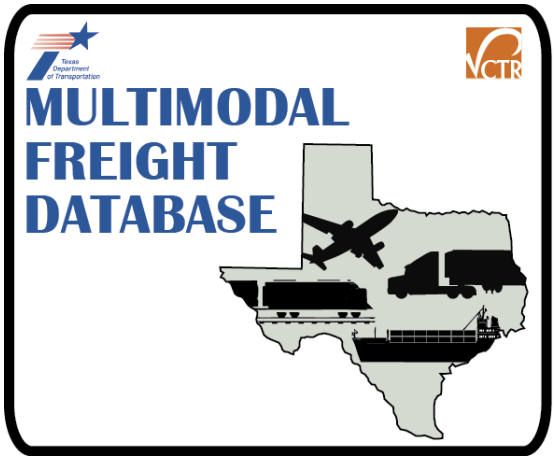
Relational Multimodal Freight Database Webinar



CENTER FOR TRANSPORTATION RESEARCH

The University of Texas at Austin

February 2012

Research Project 5-6297-01-P1

Contents

Introduction 3
Welcome and Introduction 4
Presentation 5
Questions and Answers 9
Final Remarks 11

Introduction

The relational Multimodal Freight Database (MFD) was developed as part of Texas Department of Transportation (TxDOT) Research Project 0-6297 entitled *Freight Planning Factors Impacting Texas Commodity Flows*, conducted by the Center for Transportation Research (CTR). The objective was to capture relevant publicly available freight data required for updating TxDOT freight models and studies. Implementation Project 5-6297-01, entitled *Multimodal Freight Database*, extends the work that has been conducted as part of TxDOT Research Project 0-6297. Specifically, the objectives of Implementation Project 5-6297-01 are to disseminate information about the MFD by hosting six workshops (Task 1), host a webinar to demonstrate the MFD to selected FHWA (Federal Highway Administration) freight leaders (Task 2), and update the MFD given the feedback received during Tasks 1 and 2.

This document summarizes the outcome of Task 2: Conduct Multimodal Freight Database Webinar. The objective of Task 2 was to demonstrate the MFD to FHWA freight leaders and TxDOT Division staff to gain support for the development of a national database and to discuss the integration of proprietary data into the existing MFD. The webinar was hosted by Mr. Kirk Fauver (Statewide Planning Engineer, FHWA) and Ms. Jolanda Prozzi (Assistant Director, CTR). The webinar was conducted using Adobe Connect on December 8, 2011 from 1:00 to 2:30 p.m. (CST).

Welcome and Introduction

Mr. Kirk Fauver started the webinar at 1:05 p.m. (CST) by introducing himself and sharing relevant information with the participants regarding the webinar structure. Specifically, Mr. Fauver announced that the CTR presentation would be approximately one hour followed by a 30-minute Q&A session. Mr. Fauver also asked participants to mute their telephones and explained the procedure for un-muting their telephones during the Q&A session. Finally, Mr. Fauver reminded all participants that the webinar would be recorded.

The following individuals participated in the webinar:

- Jolanda Prozzi (CTR)
- Migdalia Carrion (CTR)
- Dan Seedah (CTR)
- Dr. C. Michael Walton (UT)
- Jennifer Moczygemba (TxDOT)
- Orlando Jamandre (TxDOT)
- Kirk Fauver (FHWA)
- Jim Thorne (FHWA)
- Kate Quinn (FHWA)
- Carol Keenan (FHWA)
- Chip Millard (FHWA)
- Eric Pihl (FHWA)
- Lisa Randall (FHWA)
- Michael Sprung (FHWA)
- Ralph Volpe (FHWA)
- Spencer Stevens (FHWA)

In addition, the following individuals were invited, but unable to participate in the webinar:

- Jeffrey Frkonja
- Michael Leary
- Jose Campos
- Tony Furst
- Rolf Schmitt
- Carl Millard
- Crystal Jones
- Jack Foster
- Ed Strocko
- Genevieve Bales
- Sandra Otto

Mr. Fauver subsequently introduced Ms. Prozzi.

Presentation

Ms. Jolanda Prozzi started the presentation by outlining the origins of the relational MFD. Specifically, the database was developed as part of a larger TxDOT research project (TxDOT Project 0-6297) approximately 2 years ago. Ms. Jennifer Moczygemba was the Project Director and Mr. Orlando Jamandre was heavily involved in the MFD's development.

Ms. Prozzi proceeded to outline the topics covered in the day's presentation, including a discussion of the study objectives of TxDOT Project 0-6297 (of which the development of the MFD was a component), a brief presentation on the development of the Database, a more in-detail overview of the Database, and a few slides summarizing the feedback and suggested improvements received from potential Texas users of the MFD.

