



# Florida Statewide Model Task Force DATA COMMITTEE

Frank Baron, Miami-Dade MPO, Chair

Friday, May 20, 2005  
9:00 AM - 3:00 PM

Univ. of South Florida, College of Engineering  
4202 E Fowler Ave, Tampa, FL 33620

**CUTR** Board Room 143

## Agenda

8:30 AM Refreshments

9:00 AM Welcome and Introduction

F. Baron/H. Shen

9:15 AM Election of Vice-Chair

F. Baron

9:30 AM Committee Mission Statement **EXAMPLE DRAFT**

F. Baron

(Committee Goals and Objectives to be considered at a later time)

**MISSION:** To provide guidance to the Florida Model Task Force and the Florida transportation modeling community regarding modeling data needs to assist members in developing better models through data collection, analysis, and archiving, and modeling information management.

9:45 AM Discussion on Data Issues

All

Standard Data (Needs vs. Wants)

Standards for Data (Precision & Accuracy)

Standard Data Collection Approaches & Methods

- Standard Data Points
- Transferable Data
- Surveys & Data Collection Instruments

Standard Data Analyses

Standards for Archiving Data

- Surveys & Data Collection Instruments
- Raw Data Archival
- Analyzed Data Archives

**10:30 AM Break**

**10:50 AM Discussion on Data Issues (continued)**

All

**12:00 PM Lunch**

**1:15 PM Data News – Update on Central Office Initiatives**

Transit Office Update  
Systems Planning Office Update  
Statistics Office Update

T. Bartee/I. Ubaka  
Y. Wu/H. Shen  
G. Morgan

**1:45 PM Presentations: New Research Projects**

*Presentation: Standardize On-board Surveys and  
Archive On-Board Survey Data*

X. Chu

*Presentation: Development of State-of-the-Art Resources for  
Florida Travel Survey Data Collection and Analysis*

R. Pendyala

**2:15 PM Where Do We Go From Here? Wrap-up & Next Steps**

Frank Baron

**3:00 PM Adjourn**

**Attendees**

No.	Name	Organization	Phone	Email
1	Myung-Hak Sung	Gannett Fleming	813-882-4366	<a href="mailto:mhsung@gfnet.com">mhsung@gfnet.com</a>
2	Yongqiang Wu	FDOT/CO	850-414-4931	<a href="mailto:Yongqiang.wu@dot.state.fl.us">Yongqiang.wu@dot.state.fl.us</a>
3	Kwame Arhin	FHWA Florida	850-942-9650	<a href="mailto:Kwame.arhin@fhwa.dot.gov">Kwame.arhin@fhwa.dot.gov</a>
4	Gordon Morgan	FDOT/TranStat	850-414-4730	<a href="mailto:Gordon.morgan@dot.state.fl.us">Gordon.morgan@dot.state.fl.us</a>
5	Elaine Chang	USF Civil Engr	813-974-1738	<a href="mailto:echang@eng.usf.edu">echang@eng.usf.edu</a>
6	Fang Zhao	FIU Civil Engr	305-348-3821	<a href="mailto:zhaof@fiu.edu">zhaof@fiu.edu</a>
7	Mike Neidhart	Volusia County MPO	386-322-5160	<a href="mailto:mneidhart@co.volusia.fl.us">mneidhart@co.volusia.fl.us</a>
8	Huiwei Shen	FDOT/CO	850-414-4911	<a href="mailto:Huiwei.shen@dot.state.fl.us">Huiwei.shen@dot.state.fl.us</a>
9	Mike Maholtz	Sarasota Manatee MPO	941-359-5772	<a href="mailto:mmaholtz@sarasota-manateempo.org">mmaholtz@sarasota-manateempo.org</a>
10	Ike Ubaka	FDOT/PTO	850-414-4532	<a href="mailto:Ike.ubaka@dot.state.fl.us">Ike.ubaka@dot.state.fl.us</a>
11	Danny Lamb	FDOT/D7	813-975-6437	<a href="mailto:Daniel.lamb@dot.state.fl.us">Daniel.lamb@dot.state.fl.us</a>
12	Frank Baron	Miami-Dade MPO	305-375-1522	<a href="mailto:fbaron@miami-dade.gov">fbaron@miami-dade.gov</a>
13	Suraya Teeple	FDOT D2	904-360-5683	<a href="mailto:Suraya.teeple@dot.state.fl.us">Suraya.teeple@dot.state.fl.us</a>
14	Paul Larsen	Palm Beach MPO	561-684-4170	<a href="mailto:plarsen@pbcgov.com">plarsen@pbcgov.com</a>
15	Steve Polzin	CUTR-USF	813-974-9849	<a href="mailto:polzin@cutr.usf.edu">polzin@cutr.usf.edu</a>
16	Xuehao Chu	CUTR-USF	813-974-9831	<a href="mailto:xchu@cutr.usf.edu">xchu@cutr.usf.edu</a>
17	Jeanette Berk	API	904-823-8982	<a href="mailto:advancedplanning@bellsouth.net">advancedplanning@bellsouth.net</a>
18	Dennis Hooker	Metroplan Orlando	407-481-5672	<a href="mailto:dennish@metroplanorlando.com">dennish@metroplanorlando.com</a>
19	Kevin Feldt	JTA	904-598-8712	<a href="mailto:kfeldt@jtafla.com">kfeldt@jtafla.com</a>
20	David Schmitt	AECOM Consult	614-901-6026	<a href="mailto:David.schmitt@dmjmharris.com">David.schmitt@dmjmharris.com</a>
21	Ram Pendyala	USF Civil Engr	813-974-1084	<a href="mailto:pendyala@eng.usf.edu">pendyala@eng.usf.edu</a>
22	Amlan Banerjee	USF Civil Engr	813-974-8726	<a href="mailto:Abanerj2@eng.usf.edu">Abanerj2@eng.usf.edu</a>

