project to include monthly AADTs and daily and hourly estimations. Also, plans include leveraging additional truck metrics, such as percentage of single-unit and combination-unit truck traffic, and truck vehicle miles traveled.

SUMMARY OF PARTICIPANT DISCUSSION

Individual participants were interested in the following areas of future research: transferability of each analysis to assist in their own studies; the economic impacts of truck congestion and supply chain disruptions; implications of truck rerouting; inquiries into the validation of the work; and methods to evaluate if models are accurately representing reality.

Following the theme of the workshop, individual participants also expressed interest in understanding how these findings can be put into practice and applied to real-world scenarios faced by decision-makers at multiple levels of government.

KEY TAKEAWAYS

- GPS data are fundamental to understanding vehicle behavior and are particularly important in understanding truck trips.
- Combining GPS data with other information provides improved information to support decision-making related to truck travel behavior.

Connecting Freight Infrastructure and the Economy

Data for Planning

CASEY WELLS
Texas Department of Transportation, Moderator

MAI LE
Eastern Transportation Coalition, Recorder

In this session, experts from leading transportation establishments discussed how their organizations use freight data for strategic and operational decision-making.

- **Presidio Freight and Trade Transportation Plan: Identifying Trade-Based, Economic Development Opportunities Using a Data-Driven Approach,** Blanca Serrano-Rivera, *Texas Department of Transportation*; Davonna Moore and Nicole McGrath, *CDM Smith.*
- **Evaluating Freight System Importance—Incorporating Economic Data To Evaluate Supply Chain Movements Captured Within Freight Data,** Jenna Goldberg, *EBP*; and Derek Cutler, *EBP US*.
- **Incorporating Economic Data Into Freight Planning: A Framework For Making Sense Out Of Noise,** Derek Cutler, *EBP US*; and Jenna Goldberg, *EBP*.
- **Data For Sub-County Region Freight Flow Estimation For Freight Planning,** Paul Bingham, *IHS Markit*; and Derek Cutler, *EBP US*.

INTRODUCTION

Four presentations considered data for trade-based economic development, economic data for evaluating the importance of freight movement, including economic data in freight planning, and developing sub-county freight flow estimations. Below is the brief summary of each presentation.

SUMMARIES OF INDIVIDUAL PRESENTATIONS

In the first presentation, Moore detailed a freight planning approach and methodology using a multi-part economic analysis called “complexities.” This analysis identified the connection between system functionality and opportunities with current and future regional freight movements, economic freight impacts, and key supply chains. Moore also emphasized the impacts of freight on interconnected economic and land use compatibility which support the current and future trends of freight flow in the study region. The outcome of this analysis was to identify trade-based economic development opportunities, target and maximize the inter-functionality of freight throughout the region to increase economic growth, and provide the opportunity to look at solutions that go beyond physical infrastructure.

In the second presentation, Goldberg considered evaluating supply chain activity using economic data in conjunction with freight movements and focused on specific assets, such as highway and rail corridors, to understand what is important to the industry. It develops a comprehensive relationship view of freight, infrastructure, and the state economy. Three sector models were developed that categorize industries within the supply chain:

- Extractive industries, which are industries that extract raw materials from the ground or which harvest products from farms;
- Primary manufacturing, which are industries that process seven basic forms, using the inter-immediate products; and
- Secondary manufacturing, which are industries that take intermediate products to make complex or finished goods.

The project defined six supply chains and the presentation provided four examples of the analysis in Michigan. Goldberg also explained how to connect the freight economic data, using the *Implan* economic model to identify where industries are and what their consumers produce, which then is linked to freight flow data based on direction of the flow.

In the third presentation, Cutler covered the dependence of freight and how to incorporate freight data into economic models. He then presented two applications from the Illinois Marine Transportation and Kansas Turnpike Authority demonstrating the broader use case for the models. Both applications showed the direct impact of cargo movement on the Illinois Marine Transportation and Kansas Turnpike by industry. Cutler showed the applications of the freight economy models which included the process to enable broader applications, such as:

- Industry dynamics,
- Scenario and forecasting, and
- Enhanced spatial modeling.

Bingham discussed how to use data for subcounty region freight flow estimation for freight planning in the fourth presentation. The purpose of this work is to consider three problems:

- Traditional metropolitan area/city truck freight data sources are samples and have no data on the goods carried;
- MPOs, counties and sub-county agencies need freight modeling capabilities;
- Commodity flow data are not geographically detailed.

To support local and regional planning activities, the solutions to help resolving these problems are to construct a logical modeling framework to support detailed disaggregation and apply data that includes the economic behaviors behind the freight flows. Bingham also discussed how to interpret freight flows and sources of data. He reiterated that the most important thing is the relationship between freight and the economy. He concluded that the freight flow analysis presented enables MPOs and regional agencies to customize data, preserve logical behavior of freight flows, and is applicable at any spatial scale with a differentiable level of detail.

SUMMARY OF PARTICIPANT DISCUSSION

A participant asked about best practices for agency administration to evaluate freight data for programming. Cutler indicated that it depends on the context of the regional planning, long-term growth, understanding how goods are sourced and shipped.

Another participant asked about features of data requirements to measure economic impacts of resiliency in the long-term freight plan. Bingham responded that when considering urban freight resiliency, the important thing is the data supports the planning. The freight data community are considering local freight movement using GPS data, which does not include information about goods on board, and needs the ability to link between those movements and what is in the vehicle to understand what is actually in the region. This includes understanding the mix and value of goods on board. When considering economic impact, decision-makers would need a way of matching this to the level of economic geography to determine wider impact benefits.

PANEL

Procuring New and Innovative Data

DONALD LUDLOW
CPCS, Moderator

SAM HISCOCKS
Iowa Department of Transportation, Recorder

In this session, state and regional planning agencies discussed experiences, best practices, and challenges for procuring new and innovative freight data.

INTRODUCTION

The province of data procurement is constantly changing. There are more sources today than there have ever been, making it more difficult for agencies to make decisions on which data is needed for what price to solve existing and emerging problems. This panel focused on experiences, best practices, and challenges for procuring new and innovative freight data.

Donald Ludlow moderated a discussion with Robyn Bancroft of Ohio–Kentucky–Indiana Regional Council of Governments (OKI) and Tom Murtha of Chicago Metropolitan Agency for Planning (CMAP). Bancroft presented on the process used by OKI to identify potential datasets to procure and use, while Murtha provided an overview of specific new datasets CMAP has been working with.

Robyn Bancroft, *Ohio-Kentucky-Indiana Regional Council of Governments*

Bancroft summarized OKI's need for new data sources to comprehensively evaluate the region's multimodal freight system for the regional freight plan and provided a summary of the process used to procure the necessary data. To identify potential sources, staff documented which data could be obtained, which questions the data could answer, and which questions the data did not answer. OKI generated a list of over 80 potential datasets which were evaluated by cost, availability, and value. Staff ultimately selected 23 separate datasets for analysis. The datasets now being used will:

- Help generate a freight infrastructure project list that strategically targets investments;
- Reveal recommendations for special freight studies;
- Provide recommendations to be used in the next metropolitan transportation plan;
- Support the development of potential new, quantifiable freight performance measures;

and

- Help identify information to more accurately assess the value of funding proposals.

Three new datasets were specifically highlighted by Bancroft that OKI is now incorporating. First, OKI is relying on regional freight partners, including the Cincinnati–

Northern Kentucky International Airport, Northern Kentucky University, Central Ohio River Business Association, and the Ports of Cincinnati and Northern Kentucky to provide data on commodity movement, capacities, and bottlenecks. Second, automatic identification system (AIS) data is being used to track vessels on the Ohio River to determine usage. Third, OKI is using a new approach for evaluating highway-railroad crossings by analyzing the importance of the crossing to the public, total delay to users, redundancy, and safety.

Tom Murtha, *Chicago Metropolitan Agency for Planning*

Murtha focused on the new freight datasets that CMAP has been working with recently that stemmed primarily from two major efforts, among others. First, the development of a comprehensive regional plan, which focuses on inclusive growth, resilience, and prioritized investment to address transportation, climate, and the economy, emphasized the importance of reliable data and analysis. Secondly, the CREATE program, which focuses on rail system efficiency and limiting community impacts by reducing conflicts between highway and rail, revealed data gaps in the information being used and identified difficulties with working with different data sources.

CMAP responded by developing freight models supported by several data purchases. Regional freight system planning is supported through long-term continual data purchases (e.g., business establishment data and commercial property data), industry and government partnerships (e.g., intermodal activity data from railroads, vehicle classification data from Illinois DOT, and project development data from multiple sources), and new and innovative data procurements. Murtha provided an overview of four new and innovative data procurements:

1. Delivery, service employer, and driver surveys from Malatest;
2. Waypoint and origin/destination data from INRIX;
3. Waypoint data from the American Transportation Research Institute (ATRI); and
4. Intermodal volumes from the Intermodal Association of North America (IANA).

As part of freight model development, survey data from delivery employers, service employers, and drivers is being gathered to strengthen model inputs that typically end at distribution centers while ignoring the recent increase in delivery and service travel. INRIX data was purchased for calibration of the model. It was also used to supplement waypoint data from ATRI by providing more information on single-unit and delivery trucks to help better understand the entire transportation system. Waypoint data from ATRI were purchased to provide origin-destination data for larger combination trucks and how they behave in the region. Finally, intermodal volumes from IANA were acquired to understand the total volume of goods moving through the intermodal system and the equipment being used.

KEY TAKEAWAYS

- Data procurement is a critical topic for future freight analysis efforts.
- Continuing to explore different data sources is important, as well as better equipping agencies to analyze information and implement results for improved freight operations.
- With the extent of resources available today, agencies need to be able to evaluate how to choose which data sources are the best fit for their needs. This will require a clear definition of the problems to be addressed and a comprehensive evaluation of potential sources using criteria such as cost, availability, and value.

SPOTLIGHT

New Data and Analytical Approaches for Addressing Critical Freight Issues

Carl Vesper
Post Holdings, Moderator

Brittney Gick
Texas A&M Transportation Institute, Recorder

This session highlighted innovative data and analytical approaches to inform freight planning and investment decisions. The following presentations focused on statewide, urban, and port applications.

- **Developing Freight Data into What-If Scenario Analyses for Freight Planning: Location-Based Disruptors for Texas Freight Movements,** Dan Forbush and Kaveh Shabani, *Cambridge Systematics, Inc*.
- **A Policy and Infrastructure Evaluation Model of Commodity Flows Through Inland Waterway Ports,** Sanjeev Bhurtyal, Sarah Hernandez, Sandra Eksioglu, and Manzi Yves, *University of Arkansas*.
- **Analysis of Urban Freight Flows and Last-Mile, Retail Goods Movement Over Time Using Inference with GPS Trajectory Data,** Olcay Sahin and Monique Stinson, *Argonne National Laboratory*.

INTRODUCTION

The objective of the session was to highlight innovative data and analytical approaches to inform freight planning and investment decisions. Dan Forbush, Sanjeev Bhurtyal, and Olcay Sahin shared insights into their current efforts at addressing freight flow shifts and disruptions. Data integrity and access to relevant data is an important element in developing adequate freight flow models that allows transportation planners and agencies to better prepare for freight shifts and disruptions at all levels.

SUMMARIES OF INDIVIDUAL PRESENTATIONS

Forbush presented on his work with developing Texas DOT's Freight Forecasting and Scenario Planning (FFSP) tool, which looks at location-based disruptor scenarios when freight flow origins shift from one location to another. Texas DOT's FFSP tool uses data from Transearch, Moody's industry forecasts, and the Bureau of Economic Analysis (BEA) National Make/Use Tables, and allows users to input forecast scenarios, forecast start and end years, study area geography, and set disruptor scenario parameters to project new forecasted commodity flows. For this presentation, Forbush used the following policy question as a guide to highlight the tool's capabilities: "What if there is a change in manufacturing trends where commodity tons are produced domestically in the U.S. and fewer tons are imported?"

The FFSP tool's methodology has three stages. In the first stage, the pre-disruptor import and domestic production freight flows are identified. In the second stage and, based on user inputs, the magnitude of the change from the disruption is determined. In the third stage, based on differential factors, the change in value is calculated.

Using the sample scenario, where domestic production of commodities increased by 1% and commodity imports decreased by 1% each year between 2020 and 2040, the tool indicated that by 2040, a 22% increase in tons would occur with a 19% increase in value of domestic production and an 11% decrease in import tons and value of the commodity. The FFSP tool also provides information on the impacts of the freight flow shift by mode and direction. The FFSP uses a simple, targeted methodology to estimate the change in tons and value of commodities from one origin to another. The process is zero-sum and can be tailored by the user to address various policy questions and analyze impacts on a variety of commodity flow measures (e.g., mode or direction).

Bhurtyal presented on his current research addressing the data limitations that inland waterway ports face. These limitations include the proprietary nature of data collected by port operators since BTS only publishing port performance freight statistics for the top 25 ports, and Commodity Flow Survey (CFS) and Waterborne Commerce Statistics being limited in spatial or temporal disaggregation. This research sought to address how inland waterway port operators could optimize investment decisions and identify how these investment scenarios change based on external impacts.

