POTENTIAL APPLICATION OF GOOGLE-BASED TRAVEL TIME DATA IN TRANSPORTATION PLANNING ANALYSES

Feng Xie
TPB Transportation Engineer

Dusan Vuksan
TPB Transportation Engineer

Travel Forecasting Subcommittee
November 15, 2019

Introduction

- TPB member agencies occasionally seek travel time information in support of travel shed/accessibility analyses from or to Points of Interest (POIs):
  - How far can my residents travel in 30 minutes from their homes in rush hour?
  - How many employees can reach jobs in my activity center within 45 minutes?
- In technical terms, planners are usually looking for travel time information between a zone of interest and all other zones in TPB Modeled Area.
- Travel demand model data are usually a great source for this type of information at the regional level, and for alternatives analysis purposes.
- However, as regional travel models are not validated at the TAZ level (or "POI level"), TPB staff were interested in finding other data sources that would provide supplemental information.

Introduction (Cont.)

- TPB staff first examined other available non-model-related data resources:
  - Historical arterial travel time data gathered by cars travelling in traffic (very important resource used in the past, but has been discontinued a while back)
  - Commercial probe-based traffic data from Vehicle Probe Project (VPP) and INRIX (extremely valuable resource used in the Congestion Management Process, but TAZ-to-TAZ travel time information would require additional data processing)
- Recent rapid development of crowdsourced "Big Data" and "Big Data" tools provide an alternative source of travel time information, which is the focus of this presentation.

Crowdsourced Big Data

- Crowdsourced traffic data are provided from various sources. Many Internet companies now gather real-time traffic data through proprietary probe-vehicle networks and make them available online:
  - Google: Google Application Programming Interface (API)
  - Waze: Waze Transport Software Development Kit (SDK)
  - Uber: Uber Movement
- Two Google APIs currently provide travel time information:
  - Directions API: calculate step-by-step directions between specified locations
  - Distance Matrix API: provides travel distance and time for a matrix of origins and destinations based on recommended routes
- Quality of travel time data from Google APIs and similar portals has been examined in several studies, but additional research and validation are still needed.
  - Core of Rosslyn Transportation Study by Arlington County (presented at TFS in March 2019)

Google Distance Matrix API

- TPB exploration focused on Google API due to its data coverage and application-readiness.
- Distance Matrix API was used (as opposed to Directions API) as detailed route information was not the focus of this study.
- Zone-to-zone driving and transit times were requested in four steps:
  - Create personal developer account on Google Cloud Platform;
  - Obtain project key to use Distance Matrix API;
  - Generate latitude/longitude coordinates of TAZ centroids;
  - Make travel time requests to Distance Matrix API in Python;
- Important parameters specified in the travel time requests included: origin, destination, travel mode, departure time and traffic model.
- Cost of Google Distance Matrix API service is calculated using a pay-as-you-go pricing model by usage, multiplying price per each use.

QA/QC Analyses

- QA/QC analyses by TPB staff verified the reasonableness of requested Google travel time data but also indicated potential discrepancies:
  - "Zero-Result" Returns
  - Comparison to Google Maps travel times
  - Temporal variation of Google travel times
  - Driving paths in corridors with tolling facilities

Figure 1. The Time-of-Day Variation of Driving Times (in Minutes) to Crystal City Neighborhood (TAZ #1501) (Departing in AM on Wednesday, 4/24/19)