## MEETING NOTES AND DISCUSSION

1. The meeting commenced with a welcome from the committee chair, Frank Baron, followed by a round of introductions.
2. Mike Neidhart was unanimously voted to serve as the Vice Chair of the committee. He will assist the Chair in all functions of the committee and serve/preside/coordinate in the event that the Chair is unable to do so.
3. A wealth of information on data needs, collection, and analysis to aid in model development, model implementation, and model enhancement was disseminated recently at the TRB Transportation Planning Applications Conference in Portland, Oregon (April 24-28, 2005).
4. Data is needed to develop, implement, calibrate, and validate accurate travel demand models that will, in turn, assist in the planning, formulation, selection, and implementation of sound transportation policies and projects. For example, planning and implementing good transit projects requires good transit models; good transit models, in turn, require good data. Thus, better data is needed (than what we have currently) to develop more accurate and robust models.
5. Attention needs to be paid to any data implications associated with the transition to Cube/Voyager. With the integration of Cube/Voyager with ArcGIS, there is a tremendous potential to better manage, display, use, and analyze data of all types (e.g., zonal data, network data, stop and route data, survey data, etc.).
6. The MPOAC should get the message that data collection is very important. The MPO directors should be informed that it is very important to collect good data and that data collection should be a very high priority item. Otherwise, the MPO will not get funds for new projects. The environment for securing funding for new projects is very competitive. If an area wants to get projects approved and funded by the federal agencies, the models have to be supported and backed up by timely, accurate, and comprehensive data.
7. A central repository of survey data is needed; the Central Office has initiated a new research project with USF College of Engineering to develop such a central repository on the world wide web.
8. How would such a repository be used? Several areas might not be interested in collecting their own survey data. Instead, they would use the data available in the repository and try to apply relevant data to their local context. We need to study and understand transferability of data, both in the temporal (over time) context and spatial (across geographical areas) context. How stable are different travel characteristics over time? How similar are different areas with respect to travel characteristics? Simply assuming stability over time and space may not be appropriate. We need to understand what causes changes in travel characteristics over time and across space.

9. We can not ignore the need for better socio-economic and land use data. Great strides have been made through the use of parcel level data, InfoUSA data, GDT networks, and so on to develop the FSUTMS zonal and network data. The community should continue working on collecting and assembling better data in this arena as well.
10. However, the scope of the data committee is likely to be more focused on survey data research as opposed to FSUTMS zonal and network data. FSUTMS zonal and network data is often covered under the jurisdiction of other committees such as the Trip generation, Land use, or Network committees. As such, it may not be necessary for this committee to focus on those databases as it would be a duplication of effort. As no other committee addresses survey research and survey data collection, that will be the primary focus of the committee. However, this does NOT mean that the committee will ignore zonal and network data collection and accuracy. The committee will also address these issues, as they come up.
11. There are many questions that come up with respect to data collection.
  - What data needs to be collected?
  - What is the most important data that needs to be collected?
  - At what point in time is data no longer valid? That is, when is data too old to be used anymore?
  - How often should be data be collected/updated?

