Work Order # XX

Activity Based Model Development/Performance Management

Big Data Acquisition

Scope of Services

Version 5/4/2020

I. Purpose

The Regional Transportation Technical Advisory Committee – Modeling Subcommittee (RTTAC-MS) is looking for Big Data sources and their applicability in evaluating regional travel behavior and in estimating, calibrating, and validating regional travel demand models. The data set should also lend itself for uses related to federal performance management measures and transportation planning activities. The Committee is looking to award a contract to purchase a Twelve (12) month data set from a qualified firm who can provide a Big Data set to support the aforementioned activities.

A. As a part of this effort the RTTAC-MS will evaluate proposals based on criteria that will generally fall into the following three (3) broad categories:
    1. Review and comparison of how other metropolitan areas have used the proposed Big Data set in similar applications
    2. Identification and analysis of potential Big Data uses, including strengths, limitations, and appropriateness for various types of analytical studies
    3. Recommendation of whether the RTTAC-MS should pursue a Big Data set purchase and, if so, provide a corresponding strategy to invest in Big Data to support future transportation planning studies and analyses as well as support for regional travel demand forecasting activities.
B. The RTTAC-MS is ultimately interested in whether Big Data analytics can be used to appropriately examine travel behavior and mobility in the South Florida region, and gain insight on the strengths and limitations of Big Data sources in conducting mobility analyses.
C. Proposals are intended to provide the RTTAC-MS with necessary detailed information and understanding of the strengths, limitations, applicability, and acquisition costs of the proposed Big Data set before committing to an investment in Big Data. The RTTAC-MS seeks to determine the feasibility of using the Big Data set to:
    1. Analyze regional travel and mobility to help validate data the region collected through conventional survey and traffic monitoring programs, and to provide additional information and understanding of aspects of mobility that are not easily captured through conventional programs.
    2. Estimate, calibrate, and validate its regional travel demand model.
    3. Fulfill federal mandates for a regional Congestion Management Process (CMP).
    4. Understand the characteristics and impacts of emergent travel options and technologies on travel throughout the region, including, but not limited to, the rapid rise in the use of Transportation Network Companies (TNCs), such as Uber and Lyft, traffic information and route navigation apps such as Google Maps and Waze, and real-time arrival and trip-planning for transit, ridesharing, bike sharing, etc.

II. Background

The RTTAC-MS is made up of the three Metropolitan Planning Organizations of the Miami Urbanized Area (Broward Metropolitan Planning Organization, the Miami-Dade Transportation

Planning Organization and Palm Beach Transportation Planning Agency) and Districts Four and Six of the Florida Department of Transportation (PARTNERS). The RTTAC-MS is a subcommittee of the Southeast Florida Transportation Council (SEFTC). The PARTNERS share the responsibility of developing, updating and maintaining the Southeast Florida Regional Planning Model (SERPM). The SERPM is used by the three MPOs in the development of their respective Metropolitan Transportation Plans and is also used by FDOT Districts Four and Six for multimodal projects.

One of the critical elements in the development and update of the SERPM model is current travel behavior data. As this is an ever-evolving field with technological advances in general and urban infill trends within the Tri-County Region (Broward, Miami Dade and Palm Beach Counties), the PARTNERS recognize the need for a robust and multi-dimensional data collection effort to continue to enhance the SERPM model through better understanding the Tri-County Region's travel patterns and characteristics. The PARTNERS also acknowledge that the data collected may be able to support activities beyond transportation-related functions.

The Southeast Florida Regional Travel Study was completed in 2017 and represented a major effort for the region to collect information on household characteristics and travel activity on selected sample household's representative of the Tri-County Region. The findings from the household survey were instrumental in the development of SERPM Version 8. Additionally, the Regional Travel Study collected global positioning systems (GPS) based travel data showing origin-destination, system and freight movement at the Traffic Analysis District (TAD) level. This data was very useful for the PARTNERS to understand travel behavior and supplemented the household trip making information obtained from the Regional Travel Study. It is the intent of the PARTNERS to further explore the utility of these "big data" sources of location-based data to support the enhancements and the development of SERPM and other transportation planning activities.