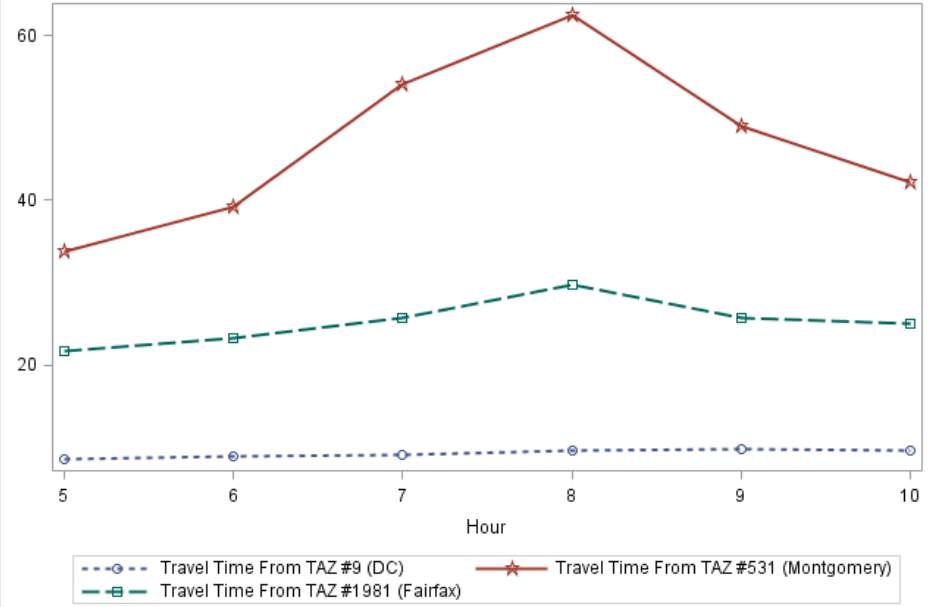
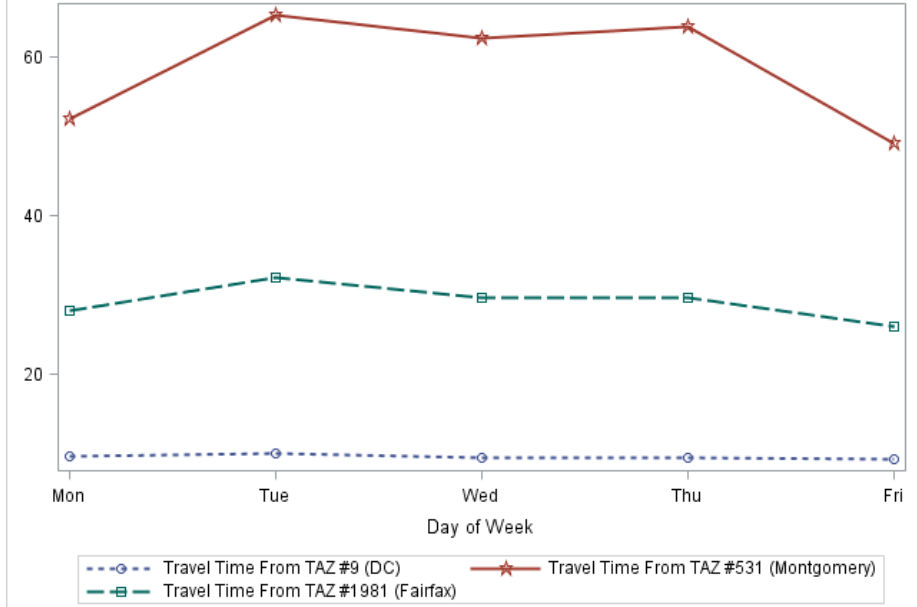


Figure 2. The Day-of-Week Variation of Driving Times (in Minutes) to Crystal City Neighborhood (TAZ #1501)(Departing at 8:00 AM in the 4th week of April 2019)



Google Driving vs. Transit Times

- Driving times and transit times from Google API are likely developed using different methods:
- Driving times:
  - "Observed" data developed based on historical averages of crowdsourced cellular phone information
  - Able to reflect real-time in-traffic congestion.
- Transit times:
  - "Estimated" times on planned transit routes developed using a public transportation planning tool based on General Transit Feed Specification (GTFS) inputs.
  - Do not reflect in-traffic congestion.
  - Do not consider transit trips with park and ride (drive) access.