As surveys and data collection tend to be expensive and resource intensive efforts, data collection needs to be prioritized. Identify the most important data and spend resources to collect it. Perhaps, cost sharing models can be developed where multiple MPOs or agencies get together to collect data that can be applied across their jurisdictions.
12. Most agencies and professionals know about data collection methods in the 1960s where classic large sample household travel surveys were conducted. Nowadays, there are new and advanced technological methods for collecting travel data. When do we use/adopt CATI (computer assisted telephone interview) and GPS-based surveys? After prioritizing data that needs to be collected, we need to identify the most accurate and sufficiently detailed method for collecting the data and any cost- or resource-sharing mechanism that can be implemented to help collect the data. We need to do this by MPO size and type; that would be a great contribution to the modeling community in Florida and a meaningful task for the committee.
13. Survey data is often used to estimate new model coefficients, parameters, and rates. Essentially, data is being used to estimate model parameters that can be used to forecast trips in the future. However, many areas use borrowed coefficients and parameters. What data is needed so that an area can customize coefficients and parameters to their specific context? Orlando is currently undertaking a major study to look at modal split parameters and model coefficients. Orlando is interested in determining how modal split models vary by income level. Essentially, Orlando is trying to determine how best to customize the model parameters and coefficients to reflect the characteristics of Orlando travelers.
14. The committee should be concerned with the following questions:
  - What data do you need?
  - What data can you borrow?

- What data can you share?
  - How can we use the data to validate models to individual areas? Essentially, we want forecasting models that can serve as robust tools for planning better transportation projects.
15. OMB is coming out with stringent requirements regarding data validity, accuracy, and precision. We need to be aware of the requirements of OMB because FTA is likely to hold any data collection effort to OMB standards.
  16. Although there is a lot of talk about transit modeling and the need for data to aid in enhancing transit models for obtaining New Starts Projects, we must take a holistic approach. We need to consider all modes of transportation, mobility and accessibility for all segments of population, all times of the day and days of the week, and so on. Such a holistic approach will help improve the model stream as a whole, including transit modeling. We need to have information that is transparent, reasonable, and defensible.
  17. We need to think about collecting data that has not traditionally been collected in surveys. Transit models are increasingly using measures such as walkability indices and transit accessibility indices to forecast ridership. We need data to develop and estimate such measures to be included in our models.
  18. We need to come up with a larger comprehensive data structure for all steps of the model. This data structure will then help guide data collection efforts in the state. The data structure should address both model building and model application processes. In general, data collection should be driven by modeling needs.
  19. FTA is very interested in onboard transit survey data, transit speed and travel time data, and highway speed and travel time data. Highway and transit speed and travel time data is very important. This data is often not collected and the data that is collected is not very accurate. We need to undertake research projects in this area of data collection.
  20. The whole idea of collecting good travel data must be conveyed effectively to MPO Boards who need to authorize funding for collecting such data. The information needs to be communicated in a manner that is understandable to a non-technical audience. MPO Board members do not understand the meaning of lifestyle variables and lifestyle trip generation models. In the past, the federal agencies have not been very demanding with respect to their scrutiny of models and data, possibly because the locals did not seek billions of dollars in funding from the feds. However, now the federal agencies are expecting local agencies to do their homework, invest in data collection and new modeling tools because literally tens of billions of dollars in investment is riding on these data and models. Also, everybody is paying attention to the fact that rapidly changing demographics – out migration, immigration of internationals, and so on warrant and demand that current data be used for model development, estimation, and application. Outdated data can not be used. If sound decisions are to be made, then one needs good data and timely data. Even the Census Bureau is moving in this direction with the American Community Survey (ACS). The decennial long form is eliminated and is now replaced by the annual ACS.