To address the research question, a two-stage stochastic model was formulated to evaluate the first and second stage decision variables. The objective of the first stage is to minimize the cost of port infrastructure (e.g., equipment and storage facilities). The objective of the second stage is to minimize shipping costs. These shipping costs include: 1) costs between counties via rail and truck, 2) ports via barges, and 3) counties and ports via rail and truck, 4) inventory holding costs at ports, and 5) penalty costs for unmet demand. The demand scenario was based on the amount of monthly demand for a commodity in a county. The base scenario commodity demand was obtained from statewide commodity flows using Transearch and the CFS data. Eight additional scenarios were modeled based on data obtained from historical Lock Performance Monitoring System (LPMS) data. The solution algorithm in Figure 3 is applied to solve for the main problem. As the main problem is solved sub-problems are also solved.

This research focuses on a case study of a section of the McClellan–Kerr Arkansas River Navigation System, which included 75 counties, 30 ports, 11 commodity groups, and nine different scenarios over a 12-month time period. The research is ongoing, and the model results still need to be evaluated, so the actual results of the model were not shared during the workshop.

Sahin presented on his work addressing gaps in the data related to urban freight flows and how truck travel patterns change over time. His work included a longitudinal analysis of truck flows. This research used a data-driven approach centered on analyzing 30 million origin–destination (OD) trip patterns for light-duty vehicles (LDV), medium-duty trucks (MDT), and heavy-duty trucks (HDT) over a combined 4-month timeframe (Spring and Winter 2015, Spring and Winter 2017) in the Chicago area using INRIX GPS trajectory data. The trip start and end points were snapped to a tolerance of 110 m (365 ft), so the exact locations where the trips were

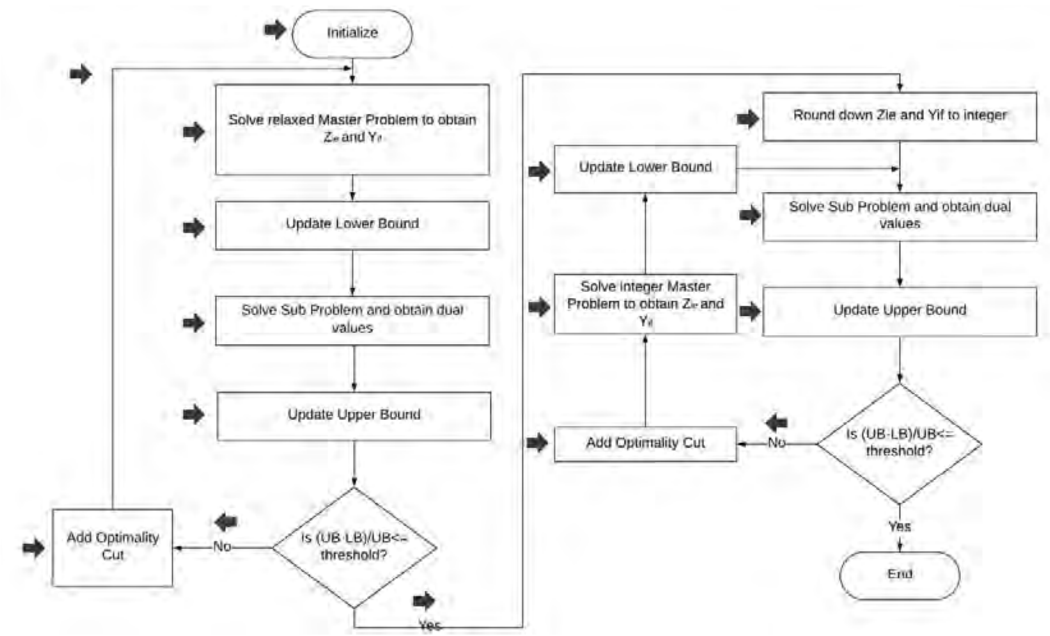


FIGURE 3 Modified benders decomposition solution algorithm.

beginning and ending are unknown. The GPS samples were typically taken every 1 to 5 min. The research used CMAP land use data to identify where the vehicles were going to and from in the area. The research used an empirical analysis to evaluate metropolitan MDT and HDT truck patterns along with how they have changed over time, including how truck trips shifted over time in response to changes in e-commerce.

The results showed that LDV trips peaked in the a.m. rush hour and remained fairly steady throughout the day, with another peak in the p.m. rush hour. MDT trips peaked much earlier in the day, with the heaviest concentration around 6 a.m., peaking around noon and then tapering off throughout the afternoon. HDT trips were heavily concentrated overnight. LDV and MDT trips had high percentages of trips beginning and ending in residential areas, while HDT trips had higher percentages of trips beginning and ending in industrial, commercial, and transportation, communication, utilities (TCU), and waste areas, suggesting that HDTs are primarily used for industrial purposes and much less for residential goods delivery. Comparing 2015 and 2017, the share of MDT trips going to and from urban retail locations declined, which possibly suggested that the rise in e-commerce created a greater decline in small shopping areas compared to larger ones. While the role that MDTs played in residential deliveries was equally important in both 2015 and 2017, the analysis showed that in 2017 MDTs expanded to more trips to other urban freight operations (e.g., warehousing and distribution centers, truck terminals, airports, intermodal yards, and other industrial areas).

The research gave insight into the nature of trips that each vehicle type performed, yielding interesting insights into retail goods distribution. The trips by land use across the four time periods suggest that warehousing employment and truck trips to land uses like truck terminals are both on the rise, which is consistent with the general rise in e-commerce. The results of the data-driven analysis could be used to provide insight into vehicle trips by vehicle type, refine freight planning tools, inform the understanding of the potential effects of e-commerce as they relate to land use, as well as inform the design of future studies.

SUMMARY OF PARTICIPANT DISCUSSION

Questions were raised about the use of the FFSP tool and the impact of freight flow shifts, as well as how the tool can address economic engagement. The FFSP tool has a function that allows exporting of the custom commodity flow forecast to the state's travel demand model, the Statewide Analysis Model (SAM), so the tool can show the impacts of freight flow shifts on routing and highway flows. Economic engagement could be monitored by flows between Texas and distant markets based on the Transearch and Moody's data used in the FFSP tool; however, it was not done for the scenario presented at the workshop.

In response to a question about who the end user of the data in the second presentation is and what they are able to do with the data, the model can help port companies better understand how to make investment decisions, such as what and where to invest. It can also help USACE prioritize actions in the case of disasters by showing what could happen. The methodology was specifically developed for inland waterway ports, not coastal ports.

Sources of the data and potential extension of the analysis in the third presentation were provided. The CMAP land use data is publicly available and 2015 data for northeast Illinois can be found at https://datahub.cmap.illinois.gov/dataset/land-use-inventory-for-northeast-illinois-2015. Having access to data from places like Amazon and Grubhub could further enhance the understanding of food and grocery delivery by vehicle type and could be applied in the models. Data used in this study is currently only available for 2015 and 2017, but as future data becomes available the study could be extended.

A number of the presenters faced challenges when it came to selecting the data source. Some challenges expressed by the presenters related to selecting data sources included the opaqueness of proprietary data (i.e., lack of clarity on how data was developed) and finding the correct commodity data that was needed. There was a general discussion on how freight disruption modeling compares to freight resiliency modeling. Disruption models focus more on freight shifts and resiliency models would require more travel demand modeling, as critical bottlenecks and routes must be identified to understand the capacity of the network.

KEY TAKEAWAYS

The different analytical approaches discussed in the session provided valuable insights into understanding freight impacts and scenarios. The following key takeaways were provided by each presenter:

- It is important to do the most with the information available at the time and plan for what the potential impacts may be, while keeping up with any real disruptions.
- The next step in the inland waterways research will be to analyze the model results and possibly try to model impacts of disruptions at coastal ports to determine if the model is applicable for coastal ports as well.
- E-commerce and last-mile deliveries are of great concern in metropolitan areas and data shared from private companies (i.e., Amazon and Grubhub) would greatly improve the models.

BREAKOUT SESSION

Freight Demand Modeling Innovations

MONIQUE STINSON
Argonne National Laboratory, Moderator

MAJBAH UDDIN
Oak Ridge National Laboratory, Recorder

In this session, presenters discussed new freight modeling approaches. Presentations covered methods for local, state, and national demand estimation, as well as an innovative data platform approach for improving operational efficiency through information sharing.

- **Rhode Island Truck Trips Origin and Destination Analysis,** Josh O'Neill and Benjamin Jacobs, *Rhode Island Division of Statewide Planning.*
- **Multi-Commodity Air Freight Origin-Destination Demand Analysis,** Rajesh Paleti, Amy Moore, and Michael Hilliard, *Oak Ridge National Laboratory*; Lindy Williams and Caleb Phillips, *National Renewable Energy Laboratory*.
- **New Method For Truck Trip Assignment For Statewide, Interregional, And Nationwide Models,** Howard Slavin, Vince Bernardin, Jonathan Brandon, Jim Lam, Andres Rabinowicz, and Bruce Spear, *Caliper Corporation*.
- **Reducing Road Freight Transport Costs And Emissions By Means Of A Data Platform For The Exchange Of Time Windows In A Pick-Up And Delivery Process,** Felix Roeper and Ralf Elbert, *Technical University Darmstadt*.

INTRODUCTION

Freight demand modeling is critical to planning for national, state, and regional movements of multimodal freight. This session presented examples of new freight modeling approaches. In particular, it covered methods for demand estimation at different geographical regions as well as an innovative data platform approach for improving operational efficiency through information sharing.

SUMMARY OF INDIVIDUAL PRESENTATIONS

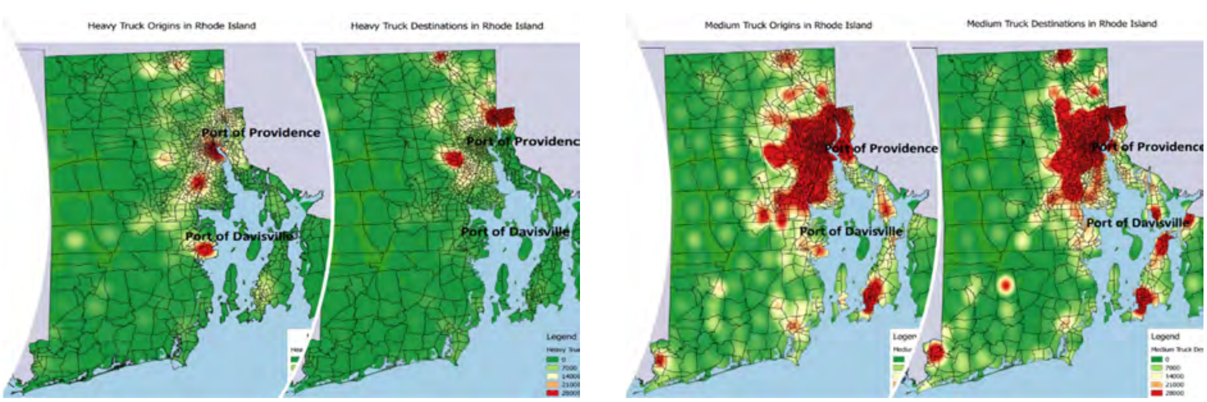
O'Neill indicated that Rhode Island's Division of Statewide Planning began using the Regional Integrated Transportation Information System (RITIS) tool in 2019. The tool was used to analyze queries about truck travel patterns, and location and density of freight generating companies in the state. The major queries were: (*a*) Where and how were the freight origins and destinations concentrated? (*b*) What were the differences between heavy truck generators and medium truck generators? (*c*) What communities hosted the highest truck origin volume? and (*d*) Which communities should be targeted for further study?

This first step in the study was an analysis of truck origins and destinations for the Port of Providence. However, it was determined that statewide origins and destinations and a Heat Mapping Perspective was necessary to determine the locations of the top freight generators and destination patterns within and outside of Rhode Island. Available data included transportation analysis zone (TAZ) geometry and trip analytics, including origin and destination by TAZ, and medium and heavy truck use. Figure 4 shows the generated heat maps with (*a*) heavy truck ODs and (*b*) the same for medium trucks.

One finding was that in addition to Amazon, FedEx, and UPS, the United States Postal Service (USPS) is a major medium truck freight generator. The results of the truck OD patterns within the state allowed the research team to focus on future community freight planning efforts. The next step is for the team to continue outreach and coordination with planners in local communities for continued dialogue about issues such as truck parking, truck idling, air quality and noise, and safety.

Paleti considered that freight demand patterns can vary across modes and different industrial sectors, depending on larger socio-political and economic factors. For example, during the initial days of the COVID-19 pandemic, surging demand in the grocery, household goods, and medical supply industries occurred relative to other sectors. For that reason, a need arose for industry- or commodity-specific air freight demand forecasting methods that are responsive to socioeconomic factors to inform transportation decision-making and supply chain management.

This study used a modeling framework involving three models: selection model, stochastic frontier model and demand composition model. The selection model determines whether an OD pair have non-zero demand. The stochastic frontier model determines the maximum total demand across all commodities for each OD pair. The demand composition model determines the composition of the demand by commodity type. The primary data source for the study was the latest version of the Freight Analysis Framework (FAF) data, which was supplemented by several other data sources, including socio-demographics (e.g., total population, population density, and median household income), transportation infrastructure (e.g., number of airports, number of bridges, and total route miles of freight railroad), and employment by industry types.



(a) **(b)**

FIGURE 4 Statewide Heat Map: (a) heavy truck OD and (b) medium truck OD.

The study also conducted a scenario analysis for the Dallas–Fort Worth International Airport. The results were evaluated in using a practical application, which considered shifts in the characteristics of air freight at the airport and surrounding area. It found that a 10% increase in total employment led to over 50% increase in the average demand across all commodity types. It also found that a 10% increase in median age of the population led to more than 50% reduction in the freight demand.