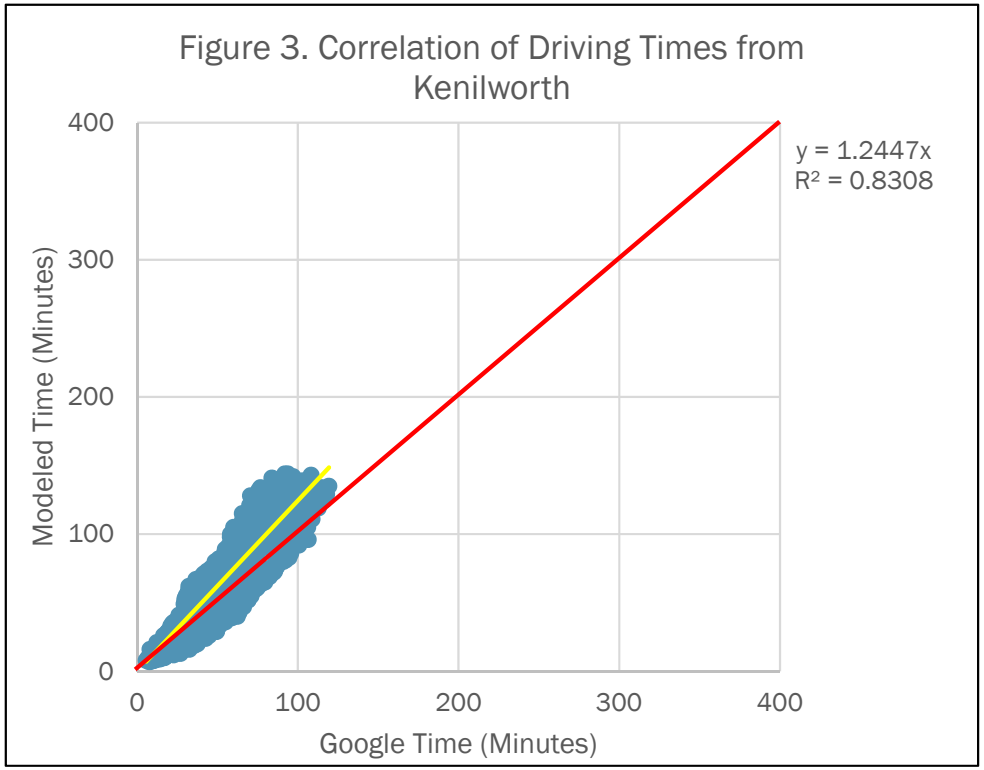
Case Studies

- Staff applied Google travel time data to travel shed/accessibility analyses for six neighborhoods in the metropolitan Washington region.
- Primarily, the analysis focused on developing 45-minute travel sheds.
- Analysis also compared Google and modeled travel times and their corresponding travel sheds.
- Six case studies led to similar findings, so only the analysis conducted for the Kenilworth neighborhood (Equity Emphasis Area) will be presented.

Driving Times from Kenilworth

- Google driving times were requested for Wednesday April 24, 2019 departing at 8 AM, while modeled 2019 AM Peak driving times were derived from TPB's Ver. 2.3.75 model with "Visualize 2045" inputs.
- Google and modeled driving times from Kenilworth are highly correlated (R-Squared=0.83).
- Modeled driving times from Kenilworth are higher than Google times (Slope=1.24).
  - Volume-delay functions
  - Other model factors such as Potomac River Crossing Penalty

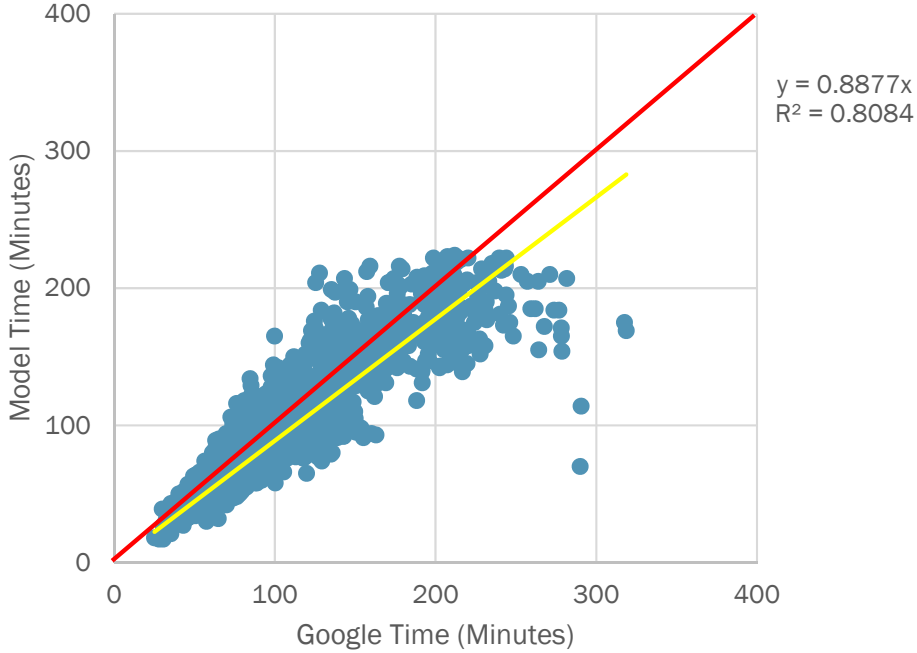
Figure 3. Correlation of Driving Times from Kenilworth