Ms. Prozzi explained that the main objectives of TxDOT Project 0-6297 were to (1) obtain freight data and (2) synthesize the information in such a manner that the document could be a basis for the development of a freight plan for Texas. More specifically, the study sought to improve the understanding of the nature of the commodities moved in Texas, the business and transportation system factors that affect shipping decisions, and the factors that impact freight competitiveness in Texas. These objectives were achieved by hosting a series of workshops with the shipper and freight planning communities in Texas. Secondary objectives of the study included the identification and documentation of the current state of the multimodal freight system in Texas, the development of recommendations for freight investments and policies to be considered by TxDOT, and finally, to explore the interest in the development of a Texas Freight Advisory Committee. Finally, Ms. Prozzi commented that, as a component of TxDOT Project 0-6297, the commodity data gathered were packaged into the relational MFD. The Database was populated with relevant data from publicly available sources.

Ms. Prozzi mentioned that the Database was developed with the intention of being a single source for Texas freight data that can be made available to the freight planning community in Texas. As a single data source, the Database needed to be capable of mapping across multiple commodity codes used in the different public databases. Other features of the Database include a user-friendly interface and the need for minimal computer processing.

Ms. Prozzi further discussed the features and limitations of the MFD. The MFD includes freight variables obtained from publicly available databases, which can be used for planning purposes, and presents the information contained in these databases in a similar format. The Database also provides the user with an overview of the information contained in each public database. On the other hand, the MFD does not contain any private or proprietary data. Ms. Prozzi elaborated that this limitation was by design to avoid concerns about dissemination. Ms. Prozzi also added that the Database to be submitted in January 2012 will contain the most recent available data, but no mechanism current exists to incorporate future data releases. Finally, Ms. Prozzi mentioned that the MFD does not provide an analysis of the data, combine any data sources, or assesses the reliability and/or accuracy of the data. Ms. Prozzi did mention, however, that an overview of each data source—including the sponsoring agency, data location, methods used, and stated limitations—was included in the MFD, but an evaluation of this information was not conducted by the research team.

Ms. Prozzi then explained that TxDOT Project 0-6297 officially ended in May 2010, but TxDOT decided to fund an implementation project (TxDOT Project 5-6297) to demonstrate the MFD to key potential TxDOT and MPO users, as well as to selected FHWA freight leaders, for review and comment. Subsequently, the MFD will be finalized and disseminated to all TxDOT Districts and MPOs. Ms. Prozzi also mentioned that the contract end date for the implementation project is January 31, 2012.

At this time, Ms. Prozzi invited the participants to ask any questions or provide comments. A transcript of the questions and comments is provided in Section 1 of *Questions and Answers* (later in this report).

In the discussion on the Database Development and Coding, Ms. Prozzi focused on the selection of the data variables included in the MFD, the selection of the data sources, and the level of data aggregation.

For the variable selection, the research team consulted with the Project Advisors and ultimately included the following variables:

- Year,
- Mode of Transportation,
- Origin and Destination,
- Port (Entry and Exit),
- Commodity Type,
- Type of Movement,
- Number of Loads,
- Containers,
- Value, and
- Weight.

In terms of selecting the data source, the research team conducted an initial review of 31 potential and relevant data sources. The list was reduced to 22 data sources after excluding private and proprietary data, as well as data sources that have been terminated and for which data are no longer collected. The remaining 22 data sources were all publicly available (with the exception of the Carload Waybill and the North American Trucking Survey, which are available to federal and state agencies upon request). Of these 22 data sources, only 11 contained all or some of the identified freight variables of interest.

The following data sources were thus included in the current version of the MFD:

- Commodity Flow Survey (CFS)—U.S. Census Bureau and the Bureau of Transportation Statistics,
- Freight Analysis Framework (FAF)—FHWA,
- North American Transborder Freight Data (TBR)—Bureau of Transportation Statistics,
- Annual Coal Report (ACR)—U.S. Department of Energy,
- Bureau Crossing Data (BCD)—Bureau of Transportation Statistics,
- Fresh Fruits and Vegetable Shipments by Commodity, State, and Month (FFV)—U.S. Department of Agriculture,
- Maritime Administration Data (MAR)—Maritime Administration (USDOT),
- National Transportation Statistics (NTS)—Bureau of Transportation Statistics,
- US Trade Data (USA)—STAT-USA and Foreign Trade Division of the U.S. Census Bureau,
- Waterborne Commerce Statistics Center (WBN)—U.S. Army Corps of Engineers, and
- Carload Waybill Public Use File (WAY)—Surface Transportation Board (USDOT).