Slavin considered truck assignments. Conventional long-haul truck trip assignments are performed using methods appropriate for passenger car traffic. The two principal methods are shortest path preload and multi-class equilibrium assignment, each with passenger car equivalents. However, multiple paths are used by each class of truck, and equilibrium methods neither apply nor do they use the correct routes.

This study used an automated heuristic that finds distinct alternatives between origins and destinations. The heuristic favors interstate routes that are most commonly used. A discrete route choice model was also used to allocate the truck trips to different routes based on route attributes, such as travel time and tolls. Both the routes and the trip volumes were compared and validated with empirical data, such as those made available by the ATRI. The resulting routes are viewed spatially and edited, added, or deleted. Figure 5 provides a comparison of FAF version 5 (FAF5) routes with routes generated using truck GPS data (i.e., ATRI).

The modeling method has been implemented in TransCAD and is supported by relevant editing, visualization and query tools. A full-featured query tool was developed to provide an easy means of accessing the outputs of large models with thousands of OD pairs, millions of paths, and numerous classes of traffic. As a future development, the researchers are considering embedding total traffic flows inside the equilibrium model.

FIGURE 5 Comparison of FAF5 routes with ATRI routes.

Roeper focused on web-based time window management systems (TWMS), which are operated by external providers and are used to coordinate loading docks to increase capacity utilization of warehouse operations. At an operational level, time windows (TWs) may be difficult to meet due to traffic congestions or other disruptions. The unfavorable allocation of TWs to freight forwarders leads to unnecessarily long tours. By exchanging TWs, freight forwarders can improve their tours. With an increasing number of participants, the possibility to exchange TW increases but also adds complexity.

This study answered two research questions: 1) which data are needed to facilitate TW exchanges in a pick-up and delivery process using an auction-based mechanism; and 2) how should the structure of a cross-provider data platform be designed to enable TW exchanges across providers? To identify further requirements, a focus group interview was conducted with nine different companies. Four additional interviews were held with persons with relevant expertise from the practice to validate and verify the modeled planning process.

The resulting cross-provider data platform leveraged the allocation efficiency of the auction-based mechanism. The platform offers the opportunity to connect different TWMS providers and freight forwarders, as illustrated in Figure 6. As more TWMS providers and forwarders participate on the platform, more opportunities arise for TW exchanges. For the exchange of TW in an auction-based mechanism, all data about the variables that influence handling time are relevant. Although relevant assumptions and constraints were addressed and the basic features of the platform were developed, the structure of the platform needs to be verified and validated.

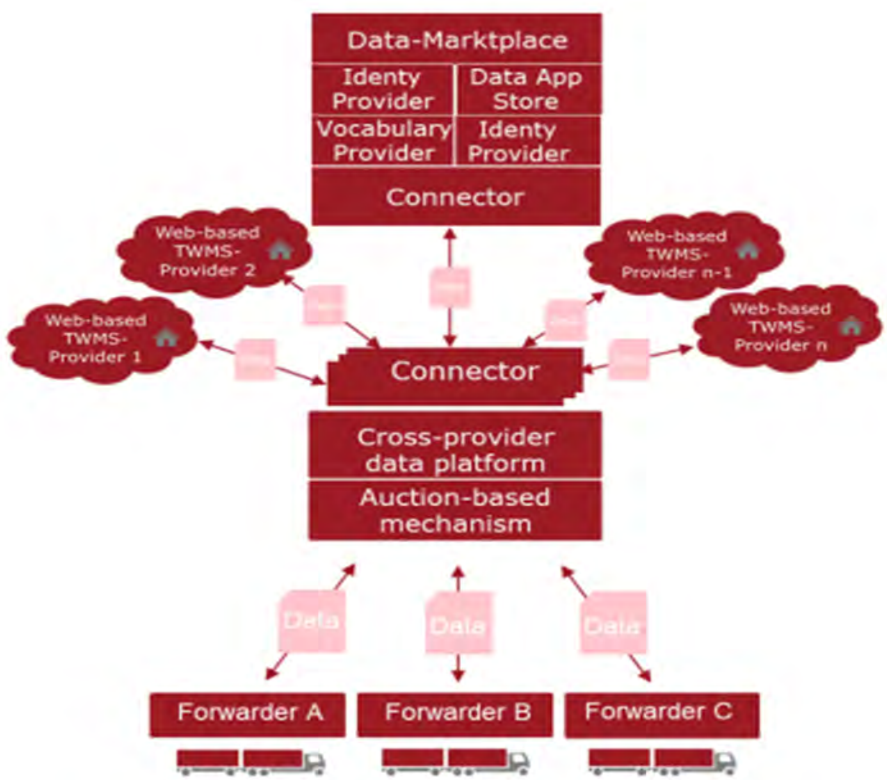


FIGURE 6 Structure of a cross-provider data platform.

SUMMARY OF PARTICIPANT DISCUSSION

One presenter commented that they are exploring a truck TW assignment system at a port for heating oil and propane deliveries and scrap metal. Without a better understanding of the financial impact to truck drivers and carriers, further consideration is unlikely. Conflicts with peak morning congestion on the interstate highway are also a concern. Another presenter responded that the use of TW is a requirement for truck carriers by warehouse operators and ports–terminals, and that associated costs are covered by the truck carriers.

O'Neill indicated that sensitivity to tolls is going to play a larger role in planning in future years. Rhode Island implemented truck tolls on highways, which began just before the pandemic. It saw an increase in truck volume using local roads to avoid tolling. The impact of this on local roads will need more analysis in the coming years.

A participant asked about consideration of changes in airport worker productivity resulting from automation or improved skills in commodity flow exchange and whether this would be most significant for high-value goods transportation. At the present time, no data are available related to changes in worker productivity.

Another participant asked whether the new truck route modeling tool improved predictions about pavement deterioration. The presenter answered that the model actually captures the routes that are used by heavy trucks. In addition, the model would permit a better estimate of pavement damage and the need for investment. The research team also observed that a lot of the paths are not the simple shortest path. This may be due to multiple drop-off points. States can also obtain better estimates of truck volumes on interstate segments using this modeling tool.

A participant asked what adjustments are taken to correct bias with the sample before use in analysis–validation for probe data. One presenter answered that they mostly were interested in whether the heuristic to enumerate routes could be validated in some meaningful way, and they were pleasantly surprised that the case study that they looked at did a remarkably good job.

KEY TAKEAWAYS

Research has been done and continues on the following innovations in freight demand modeling:

- Using probe data to analyze truck trip ODs;
- Modeling framework to predict total demand as well as the relative composition of different types of commodities using publicly available data;
- A truck trip assignment method that can find alternative routes that can be compared and validated with empirical data; and
- Exchanging TW on a cross-provider data platform is under development and could lead to lower costs and emissions for freight forwarders.

BREAKOUT SESSION

Ports and Ports of Entry

Measuring Freight Activity in Multimodal and Disrupted Environments

ERIC FREDERICKS
California Department of Transportation (Caltrans), Moderator

ANDRE TOK
University of California Irvine, Recorder

In this session, presenters discussed innovative data sources and data integration approaches for measuring freight activity for ports, inland waterways, and border crossings.

- **Assigning a Commodity Dimension to AIS Data: Disaggregated Freight Flow on an Inland Waterway Network,** Magdalena Asborno, *U.S. Army Corps of Engineers*; Sarah Hernandez, *University of Arkansas.*
- **Data Integration and Analysis to Estimate Cross-Border Freight Fluidity,** Mario Monsreal, *TTI*.
- **Using Multiple Data Sources to Analyze the Impact of the Expanded Panama Canal on Texas Ports and Landside Access,** Katherine Turnbull and Kale Driemeier, *TTI*.
- **Multi-Perspective Analysis of Pandemic Impacts on U.S. Container Trade,** Daniel Smith, *The Tioga Group*; Paul Bingham, *IHS Markit*; and Daniel Hackett, *Hackett Associates*.

INTRODUCTION

The session includes presentations that described the use of freight-related datasets and methods to model freight flow on inland waterways, estimate truck activity across the Mexico–U.S. border crossings, assess the impacts of the Panama Canal expansion and investigate benefits of lead indicators in predicting international container freight demand. A common theme amongst these presentations is their use of multiple data sources in their analysis.

SUMMARIES OF INDIVIDUAL PRESENTATIONS

Asborno presented her study which investigated the combination of AIS, the USACE LPMS and truck GPS data to estimate commodity flow at the trip level in the McClellan–Kerr Arkansas River Navigation System. It provides the framework for building a disaggregated commodity-based maritime freight database, which is used to build synthetic populations for activity-based travel demand. This presentation summarized the datasets and described the required processing for use. A data challenge was determining the number of barges pushed per tug using a stochastic model that developed scenarios of trip capacity.

Monsreal described his study which estimated the travel time of commercial vehicles moving between seven border crossings between Mexico and the United States. The study decomposed a cross-border movement into five component segments using two data sources: INRIX and HERE. The methodology involved data sampling, pre-processing, travel time estimation and data merging. This study resulted in development of the End-to-End Border Crossing Interactive Tool, which provides a visualization of performance of each border crossing segment by temporal aggregation.

Turnbull described her study which assessed impacts of the expanded Panama Canal through the analysis of data from the Panama Canal Authority, U.S. Department of Energy, USA Trade Online (U.S. Census Bureau), Texas ports from PIERS, stakeholder interviews, and other sources. Insights of impacts were given on specific commodities, vessel type, modes, and various transportation infrastructure projects.

Smith described the analysis in his study of pandemic impacts on container freight. Several years of data were analyzed to investigate the trends in container and freight demand patterns across various commodities. Several factors affecting freight capacity at the ports were described. The concept of leading economic indicators were adopted to analyze leading demand indicators for imports. Relationships between inventory, sales and imports were explored.

SUMMARY OF PARTICIPANT DISCUSSION

Freight analysis involves a combination of private and public data sources. Private and public sectors need to work together to ensure support facilities are available. However, data always indicates what has happened in the past, so the challenge is to develop better leading indicators that can help better identify and address freight congestion.

Insights on commodity movement for ocean going vessels poses a distinct challenge compared with inland waterways. While AIS data is available for both, commodity data is only available at adequate resolution for inland waterways through the LPMS. No equivalent dataset is available for ocean-going vessels.

BREAKOUT SESSION

The Truck Parking Problem

Measuring Activity to Inform Solutions

DAN MURRAY
American Transportation Research Institute, Moderator

CLAIRE EVANS
American Transportation Research Institute, Recorder

This speed session highlighted recent applications of truck probe data to measure and assess long-haul truck parking challenges. The following presentations were followed by a roundtable discussion with presenters.

- **Using Existing Data Sources to Determine Unauthorized Truck Parking Locations,** Tom McQueen, *Georgia Department of Transportation.*
- **Statewide Analysis of Truck GPS Data to Understand Truck Parking Utilization,** Makarand Gawade, *HDR*; Joel Worrell, *Florida Department of Transportation.*
- **Using ATRI Truck GPS Data for Freight Planning in a Rural Area,** Roger Schiller and Chiru Bhamidipati, *CDM Smith*; Rebecca Reyes and Blanca Serrano-Rivera, *Texas Department of Transportation.*
- **Expansion of GPS Data to Estimate Truck Parking Demand,** Sebastian Guerrero, *WSP*; Bridget Wieghart, Joseph Bryan, Rebecca Knudson, *Oregon Department of Transportation.*

INTRODUCTION

This session highlighted the use of truck probe data to address truck parking challenges. It consisted of four nine-minute presentations and a discussion at the end. A common theme was the necessity for forecasting supply and demand of truck parking in urban and rural regions.

SUMMARIES OF INDIVIDUAL PRESENTATIONS

McQueen discussed the analysis that was done in response to a review of published estimates of unauthorized truck parking in Georgia. Readily available data were used to identify and characterize unauthorized truck parking locations to gain additional insight into future truck parking demand and supply. The data included FMCSA parked truck crash GPS data, Georgia DOT data for parked truck crashes, FHWA Transportation Management Center (TMC) for truck speeds, and Georgia State Patrol truck parking warning and citation records for 2019. State land use plans and Google Street View were used to categorize and verify land use.

Three categories of truck parking locations were identified: 1) private, 2) public, and 3) unauthorized truck parking. Results showed unauthorized truck parking is happening on

Interstates and on a significant number of rural roads. This was not unexpected since many rural regions in Georgia have high density industrial land use areas (e.g., warehouses, distribution, and manufacturing) with high truck volumes. The researcher recommended fully vetting data to reveal new uses and maximizing use of readily available data from state and federal agencies.

Gawade considered obstacles in addressing the truck parking issue in Florida through identifying where parking is an issue, when issues occur, and quantifying how much additional parking is needed. GPS data from ATRI, parking supply data, parcel polygon data from Florida Department of Revenue, truck count data from Florida DOT, and Truck Parking Availability System (TPAS) data from Florida DOT were used. Aggregated data were sent to Florida DOT stakeholders for verification and recommendations, which were incorporated into an updated dataset. The data were then filtered and manipulated to obtain observations of trucks stopped for more than one hour. These observations were overlaid on spatial data to determine the number of trucks per location every hour. Florida's TPAS data validated the findings for 13 public truck parking locations in Florida. This research quantified both space and facility demand.