Walk-Acc Transit Times from Kenilworth

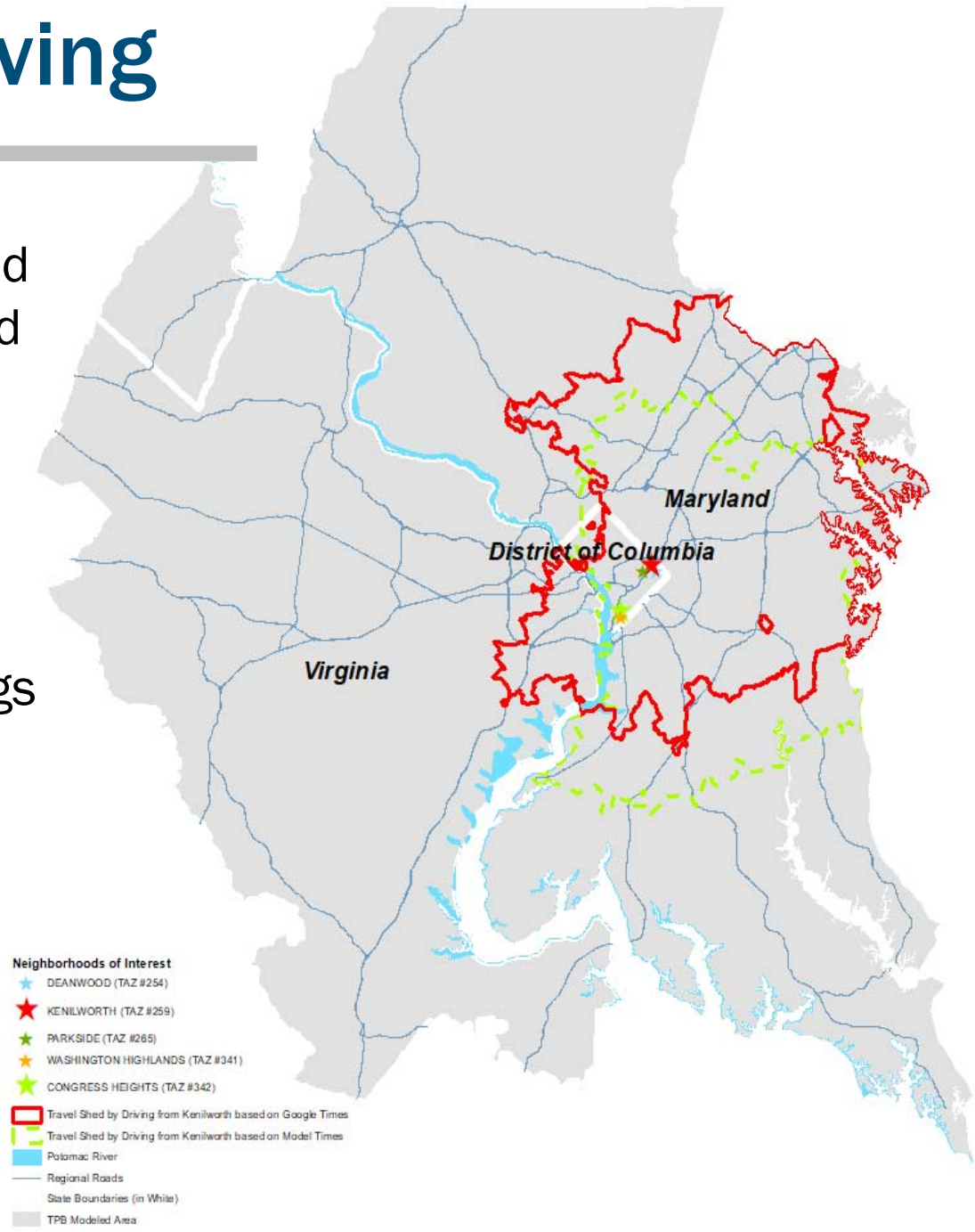
- As Google transit times do not consider drive-access transit, walk-access transit times from the model were extracted for a better comparison.
- A small group of outlier zones with extremely long Google transit times indicated a potential deficiency of Google's transit path building algorithm.
- Consistent with transit path building process in the model, outliers were removed by capping transit times at 360 minutes.
- Google and modeled transit times are also highly correlated (R-Squared=0.81) and statistically more comparable (Slope=0.89).