The PARTNERS intend to use state-of-the-art and innovative techniques for data collection that will lead to more cost-effective means of collecting vital travel behavior information. Innovative techniques refer to methods such as usage of mobile device data for anonymously identifying origin-destinations or usage of existing National Household Travel Survey (NHTS) data to supplement and potentially supplant some of the traditional methods of identifying household travel characteristics.

III. Proposal Format

Section 1 – State-of-the-practice review of the use of the Big Data set to analyze mobility and travel behavior by other Metropolitan Planning Organizations.

A. Proposal shall include the following evidence of other metropolitan areas' experience using the proposed Big Data set to support mobility and travel behavior analyses and regional travel demand modeling activities.
    1. Documented evidence of Origin-Destination matrix outputs from a travel demand model where the provided data set improved the model output results,

2. Quantified examples and explanations of the modifications made to the model inputs that improved the calibrated model outputs, and
3. Documentation and/or letters of support from MPO/Consultant(s) that the provided data set improved model results.

B. The metropolitan areas examined should be comparable to the South Florida region in size, population, population diversity, land mass, complexity of regulatory environment, or other appropriate measure(s) that would provide appropriate insights for the Region.
C. The analysis should identify the following:
    1. Evidence of MPO's using/used the proposed Big Data set,
    2. How the data set was applied, including what tasks (model calibration, congestion monitoring etc.) it helped accomplish and the quantifiable effectiveness,
    3. Whether the data set was used, or could be used, to replace existing data collection/analysis methods and programs,
    4. Whether the data set was used, or could be used, to supplement and validate existing methods and programs,
    5. Whether the data set was used as a part of a first-time methodology to address a new requirement or planning study,
    6. Lessons learned or changes in course of action in applying the data set to regional planning analysis,
    7. What the strengths, limitations, reliability, and appropriateness of the data set are and what the various analyses and applications can be used for,
    8. Methodology conveying any remedy in data limitation, bias, any over representation.

Section 2 – Key items to be addressed in the proposed data set relating to regional travel and mobility as well as travel demand modeling.

The data should depict people, freight and goods, vehicles movements, and trip origins and destinations in the Tri-County region at a Traffic Analysis District (TAD) and/or Traffic Analysis Zone (TAZ) level to provide insights and/or inferences on the following metrics and address the below considerations:

A. Travel Demand Modeling – *Person and Vehicle Origin-Destination*
    1. How are the Origin-Destination (O-D) patterns obtained; and by what modes is the data available?
    2. How is the data set generated and cleaned?
    3. What is the sampling frequency on the data set and how does the data set address the adverse effects on GPS positional accuracy caused by dense urban environments?
    4. What is the ability of data to obtain geographic O-D for trips that use a specific road segment?
    5. How does the data set determine mode choice behavior? – For example can data identify the destination choice and mode choice for areas with high concentrations of low-income and/or minority residents?

6. How does the data set distinguish between transit and auto modes? And how is this validated?
7. Can the data set identify First and Last Mile of persons, goods delivery, and door service delivery?
8. Does the data set quantify Pedestrian, Bike, and Micro-mobility trip making characteristics and patterns?
9. How are bicycle trips identified and validated?
10. Can the data give insights into understanding interactions between traffic, traffic congestion, and transit bus operations?
11. Does the data set provide travel time and speed Data (by time of day)?
12. How is information about vehicle travel speeds, time, and volumes obtained and/or validated?
13. Does the data set include understanding of the dynamically priced express toll lanes system in the region, including who is using them?
14. Does the data set understand and differentiate the volume breakdown between the adjacent general-purpose lanes and dynamically priced lanes as well as the magnitude of changes in traffic volumes in the general purpose and dynamically priced lanes that result in High Occupancy Toll (HOT) lane price changes?
15. Does the data set identify trip purpose and/or type of destination?
16. Does the data set include estimates of through travel and external travel (external-to-internal and internal-to-external), including external transit travel, which is not currently included in the travel demand forecasting model?