Schiller focused on the expansion plan for freight corridors, I-10 and I-20, in Presidio and Ojinaga Counties in Texas. The goal was increasing safety and economic development in this rural area. One consideration for increasing safety along these corridors is addressing inadequate truck parking. Data sources included the 2020 Texas DOT Statewide Truck Parking Study, 2019 ATRI GPS data, the updated Texas DOT truck stops inventory, and 2015 to 2019 crash data from the Texas DOT Crash Record Information System (CRIS).

GPS records were processed using *R* to analyze weekday and weekend patterns in truck stop durations. These were then compared with truck stop inventory data to identify truck parking demand by day of the week. From this analysis, 8% to 9% of demand was long-term (overnight) parking with approximately equal demand on weekdays and weekends. Weekday demand was markedly higher for stops less than 15 min in duration. Demand exceeded supply along the I-10 and I-20 corridors, particularly where I-10 and US-285 intersect. Little to no demand was found along the border of the two counties where truck parking supply is nonexistent. Areas identified as high demand were compared with records for parked vehicle truck-involved crashes.

Researchers concluded that GPS data works well for accurately identifying locations and durations of truck stops in rural areas. They also emphasized the necessity for evaluating impacts of insufficient truck parking supply as it relates to causation and measures of driver fatigue and parking safety.

Guerrero summarized research on estimating truck parking demand and forecasted deficiencies for 2040 using new methods for expanding GPS data. These methods included integrating ground observations with parking activity, instead of road volumes, to develop improved expansion factors. Expansion factors and their uncertainties were estimated using regression analysis using GPS sample data from ATRI and estimated demand for all rest areas and truck stops in Oregon. Video observations were obtained from cameras installed at nine rest areas and two truck stops, and through the truck parking app (Park my Truck) at 11 truck stops. Parking was recorded every 4 h for 1 week in July 2019 concurrent with GPS data. Modeled expansion factors were validated using observed occupancy data. The model produced a wide range of expansion factors. One observation was that urban areas had higher factors than rural areas. Another consideration was possible facility preference based on truck type.

SUMMARY OF PARTICIPANT DISCUSSION

Discussion began with a question on whether safety of unauthorized parking is being addressed in current research, and whether data for safe versus unsafe truck parking is used. It was noted that staged trucks, though technically unauthorized, are often safely parked. The challenges of identifying safe and unsafe truck parking data was underscored as an ongoing issue for researchers.

A second topic was urban versus rural truck parking issues. Long-term planning was emphasized as essential for allocating sufficient urban land use for truck parking demand. Additionally, it was noted that transformations in urban land use planning has resulted in more truck parking in rural areas. Parking in rural areas outside cities was pointed out as common practice because it is safer, while allowing drivers to meet delivery windows.

KEY TAKEAWAYS

- Fully vetting data to reveal new uses and maximizing use of readily available data from state and federal agencies can improve information for making decisions.
- GPS data works well for accurately identifying locations and durations of truck stops, particularly in rural areas.

SPOTLIGHT

New Approaches to Address Persistent Freight Data Gaps

KATHLEEN HANCOCK
Virginia Tech, Moderator

RAUFUL ISLAM
Virginia Tech, Recorder

In this speed session, presenters discussed innovative artificial intelligence and machine learning approaches for addressing common freight data gaps. The following presentations of applications were followed by a panel discussion.

- **A Two-Step Approach for Truck Load Factor Estimation Based on the Travel Monitoring Analysis System Data Using K-Means Clustering and Gaussian Mixture Model,** Hyeonsup Lim, Majbah Uddin, Ho-Ling Hwang, Shih-Miao Chin, Chieh Wang, *Oak Ridge National Laboratory*.
- **Evaluating Public Perception Towards Autonomous Goods Delivery Services Using Social Media Data,** Majbah Uddin, HopLing Hwang, Shih-Miao Chin, *Oak Ridge National Laboratory*.
- **Investigation of Lidar Intensity for Anonymous Fleet-Level Truck Characterization,** Koti Reddy Allu, Andre Tok, Stephen Ritchie, Zhe Sun, *University of California Irvine*.
- **An Investigation of Classification by Gross Vehicle Weight Rating Categories Using Advanced Single Inductive Loop Sensor,** Guoliang Feng, Yiqiao Li, Andre Tok, Stephen Ritchie, *University of California Irvine*.

INTRODUCTION

Public agencies tasked with freight planning and operation are constantly hindered by lack of timely data including vehicle size and weight, payload across commodity type, fleet logistics, operation time, route use, etc. Much of this information is proprietary to the private sector. As a result, freight planners rely on aggregate commodity data, costly surveys, and on-road vehicle speed and weight data. Existing prediction models developed from these data are limited in their ability to depict freight movement on the transportation network. New approaches to address these persistent freight data gaps continue to be an active field of research.

Lim and Uddin focused on leveraging currently available data using data analytics techniques to address current data needs and to gauge public perception of currently active autonomous goods delivery services. Allu and Feng were focused on leveraging datasets obtained from a new data collection system that uses different imaging and induction sensors.

SUMMARIES OF INDIVIDUAL PRESENTATIONS

Lim presented a two-step data analytics approach to improving the estimation of empty truck weight and better estimating truck payload factors using the Travel Monitoring Analysis System (TMAS) weigh-in-motion data. Truck payload estimation is difficult because of the wide range of truck body type and goods moved by the trucks. In the first step, a Gaussian mixture model is used to separate the gross vehicle weight (GVW) distribution of empty/unloaded, partially loaded/loaded to size limit, and fully loaded/loaded to weight limit trucks. The purpose of this step was to illustrate that the Gaussian mixture model can be used to separate empty trucks from the rest of the truck population. Empty truck weight can be used to estimate payload. Due to the wide-ranging body type and axle configuration in FHWA vehicle class 9, the variance in payload and the empty truckload was shown to be very high. In the second step, subgroups within vehicle class 9 were created by using the K-means clustering method. The objective of this is to create subgroups within vehicle class 9 which have similar empty weights with equivalent payloads. This method was shown to have reasonable success in differentiating between sub-group of axel configuration and empty truckload within vehicle class 9.

Gauging public reaction to new technology has always been important to transportation planning and operation. Uddin evaluated public opinion toward autonomous goods delivery services using social media data. In this research, tweets mentioning autonomous goods delivery are used to infer public reaction. The presentation showcased the potential of social media data as an alternative to traditional surveys. Two different Natural Language Processing (NLP) techniques were used to analyze language in tweets for 1) sentiment and 2) topic. Data pre-processing and filtering necessary for application of NLP were discussed. The result of sentiment analysis was presented in terms of subjectivity and polarity, where polarity of sentiment represents how positive or negative the opinion expressed in the tweet. The result of the topic analysis identified points raised by users related to the primary topic, autonomous goods delivery services. This presentation illustrated the ability to infer public opinion from social media as a potential alternative to traditional costly surveys. Uddin also indicated that using single platform-specific data may lead to bias in the result.

A participant asked about the general credibility of social media data considering the prevalence of “paid posters” and “bots” which could contaminate the dataset and portray a false representation of public perception. The presenter responded that mitigation measures were available to filter tweets posted by bots and were used in the study; however, filtering out a tweets from paid posters was still a concern.

Allu introduced a new generation of information from lidar data and its application in characterizing trucks at the fleet level. The research focused on transforming lidar data into a useable image format, which can then be used to classify fleets based on identifying marks (e.g., text or logo) on the truck body. Lidar data is independent of ambient lighting conditions, making it suitable for gathering data at any time of the day. Another factor is that the methodology is data-specific and not sensor-specific, making it applicable for a wide range of lidar sensors. In this application, the level of characterization is limited to information associated with the text or image, such as a company name/logo or an image of a product. Fleet-level truck characterization using identifying marks on the body has the potential to differentiate between long-haul and short-haul, freight and non-freight, large and small carriers, and other fleet-level activity.

A question from a participant was raised about the availability of this new dataset and another about privacy issues with truck identification that the freight companies might have. In response, Allu indicated a future consideration is anonymizing the identification marks on the vehicle before releasing the dataset to the public

Feng offered a new method for leveraging data from advanced single inductive loop detectors, which can classify vehicles into FHWA vehicle class. When a vehicle passes over an inductive loop sensor, it imprints a temporal signature on the sensors. The bagging K nearest neighbor (KNN) classifier was used in combination with the bootstrap sampling technique to process this signature. This signature can then be used to estimate the body type classification of the vehicle. Along with the GVW and axle configuration that can be obtained from the same loop sensors, the estimated body type classification can be used to estimate GVW rating (GVWR) categories defined by the FHWA. This classification method can give insight into the freight vehicle fleet operating on the network. In total, eight GVWR classes of vehicles were used in the study. The analysis showed that given a high enough sample in each category, the model provided satisfactory results in classifying vehicles. Having information about the vehicle fleets operating on the transportation network can help state and local planners in forecasting, modeling, and emission estimation.

SUMMARY OF PARTICIPANT DISCUSSION

During the panel discussion, participants were encouraged to share information about ways to fill the common freight data gaps.

The panel discussion started with a consideration of the near and long-term application and adoption strategies of the new freight data sources introduced in the presentations. Several adoption strategies were discussed. Uddin pointed out that the data used in the first two presentations are already publicly available. For example, the TMAS weigh-in-motion data is available from FHWA, and the tweets are obtainable from twitter.com. Different data analytics technology is available for researchers to obtain relevant information from these data sources. Allu stated that adopting lidar data in the near future would be difficult as this data collection method is still in the research and development phase. He also noted that the lidar data has the long-term potential for use in truck traffic assignment at the commodity level and setting adoption strategies of autonomous vehicle fleets in freight movement. Feng stated that data from induction loops could be obtained with or without the induction signature. In near-term applications, induction loop data can be used to get spatiotemporal variation of vehicle classes on the network. These data can help freight planners improve their prediction and assignment model. This classification method can also be used in enhancing vehicle emission estimation.

The discussion then considered the application of the datasets in filling other freight data gaps. Allu suggested that the lidar data can be used to fill data gaps that exists for commodities moved by similar truck body type, and the micro-trajectory of trucks obtained from lidar data can be used in emission estimation. Uddin stated that TMAS data could be used to estimate the station-level freight movement direction. This directional movement can give insight into the overall economic activity of a region. He also encouraged using different social media data to assess public reaction to new policy decisions with appropriate precaution in filtering the data for intentional bias. The mobility analysis aspect of social media data based on the geo-coded tweet of posts from other platforms was also discussed as a potential freight data source. Feng stated that loop inductive sensor data could be further used in refining the GVWR classification of vehicles.

KEY TAKEAWAYS

- The presentations and subsequent discussions highlighted the need for new innovative data sources to address the persistent freight data gaps.
- Data analytics techniques applied to nontraditional datasets have the potential to fill some gaps.
- Appropriate safeguards are needed for adopting these innovations.
- Due to the proprietary nature of freight logistics, privacy issues, and rapidly emerging new technologies, gaps in freight data will continue to exist.
- A data requirements primer at different policy and operational levels could be beneficial for informing future research directions.

PANEL

Freight and Surrounding Communities

Measuring Impacts to Inform Infrastructure Decisions

AKIKO YAMAGAMI
Los Angeles Metropolitan Transportation Authority, Moderator

ALISON CONWAY
City College of New York, Recorder

This session brought together community leaders and multidisciplinary researchers to discuss performance measures and data, as well as research needs for measuring the health, safety, and quality of life impacts of freight activity and infrastructure on local communities.

- **Juan De Lara,** *University of Southern California*
- **Sue Dexter,** *University of Southern California*
- **Robert Binder,** *University of Southern California*
- **Wilma Franco,** *Southeast Los Angeles Collaborative*

INTRODUCTION

Akiko Yamagami opened the session detailing questions to be addressed by the panel related to measuring community impacts of freight. These included:

- What do we need to be measuring?
- Who should be involved?
- What unconventional data do we need to better measure impacts?
- What is the role of the freight data community to lift the voices of communities?

Panelists introduced themselves and detailed their relevant experience.

Wilma Franco serves as the Executive Director of Southeast Los Angeles Collaborative (SELA), a network of organizations that work collaboratively in the southeast region of Los Angeles County to promote civic engagement and strengthen the capacity of the nonprofit sector. SELA works to uplift community needs through collective work; to ensure that benefits and impacts of developments and policy are informed by the community; and to support the long-term sustainability of nonprofits. The network's work is data driven and is informed by research.

Sue Dexter and Robert Binder are Ph.D. candidates at University of Southern California (USC) who have worked with SELA on collaborative research projects. Dexter's project, which builds on her prior experience at Toyota optimizing supply chains to reduce the emissions footprint, focuses on reducing local emissions impacts from drayage truck operations. Bender's work is focused on both safety and emissions, and includes both macro- and micro-scale modeling using different data sources across several Gateway cities.

Juan De Lara, the inaugural Director of the Center for Latinx and Latin American Studies at USC and a trained geographer, has extensive experience and publications examining the relationship between logistics and urban ecologies in Southern California. In his introduction, he noted a number of key questions for the panel and audience to consider, including “Why do we build spaces that support specific types of development?” and “How does a project benefit the lives of communities in neighborhoods (in addition to freight efficiencies)?” He noted that sometimes projects that do not harm surrounding communities shift damage somewhere else. He also noted that sometimes projects meet mandates, but those mandates are not always well considered. Projects are often justified on the basis of economic growth and job creation, but it is important to consider the quality of jobs, potential health impacts, and even job-related harms in a more holistic way. He noted the need to reframe development and talked about community participation that does not start from an official development agenda, but rather with what types of projects will improve community well-being.