21. Many areas are seeking transit funds from FTA (new starts projects) with an eye towards shaping urban form. MPOs should be able to include data collection in their UPWPs. Therefore a plan of action is needed for MPOs to be able to include data collection in their next round of UPWPs. Transit models in Florida are very antiquated relative to transit models from other areas that are getting approved new starts projects from FTA. Need to prioritize and provide guidance regarding data needs and collection very soon so that MPOs can get the efforts budgeted in the next round of UPWPs.
22. FHWA also places great emphasis on the need for good data and on collecting and maintaining good quality up-to-date data. FHWA Tallahassee is willing to take any message that the committee deems appropriate up to the federal level in Washington, D.C.
23. However, it is more important to convey to the MPOAC that feds are willing to give money if locals will do their homework and do things right. The message should be clear: if you want projects, you need better models. If you want better models, you need better data. These data include those collected in general travel surveys, on-board transit surveys, roadside O-D surveys, and....
24. The committee should develop a paper that clearly articulates the need for data, the type of data that should be collected, prioritize the data items, and an action plan for collecting the data on a statewide basis. This paper can be provided to the MPOAC. In addition, the paper can articulate the stipulations, recommendations, guidelines, and conditions provided by the FTA. This would convince MPO Boards of the necessity to invest in data collection.
25. But we should not forget that there are many other constituencies interested in good data. The analysis of toll facilities is an example. Florida can do an add-on for the next NHTS/NPTS. The need to address rapidly growing congestion, smaller zone sizes, bike networks, walk networks, and to make decisions not related to traditional highway issues have warranted the collection of new and different types of data. Hence the need to consider data collection in a more holistic framework. However, we should remember that the critical issue came from the transit side of the business and the need to enhance the mode choice model component.
26. We need to think about the types of decisions that need to be made and then determine what data is needed to aid in that decision making process. If we talk exclusively about transit modeling, that may be a problem because the agenda becomes issue driven. Somebody might say that transit is not important in a certain area. Thus, it is important to take a holistic view where one is concerned with a comprehensive treatment of the transportation system. Issues such as peak period spreading, toll modeling, and so on are all important. We need to write the issues down and then identify the data needs/solutions for each issue.
27. There may be opportunities to coordinate MPO and statewide efforts. The state may invest in the development of a comprehensive survey database; this effort may involve the MPO's and state getting together on a periodic basis (say, every 5 years) and sharing costs to collect a statewide travel database that can be used by all MPOs in the state. Also, the state should consider being an add-on for the next NHTS/NPTS at the national level.

28. We need to put together a presentation for the July MPOAC meeting. Danny Lamb will work with Huiwei Shen to put together a nice presentation on the importance of travel data for transportation planning and decision making. The presentation could include discussion about standards for data and the types of survey data that need to be collected. This would include general travel survey data, onboard transit survey data, and traffic counts/vehicle classification counts by time of day. Also, tie in hot topics such as growth management and concurrency and freight and goods movement. There is a need for understanding freight and goods movement and to provide for the efficient delivery of goods and services.
29. We have to realize that we are competing for a limited pot of discretionary money. When competing for funding against everybody else, it is absolutely imperative that models be backed up and supported by good quality data that is timely, relevant, and defensible.
30. We must address the whole issue of speed data. Highway and transit speed data is very important for trip distribution, mode choice, path building, and assignment.
31. Another issue that needs to be addressed is that of zone structure and zone size and the implications for measuring and determining access to transit. How well does the zone structure and size configuration lend itself to accurate transit accessibility determination? Should we be specifying minimum and/or maximum zonal sizes? The committee should undertake the specification of best practices for models and data structures including networks, socio-economic data, and so on.
32. In the transit arena, we should collect appropriate transit route and network data to get rid of TCARDS. This can be done by having good transit speed data. Other MTF committees such as the network committee, transit committee, etc. need to tell the data committee about issues and problems so that the data committee can come up with solutions and standards. Perhaps, all committee chairs can brainstorm with their respective committees and then bring back ideas to data committee.
33. In this regard, a new role may emerge for the committee. Take each issue as it is brought forward and have a task team to resolve the issue, for example, TAZ size.
34. In Florida, travel surveys are happening as we speak and have happened in the recent past. More surveys are planned in the near future. The problem is that the survey data are not defined and collected in a uniform and consistent manner. We need consistency in data collection. Sometimes, simple measures are conceptually defined differently. For example, the definition of a trip may be different from one survey to another. We can not really transfer data across areas or compare survey results across areas if consistent definitions and methods are not used. This committee will therefore focus on developing uniform and consistent standards for survey instrument design and data collection methods. The committee may need to be prescriptive in addition to recommending broad guidelines. The committee needs to specify data items to be collected and how so in order to ensure that comparability is not compromised. To get a handle on the transit side of business, we need to oversample transit users. Onboard transit surveys allow us to do that.

