Image: (Learntek/www.flickr.com/photos/153724200@N07/39835232960)

TPB Big Data Purchases

Zhuo Yang
Transportation Data Analyst

Systems Performance, Operations and Technology Subcommittee
February 15, 2024

Performance Measures and Conventional Data Collection

- Regional travel is frequently measured in trips and trip characteristics:
  - Person Trips, Vehicle Trips, Vehicle Miles of Travel (VMT)
  - Volumes
  - Trip Origin/Destination (O/D)
  - Trip Mode
  - Trip Purpose

- Historically, these data are collected through traffic counters (humans and/or machines) and travel surveys.
  - Conventional methods are costly, take time to implement, and are not updated frequently as a result.

What is Big Data?

Defined:

Big Data is . . . an information asset with such high volume, velocity, and variety that specific technology and analytical methods are required for its transformation into value.[1]

For transportation, sources can include:

- Passively collected data from mobile applications, including GPS traces and location-based services
- Unmanned aircraft/space-based radar used to monitor traffic flow
- On-board vehicle sensors
- Traffic sensors and cameras
- Smart card data

[1]De Mauro, Greco, Grimaldi, (2016) "A formal definition of Big Data based on its essential features," Library Review, Vol. 65 Issue 3, pp. 122-135, https://www.emerald.com/insight/content/doi/10.1108/LR-06-2015-0061

Big Data's Potential and Limitations

- Big Data can **provide data more frequently,** enabling more frequent analytics and decision support
- **"Black Box" nature of Big Data is a concern**—agencies/users unable to control changing data collection/processing methodologies, compromising potential for longitudinal comparisons
- **Very costly**—it is a risk to invest in a Big Data product without full knowledge of black box or control over the methods
- Data quality is **difficult to verify**

Big Data Evaluation

- **COG/TPB** hired an independent consultant, **Kimley-Horn**, to conduct an independent evaluation of Big Data products
- Purpose: Better understand (1) the state of the practice in applying Big Data in transportation planning and analysis and (2) which Big Data products have the potential to meet COG/TPB's programmatic needs
- Identify areas where COG/TPB can broaden its use of Big Data
- Completed in 2020

Big Data Evaluation - Scope

- Establish Study Work Group (membership, responsibilities, and meeting schedule) and study work plan
- Develop understanding of COG/TPB programmatic requirements and analytical/modeling processes
- Review state of the practice of Big Data use and applications by other MPOs
- Conduct an independent evaluation of Big Data sources for their potential in supporting TPB staff in meeting its programmatic requirements
- Recommend options and considerations for acquiring Big Data
- Prepare a final report

Data Needs & Research Areas

	Research Area	Description
1	Travel Demand Forecasting	• Origin-Destination (O-D) information by mode with trip purpose/destination/origin type • Region-wide speeds and volumes • Region-wide signal delay estimation/queue length
2	Travel Demand Management	• Estimating network demand using historical data • Survey of employer telework policies • Forecasts of future telework policies
3	System Performance/ Congestion Management	• Monitoring network-wide congestion and performance • Detailed network performance data for strategy development • Event traffic data
4	Transit and Active Travel	• Impact of mode choice on network level of service (LOS) (scenario analysis) • Monitoring real-time transit performance • Monitoring active travel demand and delay