Following introductions, Yamagami reiterated the need to understand how projects benefit communities and the importance of community engagement. The following questions were addressed by the panelists.

- **Where do traditional transportation analyses fall short on capturing impact on communities, resulting in insufficient understanding of freight impacts on underserved communities?**

Binder and Dexter shared experiences and lessons learned from their collaborative research with SELA. When their research project began, the university researchers had little connection to the community. Different priorities were identified during community engagement. While the research team expected pollution to be a major concern, it became clear from community feedback that traffic safety was a more critical priority. The community was concerned about trucks on local streets—particularly heavy trucks in mixed-use communities such as those where pockets of industry are surrounded by residential areas. Of particular concern were intersections with erratic truck behavior and a lot of pedestrians.

With direction from community input, USC researchers conducted a thorough analysis of collisions using the UC Berkeley Transportation Injury Mapping System (TIMS) database that consolidates safety data statewide and classifies incidents by roadway type and HDT involvement. Once researchers were able to identify dangerous intersections and corridors, they shared results to ensure that these aligned with community experience. Community engagement was critical to inform the direction of the project, but unfortunately was limited in later stages of work due to the COVID-19 pandemic.

De Lara noted that the previous speakers’ comments were very instructive about discussions that need to continue. He noted that there is a difference between community participation and community self-determination. Planning processes typically justify an existing plan. Participation is not enough; the process needs to be more focused on effective community decision-making.

He also provided an example of how community impacts might be better illustrated by leveraging multiple types of data. Referring to a map that Dexter had shared showing local emissions, he asked “what happens if we overlay a vulnerability impact?” By doing so, he noted, we might see that high emissions are also in neighborhoods right next to a school, where proximity to the source is a major public health risk factor. He also noted that analyses are often conducted at an aggregate level, but that there is a need to go down to a micro-scale to understand risk factors. He also noted the importance of contextualizing this information. Often, affected neighborhoods are home to low income, people of color, and immigrant residents who have less access to healthcare and higher levels of asthma. Once access to preventative healthcare is mapped on top of risks, problems start to magnify.

Franco noted that the community in Southeast LA is home to large numbers of both schools and factories that produce many truck trips, and that the residents are 90% Latino, and are largely immigrant and low income. She noted that the community is aware of pollution, high asthma rates, and limited healthcare access, but that traffic safety concerns have more immediate impacts. Residents are often scared to leave their homes after they see traffic incidents. She provided recommendations for improving communications with the community, including avoiding “word salads,” especially the use of technical jargon that the community does not understand. A good way to overcome this is to partner with those in the community who do understand, and are also able to meet other community members in their spaces to make this information accessible. She noted that “checking the box” is not meaningful engagement.

A participant asked if specific projects had been identified for implementation, as a forthcoming California Senate Bill could potentially provide funding. Dexter noted that the project report was in progress, with some potential areas for implementation already identified, including, zero-emissions, low-carbon pilot projects and truck geotagging to ensure vehicles are not traversing prohibited residential areas. She noted that both require local industry participation.

Bender noted that modeling was ongoing due to software-related delays, but also identified potential projects. One project was evaluating the potential to shift truck traffic onto the auxiliary Alameda Corridor, which is separated by rail from the Alamada Corridor. This would require construction of a crossover and related geometric and operational improvements. A second project was an investigation of why trucks were diverting from a primary route to a prohibited parallel route so that geometric or enforcement solutions for prevention could be identified.

De Lara noted the importance of analyzing systems at multiple scales, including looking at broader systems while identifying micro-level impacts. He provided specific examples relevant to the Alameda Corridor. The ports could only grow because of inland warehouses. As a result, many trucks were entering and idling in small neighborhoods in the Inland Empire. Aggregate-level data does not typically scale down to these individual neighborhoods. Effects can also be wide-reaching and have multiple focus points or hubs. For example, the cities of Redland and Riverside, similar to the areas immediately surrounding the ports, are largely Latino and working class. These communities sued the ports when they promoted expansion projects because of impacts 65 mi away.

- **What types of data are needed to truly shed light on project performance on community-expressed desirable outcomes?**

Dexter noted that it is difficult for researchers to access movement-by-movement data. She noted that GPS tracking data is proprietary, and that it is expensive to purchase. Identifying truck locations at specific times and understanding vehicle types and loads can aid in understanding why they are where they are. She noted that competitiveness and privacy are also concerns with GPS data, although geofencing may address these. Without this type of local data, researchers are reliant on high-level simulated data (e.g., a regional travel demand model), which is typically unable to reveal what is happening on the ground in communities.

Binder noted that during community engagement, members had a connection to the industry (e.g., a family member or friend working in a warehouse), so there is concern about discussing negative impacts. He noted that a need exists to get both industry and the community involved early to collaborate and identify opportunities for a symbiotic relationship. He also noted that understanding land use mix can inform where interactions happen, but does not identify-specific interactions.

- **How are Southern California communities reacting to increased automation?**

Dexter discussed the following advancements in automation and new technologies:

1. There is a lot of automation at the port, which is a controlled environment;
2. Manufacturers of automated trucks, currently being piloted in Arizona and Texas, have indicated that they are not yet ready to operate in an urban environment;
3. Warehouse automation, which provides substantial efficiencies in storage and retrieval through robotics, is coming quickly and will displace a lot of workers; and
4. California agencies are considering how to train responders to alternative-fueled vehicle incidents.

Franco noted that job creation is often identified as a project benefit, but that the conduits are not in place to connect affected communities to these opportunities. As a result, community disruption and whether it will displace community members are concerns. Yamagami agreed that there is disconnect between information and services. She noted that there are many academics working across 4-year universities, community colleges, and vocational schools developing training programs in many related areas, but information about these programs are often not being communicated to the people who need that information most.

- **How can academic institutions partner with community-based organizations (CBOs)/nonprofit organizations to support their data-driven decision-making? What would be a meaningful model for community stakeholder engagement, whether through academic institutions or public agencies?**

Franco noted that when CBOs are engaged as partners, the interactions should be included in the project budget, and participants should have autonomy as they contribute to decision-making on how to engage the community. It is also important to think about how to compensate community members and CBOs when building partnerships and participating in the process. She noted that during the research collaboration with USC, SELA were thought-partners and decision-makers, which enabled the team to effectively reach the community and keep them engaged. For example, with SELA's guidance, focus groups were conducted in Spanish and translated to English for the research team in the room rather than the other way around. This increased accessibility for community participants during the discussion. The project also included an advisory committee with many stakeholders (e.g., ports, Caltrans, LA Metro, academics).

De Lara stressed the importance of engaging community leaders, like Franco, as research partners. He noted that universities are sources of knowledge creation, and that professional programs and grant opportunities have been focused on doing data analytics. He noted that the SELA–USC collaboration is a great example of what can happen when communities are centered in knowledge production at universities. He also noted that universities need to do a better job of influencing how knowledge is produced; there is too much fragmentation in the university system. In particular, De Lara believes that academics who study social issues are often not in touch with those on the technical or data analytics side. Finally, he reiterated that when thinking about job creation and opportunities, it is important to consider who has access to training. It is not just a question of sharing information, but also of connecting resources to provide access.

KEY TAKEAWAYS

The session identified a number of good practices, as well as opportunities for future improvement. First, the session identified good practices for engaging communities in research:

- It is important to engage community members early, at a point when their inputs can drive the direction of a study.
- Community-based organizations can provide critical connections to community members and assist with overcoming communications challenges.
- Community-based organizations should be compensated for their time, and they should be active partners in project decisions.

Next, the session identified insights relevant to measuring community impacts:

- Impacts should be studied at multiple scales. Impacts of freight projects can be far reaching, and can produce impacts at multiple hubs. It is important to consider impacts from a broad systems perspective, identifying upstream and downstream effects on other communities. Impacts of freight projects can also be very local. Micro-scale neighborhood, street, or even building impacts need to be investigated to identify local risk factors.
- Particularly when conducting micro-scale analysis, it is important to overlay performance and externality data on community characteristics and vulnerabilities, both to identify factors that may exacerbate effects and to provide appropriate community context.

Finally, the session identified potential improvements for the practice of research and for connecting local communities to job training opportunities:

- More or different resources are needed to support community-driven research projects.
- Future research could benefit from more partnerships between researchers in the social sciences and in data analytics and other technical areas.
- A need exists for better conduit programs to connect communities to freight-related job opportunities. Specifically, when developing training programs in the logistics sector, careful attention should be paid to communication of training opportunities to local residents and ensuring access, both physically and financially, to those opportunities.

CLOSING SESSION

Key Takeaways

DONALD LUDLOW
CPCS, Moderator

KATHLEEN HANCOCK
Virginia Tech, Recorder

This is the third Innovations in Freight Data Workshop that has taken place, and each time it has become a more comprehensive and enlightening event. I am going to start by talking a little bit about what we have learned as a group, and then I will provide some comments on applications that we have seen. I will then close with a few comments on the future to segue into a closing discussion, which will help shape the research agenda going forward.

The freight data community is vibrant and growing. Approximately 180 people participated over the course of the 3-day workshop; part of a broader community of several hundred individuals, who are working in agencies, consulting firms, academia, and industry, to advance our craft, which is the vitally important practice and career choice where we operate at the intersection of logistics and public policy. Our role is to constantly push the envelope and integrate new sources of data to address the problems of the day and to translate the data into information for decision-makers.

Both veterans and many new voices were heard over the 3 days of the workshop. This community is growing because we need a resilient, safe, and sustainable logistics system. This matters more than ever with the rapid technological, socioeconomic, and environmental changes that require us to be better than we ever have been. As a result, the workshop brought people together to address a number of important and changing problems and opportunities.

One thing that we have learned is that data sources continue to proliferate. There are new sources, new vendors, and more offerings, many of which are fused from one or more datasets to produce value-added products. These allow agencies to compare and examine the most suitable data. It can also be a double-edged sword, making it more difficult to choose efficiently. A number of new products have been coming online, including enhanced data from the federal government. Examples include the enhanced FAF dataset, electronic logging device data, data from application programming interfaces (APIs), and sensor data, all of which continue to reshape our world. For those who have attended multiple innovations workshops, we have seen progress as these new data sources have moved online and become integrated into our work.

Vehicle probe data are now widely used and are potentially one of the most important observations from this workshop. Observed data have largely replaced vehicle estimation tools. It is becoming standard practice to use a combination of datasets including commercial vehicle probe data and land use data, to estimate truck movements and to address local last mile–first mile issues and performance on the transportation system. For example:

- In NJTPA, these data were used to differentiate between truck trip types, vehicle types, and package deliveries, when combined with socioeconomic data, these data provided a more complete picture of how e-commerce is affecting the region.

- Another session detailed information on truck parking and how vehicle probe data are being used to unlock truck parking observations. As recently as two years ago, we were just starting to talk about the nuances of truck parking approaches using vehicle probe data and now we are starting to identify best practices and to push the envelope to delve even further into this issue. The power of understanding the extensive use of vehicle probe data and sharing best practices for its processing and use allows researchers and practitioners to more effectively communicate its value to decision-makers, industry and government, who can move this forward to make investments that can make a difference.
- AIS data, once the domain of freight enthusiasts, are now becoming so common that many MPOs, such as the Cincinnati MPO, have a robust API system to support decision-making. A workshop participant indicated that AIS is ATRI data for the river.

An important aspect of several presentations was data fusion. Rarely are we seeing use of a single dataset. The most common recipe to address freight issues is combining several different sources to address issues. For example, the Southeast Metropolitan Planning Organization, the MPO in Chicago, uses multiple sources of vehicle probe data, each with its respective strengths, and a combination of land use data and commercial real estate data to tell the full story, bringing all of these together to relate and quantify the movement of freight and investment decisions that are made because of it.

Several presentations provided information on new data processing methods. In earlier innovations workshops, topics from machine learning to computer vision and artificial intelligence (AI) had been identified as research for the future. Many of these are now available and being used for different aspects of data processing, including standardization of data, cleaning and combining of data and the investigation of the data. The workshop provided successes and also some examples of the need for more research, such as the Oregon truck parking study where research into use of machine learning was ultimately replaced with the use of linear regression.

Another overarching theme is collaboration. Numerous datasets, provided by private companies, have supported public freight analysis and decision-making for years and continue to exemplify this collaboration. There are also growing collaborative efforts among public and private sectors and academia to unlock freight data's potential. For example, the Port of Los Angeles, with its Port Optimizer, synthesizes data from competing firms all to the same end, to make port operations more seamless and efficient. We are also seeing more collaboration between groups of agencies, universities, and the private sector to investigate freight issues. For example, in northern Kentucky, a local university and MPO partner to use data from DHL and Amazon to unlock and investigate future bottlenecks in that region.

The workshop also highlighted a number of problems and how data are being used specifically to address them, many of which reflect the day and age that we are in. There has been more focus on e-commerce, including the work performed by NJTPA related to package data. Emphasis is increasing on land use and fulfillment centers and their impact on communities nearby, including environmental justice and equity. Southern California also presented on their approach to problems related to economic impacts, truck parking, curbside management, emissions, sustainability, and resilient supply chains.

More decision-making tools are available. Strategic and operational decision-making is becoming more data driven, creating more demand for better and more refined freight data. For example, the project selection process that the CTC's Trade Corridor Enhancement Program is very data driven, and is continually seeking more data to improve its outcomes. The Port Optimizer, the interactive dashboard for Port of Los Angeles, is another example of a decision-making tool focused on operations that enables the port to make small and large changes to improve efficiency.