35. How do we deal with non-response bias? The proportion of transit users among non-respondents tends to be greater than among survey respondents. We need to also be concerned with other special populations such as visitors, seasonal residents, elderly, children, and immigrants. How do we locate visitors and seasonal residents? Many just live with family and friends. We might be missing a whole layer of trips. If any agency is doing a survey, they may wish to consult with Jim Ryan to get advice and help on survey design and content.
36. More disaggregate modeling requires more detailed data – this increases burden on the respondents and makes it difficult to gather information. We need to try and be creative with getting data from other sources. Some data is available with Department of State regarding visitors. Through collaborative efforts with the Department of Education, DMV, Department of Health, Environmental Protection, etc., we may be able to get a lot of data that we need for modeling.
37. The Public Transit Office in the FDOT Central Office is doing a detailed transit speed study to investigate the relationship of transit speeds to the transit schedule, highway speeds, dwell time, stop frequency and spacing, and so on. There has been some earlier research on transit to highway speed relationships. National aggregate data suggests that transit is half as fast as highway. There is also the contribution of access time + egress time + transit circuitry to total transit travel time. The FSUTMS models we have now tend to underestimate bus speeds; buses are actually traveling faster than represented in models. A recent Tampa study seemed to show very comparable travel times between transit and highway – surprisingly close with transit speed = 80% of highway speeds. There are some questions, however, with regard to the nature of the routes sampled and the sample size used for the study.
38. The Systems Planning Office in the Central Office has several ongoing initiatives to further enhance FSUTMS. A true ArcGIS integration within Cube/Voyager is on the horizon. It is anticipated that the next release of Cube/Voyager will include this ArcGIS integration. This integration is likely to have implications for data collection, manipulation, storage, and analysis. ArcGIS functions allow one to build relationships and interactions between land use and transportation data. We need to have a fully operational market based land use model that can generate socio-economic data accurate enough to feed into transportation model. For this, we need to take advantage of GIS functionality and databases such as tax assessors data, parcel-level data, zoning data, and other such data related to land use development and regulation. On the transportation side, there are initiatives to include a lot of new information and detail in the network.
39. The Systems Planning Office is proceeding with the time of day modeling initiative. This time of day modeling initiative will go beyond the previous study where time of day factors were derived and identified. Thus, the idea is to integrate true time of day choice modeling into FSUTMS. The RFP will be out soon. There is some interesting work on data requirements for time of day modeling, e.g., District 4 and a few others are doing innovative work in this area. In addition, the Systems Planning Office is proceeding with developing and implementing junction based modeling capabilities. Ken Kaltenbach came up with a



speed and capacity calculator that appears robust in this context. Instead of traditional look-up tables, his method examines road configuration and signal controls to come up with speed and capacity. He recently demonstrated a program that generates an intersection based on one-digit FT and AT and calculates speed and capacity; the tool looks very encouraging. The method can be introduced to model task force if it works well. We need turning movement counts and detailed signal control data/indications to do junction based modeling. Note that junction based modeling also allows getting better highway speeds.