Figure 4. Correlation of Walk Access Transit Times from Kenilworth



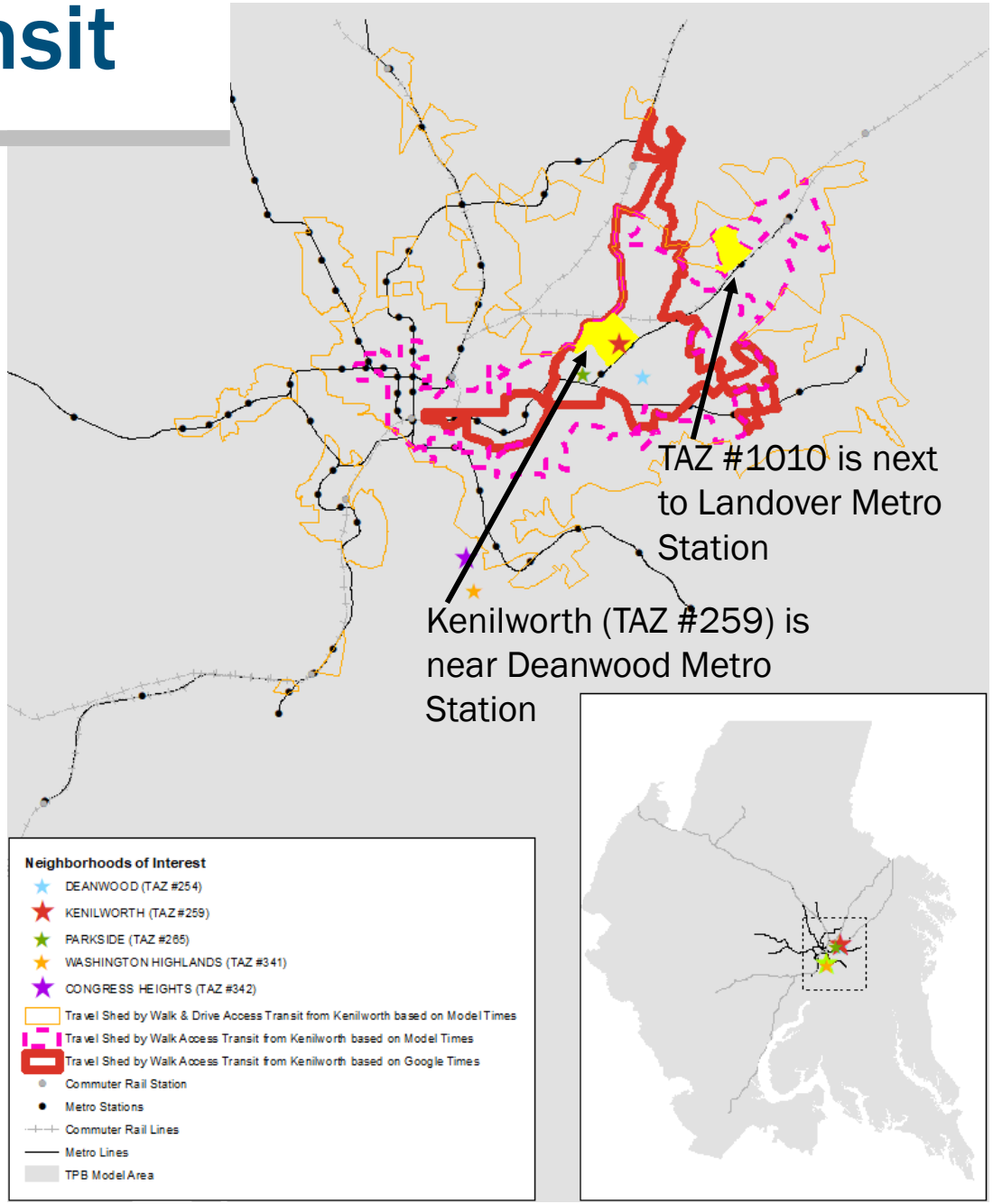
Travel Sheds by Driving

- 45-minute travel sheds from Kenilworth were developed based on Google (red line) and modeled (green line) driving times.
- Travel shed based on modeled times is largely confined to DC and Maryland.
  - Low speeds on river-crossings
  - Potomac River Crossing Penalty
- Travel shed based on Google times appears to be more realistic showing some accessibility in Virginia.



Travel Sheds by Transit

- Transit travel sheds were developed based on Google (solid red line), modeled walk-access (dashed purple line) and modeled drive & walk access (solid orange line) transit times.
- Excessive walk times were found on the Kenilworth end in the Google data, as TAZ #259 centroid is located in Kenilworth Aquatic Gardens, indicating a need to revisit centroid location coding.
- TAZ #1010 is outside Google travel shed as the centroid is separated from Landover Metro by US 50, but is inside modeled travel shed due to airline-distance walk access links in the model.



Key Findings

- Travel times requested from Google API show reasonable temporal variations that reflect in-traffic congestion and transit schedule changes.
- While Google driving times are "observed" data developed based on historical cellular phone information, Google transit times are "estimated" using a public transportation planning tool.
- Google and modeled driving/transit times are highly correlated.
- Modeled driving times are higher than Google times, mainly due to volume delay functions, river-crossing penalty and other factors.*
- Google and modeled transit times are statistically more comparable, as both were estimated times using the same or a similar set of GTFS input data for this region based on published schedules.

* Please refer to Roden, David B. "Assistance with Development and Application of the TPB Travel Demand Model." presented at the July 19, 2013 TFS meeting.

Key Findings (Cont.)

- Driving times from Google API may be applied to travel shed/accessibility analysis. Case studies indicated that travel sheds developed based on Google times showed reasonable coverage.