Visualization continues to improve through use of graphic interfaces, story mapping and other kinds of digital depictions. These interactive tools are making the data more accessible and useable. It still means that the freight data practitioners are often crafting these tools, but they can now focus on uncovering the data of interest and making it more applicable, transparent, and widely understood.

Finally, the following presenters provided thoughts related to the future based on observations during the workshop.

- Freight data continue to have tremendous potential and this community continues to rise to the occasion. Robyn Bancroft said that using freight data for decision-making is a dance rather than a straight walk. It is a nuanced activity that requires us to be able to interpolate and bring together the best sources and methods to make things work.
- Many data gaps still exist. Some of these are the same gaps that have been identified for years, such as the ability to conflate cargo with conveyance. The freight data community is getting better at inferring this through land use, business, and truck movements. The picture on understanding equipment, chassis, and connections is still missing. Major gaps remain between rural and urban data. This is increasing in importance with the increasing focus on rural equity and rural transportation.
- In the future, the community will continue to face a proliferation of data sources, difficult decisions about which data to procure and how to most efficiently and effectively apply these data. Considerations include potentially high costs, specialized software, staff training and analytical expertise. Balancing these results in difficult decisions about allocating resources appropriately including whether to outsource some of that work. This goes back to understanding what is needed and what is available and how one meets the needs of the other. There is not a lot of readily available information to enable these decisions.
- The freight data community will continue to see more use of AI, machine learning, data collected through instrumentation of transportation systems and vehicles, and data from imagery and computer vision.
- We will also need to train and attract the best people to freight operations and planning who are able to understand the different dimensions and to share in our passion for moving this forward. All of this is happening, and will continue to happen, around the backdrop of a continued standardization of the approaches that are used for analyses like uses of commercial vehicle probe data for identifying and measuring bottlenecks, congestion and truck parking.

AUDIENCE PARTICIPATION

Alison Conway led the participant discussion with a series of questions. The bullets provided under the following questions were provided by individual participants.

- **Were there any new types of data or new methods of data collection that you learned about during the workshop?**
  - New types of sensors. Lidar and its potential future applications to address gaps in making the connection between vehicle information and commodity information on a vehicle, specifically for identifying logos and fleets from those logos.
  - Approaches to measuring the community impact of freight. Layer information from other sources that we might not have traditionally thought of, such as public health data and public health access data to understand how communities are going to be affected by a project.
  - Using GIS to overlay different pieces of information as a key tool.

- **Were there any new methods that you learned about during the workshop?**
  - Machine learning is moving into the forefront of how we are analyzing data.
  - Use of leading indicators to predict future demand. This is a good example of how data is not just a source, but how it can be used to inform both the past and the near future, and how it can be used for making decisions.

- **What did you learn about identifying how data are informing different types of decisions and conversely, which types of decisions do we not yet have the data we need?**
  - For the NJTPA project, which was an innovative project looking at e-commerce, many municipalities are trying to answer the question "Is this proposed e-commerce project good or bad for my community," which comes down to a tradeoff between jobs and traffic and its corresponding environmental impacts. Answering that question requires the ability to predict what future e-commerce activity is going to be happening in a facility that might be "built on spec," meaning it is unclear who is leasing it or if it is going to be a return center or a distribution center. As a result, the data that we have right now are not necessarily supporting decisions that need to be made at that municipal level.
  - The Port of New York and New Jersey developed GIS software, which integrates traditional environmental data about what is happening underwater with AIS data to understand where ships are entering and exiting the port. By combining these two data sources, they were able to better inform the process of planning for dredging. Combining the datasets, enabled decision-makers to understand the drafting depth of the ships, as well as the frequency of their movements in port. By visualizing the combinations, they could understand the complexity of what was going on, providing a very specific kind of operational or maintenance decision.
  - Data gaps related to understanding the impacts of freight on local communities still exist. There is progress but what more can be done? An example could be something similar to the Health Opportunity Index in Virginia, which drills down to the Census Tract level on community impacts. A new frontier for the freight data community is integrating that kind of information and fusing it with other data sources to tell the story of community impacts in a more meaningful way.
  - A good example is Georgia DOT's effort on designated truck parking using multiple datasets to address the types of decisions that Georgia needs to make in their investment in public truck parking and/or partnering with industry to add truck parking. This is a specific, but very critical decision that a lot of agencies are facing.
  - Layering earthquake and other natural disaster data to understand risks.

- **Who was not part of this workshop that should have been, or who do we need to engage with to move the conversation forward?**
  - We tend to be siloed into topic areas. As we consider issues like public health, we should be connecting to those communities.
  - Weight-based, road-user-charging policymakers specifically for tolling.
  - Professionals in electrification, which, as we move towards integrating alternative fuel vehicles and electric vehicles, is a consideration as well as the interdependencies between freight systems and energy systems.
  - E-commerce companies like Amazon and Walmart.
  - Those involved in taxing.
  - Federal Emergency Management Agency and their community lifeline.

- **Do you have any suggestions for the next workshop?**
  - Addressing the gaps between the perspectives of planners, engineers, economists, and the public, with the goal of learning about and contributing to helping these groups better communicate with each other.
  - Case studies on specifically how data and tools were used in decision-making.
  - Focusing on freight-related activities that we want to learn more about.
  - Helping different stakeholders to communicate with each other. Understanding who is actually incentivized to participate in these discussions and who is not. Identifying what incentivizes stakeholders to initiate these discussions.

- **What research would help advance our practice and help address the remaining data gaps?**
  - Data procurement. This is an issue that needs to be addressed in many agencies. Taking best practices from organizations that have been struggling with these challenges and have identified both good methods of doing so, as well as what challenges continue to exist, so that they can be avoided in future.
  - Data-sharing agreements. Sharing best practices and communicating how to structure data sharing agreements and defining what template(s) might look like for different types of agreements. The Eastern Transportation Coalition, formerly the I-95 Corridor Coalition, has done some good work on data-sharing agreements, so they may be a good source to look to for more information on that.
  - Business activity data sources. How do we develop business activity data sources that can be used to develop, standardized measures for freight generation using truck activity inference data, possibly using body type information and creating a broad vehicle typology?
  - Standardizing data derivatives. For example, converting vehicle probe data to truck trips. As our practice moves forward with new data sources, we have seen a lot of standardization start to happen in metrics and data processing to obtain these metrics. What can we learn across locations and geographies?
  - Best practices for using data for decision-making.
  - Accessible tools for analyzing probe data. For example, what Python packages could be implemented by lay users?
  - User- friendly tools. Making data useful and repeatable for both analysts, who are working with the data, and the public for understanding the story the data tell.

– Data visualization platforms: Identifying characteristics of these platforms that ensure that they are useful to different users to produce the information needed for different types of decisions and at different levels of decision-making.

Interactive Showcase

In this interactive poster session, presenters shared recent tools and applications for freight data analysis and visualization. Presenters interacted directly with attendees.

TRUCK CONGESTION ANALYSIS TOOL

L.D. WHITE
Texas A&M Transportation Institute, Presenter

The Truck Congestion Analysis Tool (TCAT) is a web-based planning tool for visualizing, analyzing, and monitoring truck mobility in Texas. The underlying data were developed as part of the Texas Top 100 Congested Roadways effort between the TTI and the Texas DOT. Through TCAT, practitioners have access to many of the traditional mobility performance measures, including annual delay, delay per mile, congestion cost, travel time index, planning time index, and others. TCAT also provides practitioners context information regarding truck congestion. These performance measures are based on traffic volumes from the Texas DOT Roadway-Highway Inventory (RHiNo) road network and proprietary traffic speed data (INRIX). Practitioners also have access to the Top 100 Congested Roadways information and the more comprehensive RHiNo-based roadway segment dataset that allows for evaluation at different geographic scales including individual roadway segment, custom/user generated corridors, or regional summaries based on the selected region. Significant geographical regions are pre-defined in TCAT, including Texas DOT districts, Texas counties, MPOs, major regions of interest, and major corridors–roadways of interest. TCAT has two distinct analysis levels: road segment and regional. The road segment analysis level provides access to the performance measures associated with each road segment. Custom corridors of multiple road segments can also be created using from here/to here selections similar to adding pushpins to a map. The regional level provides practitioners quick access to summaries of the mobility performance at the much broader area level. TCAT has summaries at pre-defined aggregation levels (i.e., Texas DOT district, county, roadway classification, rural–urban) and a custom summary that provides the ability to aggregate performance measures by multiple categories. As an extension of the region summaries, annual truck congestion report cards are available for the Top 100 Truck Congested Roadways, as well as each of the pre-defined regions. These report cards were designed to quickly examine performance trends.

One of the goals of TCAT is to provide context information regarding truck congestion in Texas. A layer that is currently active in TCAT is road projects from Texas DOT's Unified Transportation Program (UTP). The UTP is Texas DOT's tenured project plan that guides the development of transportation work across the state. TCAT accesses the project information through the Texas DOT Data Portal to ensure consistent output across platforms. In addition to the UTP projects, the context information in TCAT will be expanded to include additional map layers of various freight infrastructure and truck generators. These layers will include location information of maritime and inland ports, international ports of entry, intermodal facilities, major distribution centers, warehousing hubs, border crossings, and energy sector considerations (i.e., disposal wells, sand mines).

DREDGIS+: DATA-DRIVEN DREDGING PLANNING FOR THE PORT OF NEW YORK AND NEW JERSEY

JEFFERY BRAUNER
Port Authority of New York & New Jersey, Presenter

Planning and coordinating maintenance dredging at the Port of New York and New Jersey, the third largest container port in the United States, is a growing challenge. Berths are steadily shallow, while the fleet of vessels increase in size and require deeper water. Dredging costs increase, while budgets, especially post COVID-19, decrease. At the same time, the port is busier than ever, with a volume of 7.6 million TEU in 2020, a 1.5% increase over 2019, and some months seeing greater than 20% growth over 2019, despite the ongoing pandemic. Dredging the right berths at the right time is more critical than ever since, without proper depth, ships cannot call, and goods cannot move. To meet this challenge, DredGIS+ was developed. DredGIS+ is a combination of two distinct, but complementary datasets; AIS vessel and berth call data collected by MarineTraffic, visualized in Tableau, and hydrographic survey data collected by multibeam sonar, visualized in ArcGIS. Together, these provide the Port Authority's dredging team with a holistic, data-driven approach to supporting optimal decision-making related to what berths to dredge, when to dredge, and why. Uses of DredGIS+ include:

- Comparing historic survey data to analyze change over time to readily assess hot spots where berths are quickly silting up and prioritizing those for dredging;
- Analyzing the draft of vessels calling at different berths when evaluating whether to dredge one berth over another; and
- Evaluating berth availability over time, especially at busy container terminals, to determine the least disruptive time to dredge.

The 2021 dredging program has been developed based on insights from DredGIS+. DredGIS+ helped identify areas of the port that are silting up quickly, flagging berths for inclusion in the dredging schedule. Due to limited funds and prioritization of container terminal tenants, all berths cannot be dredged each year. To make the most of limited resources, MarineTraffic data was leveraged to determine which auto and bulk berths should be prioritized ahead of others, based on vessel frequency and draft. As an example, two berths were identified that required a similar amount of dredging. Berth 11 had 16 vessel calls with a draft greater than 30 ft, with a maximum draft of 34 ft from July to December 2020. Berth 8 through 10 had only six vessel calls in that range, with a maximum of 32 ft of draft. DredGIS+ helped inform the decision to direct funds to dredge Berth 11. Data now supports and reinforces prior experience and knowledge of the port using expertly visualized information in creative, innovative, and useful ways to support dredging decisions to keep the Port of New York and New Jersey running and keep the region moving.

MINING EXISTING TEXAS OVERSIZE/OVERWEIGHT TRUCK TRIP PERMIT DATA FOR FREIGHT TRANSPORTATION PLANNING

WILLIAM HOLIK
Texas A&M Transportation Institute, Presenter

This research consisted of an innovative data mining process conducted by the TTI and sponsored by the Texas DOT on existing oversize/overweight (OS/OW) permit applications data to improve statewide understanding and planning for OS/OW truck movements. The Motor Carrier Division of the Texas Department of Motor Vehicles (DMV) administers the OS/OW truck permitting system in Texas through the online Texas Permitting and Routing Optimization System (TxPROS). It is used for request, approval, issuing of permits, and designation of assigned routing for OS/OW loads. For this project, Texas DMV provided data on all permits issued from CY2015 through CY2020. Texas DMV provided two files: a permit report describing the applying company, type of load, dimensions, and GVW, and a GIS shapefile of the proscribed, permitted route. The raw permit data required extensive processing to remove extraneous inputs and common misspellings by applicants to identify common industry loads and commodity types. Statistics about the number of OS/OW permits issued, types of permits, load types, and time series data were generated from the processed data. The permit data were then joined to the route shapefiles to expand analyses on OS/OW movements for the top four industry types (construction, manufacturing, manufactured housing, and oil and gas) and specific load characteristics (loads over 250,000 pounds GVW and over 16.5 ft in height). This was then spatially joined to the Texas roadway network files to obtain total GVW and number of permits issued for each segment. This process was done for each month and then summed into annual summaries for each roadway segment level. The outcome of this process was a modified roadway network layer that contained roadway attributes and the number of OS/OW permits and total GVW for each month of the analysis period for each road segment. Roadway attributes include AADT, annual average daily truck traffic, crash data information, and congestion measures which allowed spatial comparisons with the OS/OW permits and GVW. OD trends for OS/OW permitted loads were analyzed to understand load shipments for various industries. ODs from special generators such as seaports and international border crossings were extracted to further analyze shipping data. This information can be used to inform freight planning decision-making as it provides a unique and targeted view of the impact of OS/OW loads on Texas DOT infrastructure.