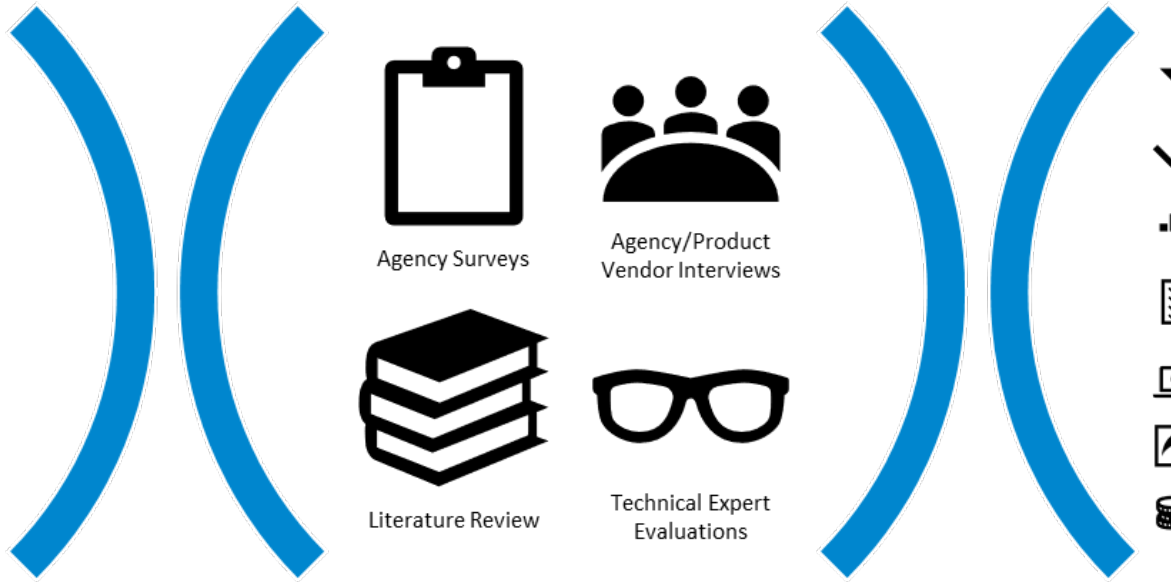
Data Needs & Research Areas (cont.)

Research Area		Description
5	Transportation Network Companies	• Improved understanding of usage and demand for TNCs: temporally, geographically, O-D patterns • Socioeconomic and demographics of TNC riders • Relationship between TNCs and transit (e.g., first/last mile problem)
6	Traffic Counts	• Obtaining permanent/temporary traffic count data • Variability analysis of traffic counts (e.g., day-by-day) • Methods to validate traffic counts collected in traditional methods • Vehicle classification data
7	Connected and Automated Vehicles	• Impact of CAVs to vehicle occupancy • Impact of CAVs on multimodal travel demand • Impact of CAVs on infrastructure (e.g., parking) and revenue streams • Available data from CAVs
8	Other Research Areas	• Regional freight and commercial vehicle travel forecasting • Household survey data validation and supplementation • O-D patterns based on population demographics • Improved project selection

Big Data Evaluation Methodology

- Travel Demand Forecasting (TDF)
- Travel Demand Management (TDM)
- System Performance and Congestion Management Process (CMP)
- Transit and Active Travel
- Transportation Network Companies (TNC)
- Traffic Counts
- Connected and Automated Vehicles (CAV)
- Other Research Considerations

Consider TPB research areas



Evaluate value by research area from **4 input sources**

- Applicability for Research Area
- Data Reliability and Validity for Research Area
- Data Coverage
- Technical Staff Competencies and Training Requirements
- Information Technology Requirements
- Data Sharing Restrictions
- Data Cost

Score each product based on **7 evaluation criteria**

Source: Kimley-Horn

Report Components

SURVEY & INTERVIEW
OF MPO AGENCIES

EVALUATION
OF BIG DATA PRODUCTS

DEVELOPMENT OF A
PROCUREMENT PLAN

Source: Kimley-Horn

What's Next?

4 location-based Big Data products were identified to meet half or more of TPB's programmatic requirements

Kicking the Tires! TPB is acquiring 1-year subscriptions for all 4 to conduct analysis and evaluate

- Establish internal Big Data user group
- Determine which Big Data product to support TPB's requirements over the long term.

Zhuo Yang

Transportation Data Analyst
(202) 962-3370
zyang@mwcog.org

mwcog.org

777 North Capitol Street NE, Suite 300
Washington, DC 20002