- Transit times from Google API should be used only in analyses where drive access transit is not a critical factor, e.g., environmental justice studies that focus on low-income residents with low vehicle availability; if cost is an issue, modeled walk access transit times provide a comparable alternative.
- Analyses of Google travel time data indicated potential discrepancies associated with its driving path building (e.g., paths related to toll facilities) and transit path building (e.g., unrealistically long trips) processes.
- The "black-box" nature of Google travel time data is a concern for practitioners as Google does not disclose substantive details on the algorithms it uses for developing the data.

Conclusions and Next Steps

- Relatively easy access to Google API makes it an attractive commercial source of travel time information for practitioners.
- Similar to many other Big Data, the validity of Google travel times should not be assumed without a comprehensive comparison to ground-truth data.
- Based on the testing by TPB staff in limited capacity, Google Distance Matrix API provides reasonable travel time data and may find applications in transportation planning analyses.
- As the next steps, TPB staff may:
  - Explore the possibility of using the travel demand model in conjunction with Google Distance Matrix API to forecast future-year travel sheds, since Google does not provide future-year travel time forecasts;
  - Explore the possibility of developing a matrix of travel times for the modeled area using the Google data;
  - Explore Google travel times by non-motorized modes;
  - Explore the use of Google Directions API in subarea- or corridor-level studies.

Acknowledgements

- James Li, Daivamani Sivasailam, Erin Morrow, Mark Moran, COG/DTP staff members
- Kanathur Srikanth, DTP Director

Feng Xie

TPB Transportation Engineer
(202) 962-3259
fxie@mwcog.org

Dusan Vuksan

TPB Transportation Engineer
(202) 962-3279
dvuksan@mwcog.org

mwcog.org/TPB

Metropolitan Washington Council of Governments
777 North Capitol Street NE, Suite 300
Washington, DC 20002

National Capital Region
Transportation Planning Board