ANALYSIS OF TRUCK MOVEMENT DURING EVACUATION BY TTMS

KATERINA KOLIOU
Florida Atlantic University, Presenter

The goal of this research was to investigate the movement of vehicles, by classification with an emphasis on freight, during two major evacuation events: Hurricanes Irma (2017) and Michael (2018) using data from Telemetered Traffic Monitoring Sites (TTMS). Response to Hurricane Irma has been referred to as the largest evacuation in the history of the United States and resulted in approximately 6.5 million Floridians being placed under either mandatory or voluntary evacuation orders. Hurricane Michael was the strongest storm, measured by wind speed, to strike mainland Florida in recorded history. The goal of an evacuation is to avoid injuries, loss of life, and, to a lesser extent, property damage and economic loss. Thus, a primary objective is to move all evacuees outside of a threat area as safely and as quickly as possible. This research sought to identify where and when different classes of vehicles were traveling leading up to hurricane landfall and post-storm re-entry. The research methodology was divided into three phases: data collection and management, spatial analysis, and temporal comparisons. For data collection and management continuous-count station data was obtained from TTMS sites across Florida. The second phase used GIS to display heat maps of where and when traffic varied across the state. The third and final phase was a statewide quantitative investigation into which vehicle classifications were statistically different on which dates. The multilevel analysis revealed that the effect of a hurricane on transportation started 2 to 3 days before landfall and return to pre-storm levels after 3 to 4 days. Significant differences in volumes were observed on landfall days across all vehicle classifications. Of the more significant findings, the research results showed that commercial-use vehicles may have underutilized rest areas during the evacuation or perhaps these rest areas were closed. The need for effective evacuation plans is critical. Truckers are driving longer distances and possibly longer hours before hurricanes so state planners could reconsider the role of rest areas in evacuation plans. Making these areas safer for truck parking during a storm or by repurposing these facilities to better serve evacuees if truckers decide not to use them. Another finding was that changes to traffic patterns for commercial-use vehicles occurred earlier and lasted longer than changes for personal-use vehicles. Commercial-use vehicles and vehicles for personal use operated in uniquely different ways pre and post evacuation.

IMPROVING COMMODITY FLOWS WITH PRIVATE-SECTOR BILL OF LADING DATA

MARK BERNDT
Quetica, Presenter

Two main sources of commodity flow data commonly used for public sector freight studies are available in the United States: FAF from FHWA and Transearch, a proprietary data product from IHS Markit. Both FAF and Transearch provide OD estimates of tonnage and value for freight transportation by mode. Both rely on the CFS conducted every 5 years to establish baseline information about commodity movements. The Census Bureau uses a stratified sample of 100,000 businesses nationwide in conducing the survey. Each business in the CFS sample is asked to provide shipping records for one week each quarter during the survey year. The response rate during the past two cycles (2012 and 2017) has fallen between 55% and 60%. In 2012, the survey resulted in a total of 4.6 million bill of lading (BOL) records, which were then used to estimate total national freight flows. In 2019, Quetica started sourcing and processing private-sector freight data for use in public-sector planning efforts and now maintains access to over 30 million BOL annually. The following provides a brief summary of how different clients have used commodity flows enhanced with BOL data:

- In 2020 Quetica purchased 4 million BOL shipping records on behalf of Texas DOT that were integrated with Transearch. Enhanced Transearch data is being used to estimate, analyze, and document the economic role of freight across Texas, as well as at regional and corridor levels. The findings are planned to raise awareness of the importance of safe and efficient freight transportation among state officials, local governments, economic development groups, and the public. The enhanced Transearch was used to estimate annual transportation costs for Texas shippers statewide, within the Texas Triangle, and for each of Texas DOT's 25 districts.

- For the Missouri's State Freight and Rail Plan update, Quetica acquired over 7.2 million BOL records for Missouri DOT. The BOL dataset includes shipments originating or destined to locations in Missouri, as well as shipments originating or destined to six surrounding states. The integrated Missouri Transearch dataset provides more detailed analysis regarding key freight facilities in the state, freight transportation spending estimates at the corridor-level for Missouri's National Highway System, and insights about the competitive nature of Missouri's modal options as compared to neighboring jurisdictions.
- Minnesota DOT is completing several major construction projects on the I-94 Corridor between St. Cloud and Minneapolis. Minnesota DOT's Metro District sought a better understanding of the businesses that rely on the corridor by integrating BOL data with StreetLight, business establishment, parcel information, and county-level FAF4 data to characterize the corridor's trip types and estimate annual spend on transportation services in the corridor. Results showed businesses spent more than $1 billion annually to move freight through this particular stretch of the I-94 Corridor.

IDENTIFYING SPATIOTEMPORAL ACCIDENT CLUSTERS AND ASSOCIATED TRAFFIC IMPACTS

STEVEN PETERSON
Oak Ridge National Laboratory, Presenter

The U.S. transportation system is subject to several significant issues and problems including fatalities, injuries, environmental impacts from vehicle emissions, and traffic congestion. Among these, traffic congestion impacts fuel use, productivity, emissions, and mobility. Recent studies, like from the TTI (Schrank et al 2019), estimate that the average driver wastes 54 h and almost $1,100 per year sitting in traffic, with the total cost of congestion in the United States topping $160 billion. For freight trucks, the ATRI (Hooper 2018) estimates congestion adding 1.17 billion hours of delay to freight movements and $74.5 billion annually in additional congestion-related costs, with increases in traffic accidents being a major contributor to trucking congestion. Further, freight shippers have limited options in reducing freight truck demand in a healthy economy.

This study addresses the relatively straight forward hypothesis that increased levels of congestion contribute to increases in accidents and that these accidents have correlated impacts on travel times, increased congestion, and diminished traffic mobility. Further, certain accident types may have greater impacts on traffic flows than other accidents. The study focuses on accidents involving freight trucks and their associated impacts on freight mobility. It also includes consideration of secondary accidents involving trucks. The study area compares accidents in two high-volume truck corridors in the state of Tennessee: the I-75/I-40 corridor through Knox County (Knoxville) and the I-75/I-24 corridor through Hamilton County (Chattanooga). An important aspect of traffic mobility and travel behavior modeling is the relationship between congestion and delay-causing incidents which have travel time impacts. To study this relationship, temporally and spatially consistent databases of travel times, speeds, network configurations, and traffic incidents are required. This study uses the NPMRDS, the RITIS, and the Tennessee Integrated Traffic Analysis Network, an incident database available from the Tennessee Department of Public Safety to model high-level system performance as well as specific roadway segment performance over time. The study identified truck accident clusters using network kernel density estimates and then used Bayesian estimation to assess factors such as travel time, average vehicle speed, and traffic volume trajectories, determining the spatial and temporal impacts of truck accidents on subsequent traffic flows.

IMPROVING FIRST- AND LAST-MILE FREIGHT MOVEMENT EFFICIENCY THROUGH CONNECTED VEHICLE APPLICATION

KANOK BORIBOONSOMSIN
University of California at Riverside, Presenter

Medium-duty and heavy-duty vehicles account for the second largest share of energy consumed by the transportation sector in the United States, and their share is projected to grow as a result of increasing e-commerce activity. At the same time, the transportation sector is going through transformational changes driven by automated, connected, electric, and shared vehicle technologies, which have opened new opportunities for significantly improving vehicle energy efficiency.

A recently completed research project combines connectivity and automation to support energy efficiency improvements in freight trucks using a connected vehicle application called Eco-Drive. Traffic signal controllers at 10 intersections on two urban freight corridors, Alameda Street and Wilmington Avenue, near the Port of Los Angeles were instrumented with an integrated router/modem to send signal phase and timing (SPaT) information to a Traffic Signal Information System (TSIS) server. The Eco-Drive mobile application onboard a truck can then receive the SPaT information from the TSIS server and drivers can use it to optimize the truck's speed profile as it approaches and departs each connected intersection. Results from the evaluation of Eco-Drive performance show that driving with Eco-Drive resulted in 6% to 15% lower fuel consumption than when driving without it. These savings varied by corridor: on Alameda Street, Eco-Drive helped the driver better comply with the speed limit (45 mph) and reduced speed fluctuations around the cruising speed, resulting in higher fuel efficiency; on Wilmington Avenue, Eco-Drive helped reduce the number of stops at connected intersections, which resulted in lower mean acceleration and deceleration values. In addition to fuel savings, driving with Eco-Drive on Wilmington Avenue resulted in travel time savings. These results demonstrated Eco-Drive's potential to significantly improve travel and energy efficiencies of freight movement by medium-duty and heavy-duty vehicles during the first and last miles of their trips.

MINNESOTA DEPARTMENT OF TRANSPORTATION METRO DISTRICT FREIGHT DATABASE TOOL FOR COMMUNICATING FREIGHT DATA TO PUBLIC AND PRIVATE STAKEHOLDERS

CHRISTOPHER RYAN
HDR, Presenter

The Minnesota DOT's Urban Freight Perspectives Study is focused on enhancing freight mobility in the Metro District by improving communication with freight stakeholders and integrating freight considerations into the district's project development process. The study focuses on how to integrate key takeaways from past freight planning work into Metro District's project development process. A key tool being developed for this project is a "living" database with segment-level and census-block-group-level freight information to help district project managers develop projects that integrate freight issues into the design process and address freight impacts during construction. The database is being developed as an ArcGIS Online web map series consisting of multiple StoryMaps, interactive maps, and embedded Tableau dashboards. The data used in this tool includes StreetLight Insight location-based services data, FAF data, Census County Business Pattern Data, Longitudinal Employer-Household Dynamics data, BTS T-100 air cargo data, Data Axle Business Information, and location-specific information collected during a series of stakeholder interviews. Many aspects of the data analysis process were expedited by the use of APIs to access StreetLight and Census data, and the use of software, such as RStudio to efficiently process the data. Currently completed sections of the freight database include:

- Assessment of segment-level truck speeds, congestion, and travel time reliability. Users can use the tool to select specific roadway segments of interest.
- Top Route Analysis for every block group in the Metro District and for roadway segments above a specific functional classification. Selecting the areas or roadway segments displays the most-used routes used by trucks to arrive at the selection and their most-used routes after departing the selection.
- Truck Activity Profiles for every block group in the Metro District highlighting hourly truck activity for medium and heavy trucks for arrivals and departures, metrics such as average trip length/speed and trip circuity, and the locations of key ODs for those truck trips.
- Previous Study Summaries including interactive maps displaying key study findings such as top freight bottlenecks and previously identified key freight routes.
- An interactive map of the current Minnesota DOT 4-year construction program highlighting projects anticipated to have an impact on freight movements.

Sections currently under development include:

- A summary of multimodal freight facilities and commodity flow information.
- A summary of freight industry employment trends by three-digit NAICS code.
- Additional analyses and maps determined through coordination with Minnesota DOT and the Study Advisory Committee.

NORTHERN ALAMEDA COUNTY TRUCK ACCESS MANAGEMENT PLAN: UNDERSTANDING FREIGHT PATTERNS WITH BIG DATA

TALIE LERNER
StreetLight Data, Presenter

Annual average daily truck traffic (AADTT) provides important information to characterize truck movements, which is critical for freight infrastructure design and performance assessment and is closely tied to highway planning, truck safety analysis, and air quality monitoring. An accurate and reliable way for calculating AADTT can be performed using data recorded by permanent counting stations or video cameras combined with vehicle classification techniques. However, stationary sensors are scarcely deployed in the United States and can only directly measure AADTT on road segments where sensors are installed. Location-based data (LBD), created by mobile devices and connected vehicles, represent an emerging data source of rich travel information. These big data sources are independent of stationary sensors, and can be collected on any road segment where people travel. As a result, they have the potential to provide traffic volume information across a large spatiotemporal road network, requiring much lower costs and labor when compared to other methods. Not only can these big data describe general traffic trends, they also aid in modeling expected traffic patterns across different classes of vehicles. The StreetLight process for developing AADTT from big data is outlined below.

There are 1,841 unique permanent-count reference sites from 33 different states across the contiguous United States were used to train machine learning models constructed using big data techniques to estimate the 2019 AADTT for three different vehicle classes: personal-use (PU) vehicles, single-unit trucks (SU) and buses, and combination-unit trucks (CU). Each category of vehicle classification was best modeled by different big data source data and contextual features (trip length, weather). Model error was lowest for PU and CU vehicle classes, which represent distinct classes, as compared to single-unit trucks and buses. Because big data techniques rely on inferences from a sample population, model accuracy was strongly related to traffic volume within each classification category.

Extensive cross-validation was performed and results show that using LBD for the estimation of AADTT can achieve a mean absolute percent error of 8% to 13% for SU, 16% to 19% for SU, and 7% to 13% for CU vehicle classes across different classifications of roads. This is comparable to the factoring methods for AADT using short-term counter data. Analyses were also conducted as part of research for FHWA's ongoing Pooled Funds Study "Exploring Non-Traditional Methods to Obtain Vehicle Volume and Class Data." Results support LBD big data as a viable option for inferring truck traffic volumes. Widespread use could save agencies resources, facilitate freight analysis and design, inform roadway safety metrics, and help build the foundation for smart communities. Continuous counters and sensors remain a critical resource for traffic engineers and are important for deriving AADTT derived from Location-Based System (LBS) big data.