B. Transportation Network Companies (TNCs) – *Ride Hailing Usage*
    1. Does the data set quantify the magnitude and O-D patterns on TNC trips?
    2. Can the data set help with assessing the existing and forecasting the future relationship between TNCs and transit travel (i.e., understanding where TNCs support/enhance transit use and where TNCs compete with/supplant transit travel)?
    3. Can the data set inform how and where (across the region, by land use, by residential and commercial density) TNCs are influencing Vehicle Miles of Travel (VMT)?
    4. Does the data set provide temporal distribution of TNC ridership (in rush hours, weekday/weekend, monthly, and quarterly)?
    5. Does the data set give insights into the demographics of TNC riders in the region, including but not limited to income, race/ethnicity, and gender?

C. Travel Demand Management (TDM)
    1. Can the data set aid in evaluating the influence of employer-based TDM policies on travel behavior (e.g. obtaining more insight about the influence of teleworking policies on trip generation)?
    2. Does the data set support TDM Model Calibration?
    3. Can the data set estimate the current and forecast future share of the workforce that are telecommuting on a given day?

D. Connected Autonomous Vehicles (CAVs) – *Autonomous Vehicle Usage*

1. Does the data set have any information on O-D information about CAVs, both personally owned and fleet vehicles?
2. Can the data set inform any insights on the impacts of CAVs on parking, including revenue generation?
3. Any general availability of data on CAVs.

E. Traffic Counts
   1. Does data set provide any permanent/temporary/hourly traffic count data?
   2. Does data set use traffic count data collected by State DOTs, or related partners, for data validation?

F. System Performance/Congestion Management Process
   1. Can the data be used for monitoring and/or evaluating transportation system performance?
   2. How can the data help with understanding the impacts of congestion management strategies?
   3. Can the data help with recognizing trends in new congestion management strategies and begin to predict/track their impacts?
   4. Can the data set help with understanding impacts of specific events, such as major traffic incidents, severe weather events, or major scheduled special events, as well as the impacts of strategies employed to manage transportation impacts of these events?

G. Other Considerations
   1. Does the data set include regional freight and commercial vehicle travel?
   2. Intercity bus or any transit travel.
   3. Can the data set help with validating data collected from household travel surveys and OD transit surveys?
   4. Can the data set be integrated with household travel surveys to provide both long distance and local travel behavior data?
   5. Does the data set allow for real-time monitoring of traffic/transit/rail data?
   6. How can the data set proposed be used for transportation related project selection?
   7. How can the data set help in understanding how the use of alternative modes for commuting (bike, walk, rideshare, transit, car/vanpool, first/last mile, scooter, telework) affects the overall network in terms of major highways and arterials and how the level of service (LOS) would look if the individuals using those modes drove alone instead?
   8. How does the data set address the inherent bias towards younger more affluent commuters by using smartphone data?
   9. How is data filtered, corrected and validated?

Section 3 – Costs and Technical Requirements

Proposal should also include the following information:

A. The commitment that would be required of the RTTAC-MS to procure Big Data—cost breakdowns of data purchase/subscriptions options.
B. The ability/limitations of the RTTAC-MS to partner and/or share data with other stakeholders, including licensing restrictions.
C. Minimum IT requirements to adequately store, process, query, analyze, and document results either locally or in cloud-based platforms.
D. Technical staff competencies required to manage and conduct Big Data analyses.
E. Training requirements to ensure continuity of staff competencies to manage and conduct Big Data analyses.
F. Assurances of a consistent methodology used to develop and maintain Big Data over the long term (years, possibly decades) to ensure the integrity of time-series analyses.
G. Proposal should also be cognizant of partnering opportunities that may exist with other agencies to leverage resources and gain efficiencies. For example, if one agency already subscribes to a recommended Big Data source but that license only covers a portion of the region, then what incremental investment might be required of the RTTAC-MS in order to have access to this Big Data for the entirety of the region. Similarly, if one agency already subscribes to a recommended Big Data source that covers the entire region but has limited licensing, granularity, and sharing restrictions, what incremental investment might be needed to modify the terms of the license to expand the scope of use and access of the Big Data.

IV. SCHEDULE

The duration of this contract will depend on the pricing options presented to the RTTAC-MS Committee.

VI. ESTIMATED BUDGET

The estimated budget for this contract is a lump sum amount of **$300,000**; comprised of the below agency funding contributions:

Agency	Funding Contribution ($)
FDOT	150,000
MPO	25,000
TPO	60,000
BC	25,000
TPA	40,000
TOTAL	300,000