40. The Systems Planning Office has completed Cube workshops for the fiscal year. The comprehensive workshop is based on interim standards and enhancements to FSUTMS. As some areas are still in the transition to Cube/Voyager, it was necessary to keep some old Tranplan material in the workshops. However, in the next fiscal year, courses will be revised to include new developments and delete old material. The Cube Transition workshop is being discontinued as all transitions are now complete. New workshops related to model calibration, model validation, executive summary, ArcGIS, and understanding Voyager will be introduced. There will be another workshop in July in Orlando or Ft. Lauderdale. There might be one or more one-day workshops on specific topics such as junction based modeling, public transit modeling module, etc. in Cube/Voyager so that the modeling community can become familiar with the Voyager tools.
41. The Systems Planning Office is also working with Citilabs to set up a highway-only flowchart that would be standard in FSUTMS. There is also ongoing work on the development and identification of reasonableness checks – green, yellow, and red indicators that would serve as warnings.
42. One needs to be careful about specifying and developing models that are data intensive and data hungry. First of all, there are costs associated with data collection. Second, more data may not lead to better models. There are essentially two types of errors – model specification errors and data measurement errors. If you exclude variables or misspecify models, then travel behavior is not represented accurately. But if you include many variables, there is measurement error associated with every variable included in the model. So, the total measurement error over all variables will be very large. The overall accuracy of the model will not necessarily increase with the addition of more explanatory variables and dimensions. Essentially, there will be diminishing marginal returns. One should also be careful as to whether the model specification is merely correlative or truly causal.
43. One should resist temptation to develop very complicated models – avoid over-specification of models. The modeling community should use resources wisely. One does not always need a new survey to collect raw data; it may be possible to borrow or transfer data as appropriate. Statistical tests can be applied to determine if the data is applicable to a local context. We need to provide guidance to the community on how to apply statistical tests and measures to determine if survey data can be transferred to their specific context.
44. The Transportation Statistics Office of the Central Office has only highway data. This includes detailed roadway data, roadway geometrics, roadway jurisdiction, and traffic count data from 300 count stations. There is a nightly download of traffic count data and the

Statistics office maintains much of this traffic data. Many traffic counts are 1-2 day counts. The Statistics office applies seasonal and other factors to get AADTs, number of trucks, and so on. The GIS section in the Statistics office develops tools to help display and analyze data. They provide GIS coverages to everybody. The office provide data, reports, GIS information, and so on. Much of the information is available on the internet [www.dot.state.fl.us/planning/statistics](http://www.dot.state.fl.us/planning/statistics) . The Statistics office also supports other offices by providing data and information for supporting modeling, project selection, and the SIS. The Statistics office also coordinates storing and display of data. There are numerous performance measures for SIS; the data provided by the Statistics office contributes to helping other modes develop their own performance measures. The Statistics office is also a user of data. The office uses HPMS data and MPO forecasts to provide roadway traffic forecasts to FHWA.

45. The raw count data from the traffic count stations is generally hourly by direction; but the data is compiled and aggregated to the link level in the RCI database. As the transportation planning community is being asked questions that requires more detailed traffic data, is it possible to publish data as disaggregate as 15 minute counts by direction by lane? Fang Zhao completed a research study that looks at seasonal variations in counts from permanent sites and how that can be used to adjust counts at other sites to account for seasonal variation. Richard Reel in the Statistics Office would be the contact to find out if more detailed data can be made available to the modeling community.
46. There are other interesting initiatives and ideas that need to be considered. Getting data may become possible using new technology. Transponders and GPS technology, cell phones, and other devices can be used for travel data collection. The FIU EE department project involves investigating the use of cell phones for travel data collection. CUTR has a similar project underway. The Central Office Systems Planning office is planning a small scale O-D study for SIS implementation.
47. Drs. Xuehao Chu and Ram Pendyala made presentations on their new data research projects. The powerpoint presentations are attached.

### **Action Items**

The MTF Data Committee identified several action items as follows:

- Make a presentation to the MPOAC about the need for detailed and accurate data to support transportation planning and decision-making and the importance of allocating resources for data collection. This presentation will be made by Danny Lamb in coordination with the Central Office (Huiwei Shen) and the MTF Data Committee Chair (Frank Baron).
- The Committee, in coordination with the FDOT Central Office, should sponsor research that results in a paper that clearly articulates the need for data, the type(s) of data that should be collected, list that data in order of priority, and recommend an action plan for collecting that data around the state –for both statewide data, and local (urban area and regional) data.

- The Committee, in coordination with the Central Office, should produce a research paper that recommends (perhaps even prescribes and proscribes certain principles and practices) application of statistical tests and publication of results for all survey activities associated with modeling.
- The Committee, in coordination with the Central Office and other MTF Committees, should develop an overall, larger and comprehensive data structure for the Cube-Voyager model with a listing of all types of data formats: e.g., the various formats for the various GEN models used around the state; it would be an excellent idea to develop – and publish – a compendium of data formats – both input (I) and output (O) – for all the standard models in the state (not ‘special’ cases, like that developed for the Miami North Corridor study; rather those data formats for the general ‘Miami model’ should be listed).