ESTIMATION OF CRUDE OIL MOVEMENTS FOR THE FREIGHT ANALYSIS FRAMEWORK USING DATA FUSION

YUANDONG LIU
Oak Ridge National Laboratory, Presenter

A major hurdle in freight demand modeling and freight mobility studies has been a lack of adequate data on freight movements. The FAF integrates data from a variety of sources to create a comprehensive picture of freight movement among states and major metropolitan areas by all modes of transportation. This study focuses on the estimation of crude oil movement in the latest version of FAF. The movement between FAF zones was estimated by assembling crude oil production, attraction, and movement information from several diverse sources. Compared to previous versions of FAF, several major data sources with higher geographic resolution and coverage were identified and fused to improve the comprehensiveness of the crude oil movement estimation. These data contain production, attraction, and movement data at various geographical levels for different modes and include:

- Energy Information Administration:
  - Crude Oil Movements by Pipeline, Tanker, Barge and Rail between Petroleum Administration for Defense (PAD) Districts,
  - Refinery Receipts of Crude Oil by Method of Transportation,
  - Company Level Imports of Crude Oil, and
  - PAD District Exports of Crude Oil by Destination;
- Waterborne Commerce Statistics Center–USACE:
  - Commodity Movements from the Public Domain Database,
  - Waterway Network Link Commodity Data, and
  - Foreign Cargo Inbound and Outbound Data;
- Surface Transportation Board: Carload Waybill Sample
- Online Databases: Crude Oil Production Data at both county level and state level
- Production Data Reported by State Agencies: Crude Oil Production Data at both county level and state level.

A set of data fusion techniques were developed, based on transportation network flow theory with engineering judgement, to generate national crude oil movement data at the state and metropolitan level for domestic, import, and export flows.

The above agencies are collecting and publishing crude oil related data independently for their own purposes. By fusing these data, from both government agencies and the private sector, a more comprehensive, holistic national crude oil movement estimation can be produced. This improved movement estimation could help state energy offices and transportation planning agencies better understand crude oil flows and trends to plan for future energy management and to identify strategic network expansion.

SCAG HEAVY-DUTY TRUCK MODEL ENHANCEMENT USING MULTIPLE SOURCES OF FREIGHT DATA

STEPHEN YOON
Southern California Association of Governments, Presenter

The Southern California Association of Governments (SCAG) is the MPO for the six-county region comprising Imperial, Los Angeles, Orange, Riverside, San Bernardino, and Ventura Counties. As the MPO, SCAG is responsible for preparing a long-range transportation plan that facilitates the movement of people and goods. In developing this plan, SCAG identifies and promotes effective and practical policies and planning strategies that improve the region's transportation system while addressing challenges, such as air quality degradation that results from harmful emissions generated by goods movement activities that impact the overall quality of life for residents in the SCAG region. SCAG is required to develop and maintain an up-to-date HDT model and associated freight data, to evaluate important goods movement policy choices and related infrastructure investment strategies.

The HDT model is SCAG's primary analysis tool to support policy decisions and, therefore, requires continuous improvement to ensure that it reflects current trends in the freight industry and travel demands from diverse modes of transportation. Current model improvements are focused on clusters of key goods movement activities including, but not limited to:

- International trade flows through the region's seaports, airports, and ports of entries including border crossings;
- Domestic trade flows linking the region's manufacturing and warehousing sectors with local and national trade partners;
- Consumer goods from both domestic (including local) and international trade flows; and
- Local service, construction, and pick-up, and last-mile delivery vehicles/trucks.

To update the HDT model, SCAG plans to acquire representative freight data from available data sources, including truck GPS devices and LBS data. Current freight movement trends would be assessed through data validation procedures. Model components are anticipated to be updated with combinations of datasets as follows:

- Updates to existing secondary and tertiary truck trips from the ports to include links to the Regional Warehouse Demand Model;
- Updates to the internal HDT model data with the most recently available land use and trip purpose relationship data to represent the continuously changing market environment;
- Data updates to improve the regional truck vehicle-miles-traveled (VMT) distribution rate by county and air basin;
- Adds micro-level commodity details in the internal HDT model to allow for evaluating how corridor performance ultimately affects the supply chains of specific industries.

These data enhancement activities will result in the following critical improvements: 1) improved representation of the international trade sector by developing better techniques for integrating secondary and tertiary truck trips associated with port cargo movements (forecast is generated by the Port Transportation Analysis Model (PortTAM)); and 2) improved performance of the internal HDT model by focusing on intra-regional trade trends and supply chains associated with updated information about regional distribution centers, warehouses, and manufacturing industries based on updated land uses.

APPLIED FREIGHT FLUIDITY FOR STATE DOT DECISION-MAKING

NICOLE KATSIKIDES
Texas A&M Transportation Institute, Presenter

This research demonstrates ways that state planners, operators, and lead decision-makers can visualize freight fluidity—the trips that goods make from origin to destination and where bottlenecks occur—to support decisions for programming and operations. It includes multimodal analytics that have not been typically included in fluidity analytics before. This work expands on freight fluidity efforts in Canada, and by FHWA, Maryland, Colorado, and Texas, to develop applied resources that public sector staff can use to understand how freight moves. Specifically, this research develops approaches for: 1) economic analysis (understanding trading partners and economic development opportunities and key freight routes); 2) framing fluidity analysis (determining what is important to measure); 3) using relevant tools and resources to assess freight fluidity; and 4) connecting highways to multimodal trip legs at ports and intermodal facilities. This work further describes approaches by demonstrating how to do these analyses quickly or in a hurry yet still be defensible. This is important for public sector analysts who are often asked to justify projects or requests with few resources. This research describes the most beneficial processes for understanding freight bottlenecks and translating that information into simple, readily understandable visuals and information that can be used in discussions about how to address problems through capital or operational investments. This research also demonstrates advances in freight fluidity analysis by demonstrating how to combine highway data with port and vessel mobility data sources to present a multimodal view of freight flows, bottlenecks, and impacts. This work can improve mobility on the nation's transportation network by providing a synthesis or consolidation of freight fluidity approaches in addition to taking fluidity a step further and adding in multimodal elements that have not been previously included in existing fluidity analytics. The goal is to help public sector planners, operators, and decision-makers incorporate freight fluidity into their day to day work so that freight can be considered easily and intuitively.

PILOT INTEGRATION OF BIG DATA INTO FREIGHT TRANSPORTATION PLANNING AND OPERATIONS

DAN SEEDAH
Jacobs, Presenter

State and other government planning agencies are exploring big data sources to facilitate better decision-making in freight transportation planning and operations. Limited funds, differing planning horizons, proprietary information, insufficient human resources, and misinformation about available data have resulted in a disconnect about the acquisition, integration, and use of freight-related big data. As part of NCHRP Project 08-119, Data Integration, Sharing, and Management for Transportation Planning and Traffic Operations, a pilot product is being developed to showcase how big data from the private sector can be acquired, integrated, and used to address a limited number of freight planning and operations use cases. The pilot product 1) provides a list of identified state and other government stakeholders' needs and use cases, 2) identifies which big data sources are available from vendors, and 3) provides recommendations on how the data can be integrated and used to address the selected use cases. This pilot product was showcased at the 2021 TRB Innovations in Freight Data Workshop for feedback and comments from conference participants before final deployment.

APPENDIX A

List of Abbreviations and Acronyms

AADT	Annual Average Daily Traffic
AADTT	Annual Average Daily Truck Traffic
AI	Artificial Intelligence
AIS	Automatic Identification System
ALE	Accumulated local affects
API	Application Programming Interface
ATRI	American Transportation Research Institute
BEA	U.S. Bureau of Economic Analysis
BOL	bill of lading
BTS	U.S. Bureau of Transportation Statistics
CFS	Commodity Flow Survey
CMAP	Chicago Metropolitan Agency for Planning
CRIS	Crash Record Information System
CSTDM	California Statewide Transportation Demand Model
CTC	California Transportation Commission
CU	Combination-unit trucks
DOT	Department of Transportation
EMFAC	EMission FACtor
FAF	Freight Analysis Framework
FFSP	Freight Forecasting and Scenario Planning
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carriers Safety Administration
GBM	Gradient Boosting Machine
GIS	Geographic Information System
GPS	Global Positioning System
GVW	Gross Vehicle Weight
GVWR	Gross Vehicle Weight Rating
HDT	heavy-duty trucks
IANA	Intermodal Association of North America
IoT	Internet of Things
KNN	K Nearest Neighbor
LBS	location-based system
LDV	light-duty vehicle
lidar	light detection and ranging
LPMS	Lock Performance Monitoring System
MDT	medium-duty trucks
MPO	metropolitan planning organization
NAICS	North American Industry Classification System
NCHRP	National Cooperative Highway Research Program
NHTS	National Household Travel Survey
NJTPA	New Jersey Transportation Planning Authority

NLP	Natural Language Processing
NPMRDS	National Performance Management Research Dataset
OD	Origin-Destination
OKI	Ohio-Kentucky-Indiana Regional Council of Governments
OS/OW	Oversize/Overweight
PAD	Petroleum Administration for Defense
PortTAM	Port Transportation Analysis Model
PU	personal-use vehicles
RHiNo	Roadway–Highway Inventory
RITIS	Regional Integrated Transportation Information System
SAM	Statewide Analysis Model
SCAG	Southern California Association of Governments
SDCI	Seattle Department for Construction and Inspection
SELA	Southeast Los Angeles Collaborative
SPaT	Signal Phase and Timing
SU	single-unit trucks
TAZ	transportation analysis zone
TCAT	Truck Congestion Analysis Tool
TCU	truck, communication, utilities, and waste
TEU	20-ft equivalent units
TIMS	Transportation Injury Mapping System
TMAS	Travel Monitoring Analysis System
TMC	Traffic Management Center
TPAS	Truck Parking Availability System
TRB	Transportation Research Board
TSIS	Traffic Signal Information System
TTI	Texas A&M Transportation Institute
TTMS	Telemetered Traffic Monitoring Sites
TWMS	time window management systems
TxPROS	Texas Permitting and Routing Optimization System
UPS	United Parcel Service
USACE	United States Army Corps of Engineering
USPS	United States Postal Service
UTP	Unified Transportation Program
V2X	vehicle-to-everything
VMT	vehicle miles traveled

APPENDIX B

Web Links Provided During the Workshop

The following is a list of links to data, dashboards or additional resources that were provided during the workshop either in presentations or the chat box.

Agricultural Transportation Open Data Platform: https://agtransport.usda.gov/

Agricultural Transportation Research webpage: https://agtransport.usda.gov/stories/s/Agricultural-Transportation-Research/sqke-p23b/

BTS VIUS webpage: https://www.bts.gov/vius

Caltrans PeMS Traffic Data: https://pems.dot.ca.gov/

Central Ohio River Information System: https://centralohioriverbusinessassociation.com/coris/

CMAP land use data: https://datahub.cmap.illinois.gov/dataset/land-use-inventory-for-northeast-illinois-2015

Fact #621: May 3, 2010 Gross Vehicle Weight vs. Empty Vehicle Weight: https://www.energy.gov/eere/vehicles/fact-621-may-3-2010-gross-vehicle-weight-vs-empty-vehicle-weight

FAF Payload factors https://ops.fhwa.dot.gov/publications/fhwahop20011/chap12.htm

Florida DOT Parking Supply Dashboard: https://hdr.maps.arcgis.com/apps/webappviewer/index.html?id=1e827e785913431982b52104646fd20d

Florida DOT Truck Parking Resources: https://www.fdot.gov/rail/studies/truck-parking

Freight Finder: https://www.freightfinder.com/

GeoTab: https://www.geotab.com/

IMPLAN – Economic Model: https://implan.com/

LA Port Optimizer: https://www.wabteccorp.com/digital-electronics/network-logistics/port-optimizer

Location Utilization Dashboard: https://hdr.maps.arcgis.com/apps/opsdashboard/index.html#/623116a8deeb436bbb37e32e5b807aea

Market Data: EMSI: https://www.economicmodeling.com/data/

Municipal Freight Scans: http://www.planning.ri.gov/planning-areas/transportation/freight-scans.php

Oregon Commercial truck Parking Study, 2020:

Oregon Commercial truck Parking Study, 2020:
https://www.oregon.gov/odot/Projects/Pages/Commercial-Truck-Parking-Study.aspx

PIERS: https://ihsmarkit.com/products/piers.html

Pilot Integration of Big Data: https://transportstudies.org/apps/freight_big_data/index.html and
https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4543

RITIS: https://ritis.org

Statewide Truck GPS Analysis: ArcGIS Story Map:
https://hdr.maps.arcgis.com/apps/Cascade/index.html?appid=b760ec83e85544b899724bd3910af45b

TCAT: https://tcatwebprod.z14.web.core.windows.net/

Texas Annual Truck Congestion Report Cards:
https://tcatanncongrep.z14.web.core.windows.net/anncongrep/dashboard

Texas Top 100 Congested Roadways: https://mobility.tamu.edu/texas-mostcongested-roadways/

Title 13 CFS: https://bit.ly/3llzocl

Transearch: https://ihsmarkit.com/products/transearch-freight-transportation-research.html

USDA: https://agtransport.usda.gov/

Verification, Refinement, and Applicability of Long-Term Pavement Performance Vehicle Classification:
https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/ltpp/13091/002.cfm

TRANSPORTATION RESEARCH BOARD
500 Fifth Street, NW
Washington, DC 20001