Ms. Prozzi then presented several tables to demonstrate the differences in the level of aggregation used in the 11 data sources. A coding system was developed to present the data uniformly and at the same levels of aggregation. Specifically, in terms of the Origin and Destination data, Ms. Prozzi pointed out that none of the data sources captured a substantial amount of disaggregated (i.e., county-level) data. In terms of mode of transportation, the data sources also vary substantially as to which modes are included. For example, the Waterborne Commerce Statistics

Center captures information only for a single mode (i.e., water), whereas the Freight Analysis Framework captures data for all modes of transportation. In terms of shipment data, most data sources include information for either weight, value, or both. In terms of commodity information, most data sources typically use either the commodity or industry classification system (although databases that use a unique classification system do exist). The classification systems include the Harmonized Commodity Description and Coding System (HS), the North American Industry Classification System (NAICS), the Standard Classification of Transported Goods (SCTG), and the Standard Transportation Commodity Code (STCC). The research team thus developed a data map to document how commodity data from different databases were coded in the MFD. This data map was subsequently used to convert commodity codes from HS, NAICS, SCTG, and STCC to the commodity codes used by the research team. Although initial discussions with TxDOT revealed the need for only the nine commodity codes included in TxDOT's Statewide Analysis Model (SAM), the research team ultimately used 41 commodity codes (i.e., the two-digit STCC codes plus two codes to denote unknown or not reported information). Ms. Prozzi illustrated the research team's coding efforts by showing excerpts of the original North American Transborder Freight Data and Carload Waybill Public Use File data alongside excerpts of the same data after they were re-coded by the research team.

In terms of the software platform for the MFD, the research team used Visual Basic to develop the queries and relied on Wizard Access to develop the forms and reports. The MFD also includes a user-friendly interface that allows the user to search an individual database or to specify a number of freight variables to be searched in multiple databases. To conclude this section, Ms. Prozzi explained that, in hindsight, using Wizard Access to develop the forms and reports was not the best option and that the research team is currently working on improving the software's user-interfaces.

At this time, Ms. Prozzi invited the participants to ask any questions or provide comments. A transcript of the questions and comments is provided in Section 2 of *Questions and Answers* (later in this report).

Ms. Migdalia Carrion subsequently continued with a review of the software product. Ms. Carrion started by explaining how to install and use the MFD. She explained that an automated run should appear if the MFD CD is placed in the computer's CD drive. She also explained how to install the MFD if the automated run does not appear and manual installation is required. Ms. Carrion then explained the software features of the MFD. A question mark icon appears in all the MFD screens. When the user clicks on the question mark, the MFD User Manual opens in Microsoft Word. The MFD User Manual provides step-by-step guidance on how to use the MFD. The Manual also contains two appendices. Appendix A contains an overview of each data source, including the sponsoring agency, data location, methods used to gather the data, and stated limitations. Appendix B contains the code developed by the research team and used in the MFD. The "About the Databases" button that appears on the second screen of the MFD links to an .html file containing the information provided in Appendix A. The latter makes it easier for the user to locate information regarding a particular data source. Finally, the "Summary" button allows the user to populate a query in a timely manner. The "Summary" report contains a number of hyperlinks that can be expanded to display the data by clicking on the hyperlinks. The report can be saved, copied, and pasted into either Excel or Word.

Mr. Tim Thorne then assisted Ms. Carrion in sharing the desktop view of her computer with the participants. Ms. Carrion proceeded to demonstrate the MFD by guiding participants through an example. At this stage, Ms. Carrion invited the participants to ask any questions or provide comments. No questions or comments were provided.

Ms. Prozzi continued the presentation by outlining some of the feedback and suggestions for improvement provided by potential MFD users in Texas. Ms. Prozzi explained that the MFD was demonstrated in a series of workshops to potential TxDOT and MPO users at the following TxDOT District offices:

- El Paso (September 28, 2011),
- Tyler (October 5, 2011),
- San Angelo (October 12, 2011),
- Lubbock (October 19, 2011),
- Houston (October 26, 2011), and
- Corpus Christi (November 9, 2011).

During each workshop, the data included in the MFD and the development of the MFD was discussed. A sample query was created to display the functionality of the MFD and potential users were then encouraged to run pre-designed queries. Once the potential users had time to review the MFD and its features, participants were asked to provide the research team with feedback and suggestions for improvements to the MFD. Approximately 10 to 20 potential users attended each workshop, with more than 20 potential users attending the Houston workshop.

Ms. Prozzi proceeded to demonstrate the feedback obtained from potential users. Overall, the feedback received was very positive. For example, a Houston participant pointed out that the MFD could replace some of the private data that cannot be readily updated. Several other participants also commented on the user-friendliness of the interfaces. A San Angelo participant mentioned that some of online databases cannot be accessed at TxDOT because of firewalls. Having the MFD thus makes the information easily accessible to planners. Other participants new to the field of freight planning commented that having all the freight data in one database helps them to gain better insight in the area. Also, the information about each of the databases included in the MFD helps potential users to understand the differences among data sources.

A number of potential users also highlighted several opportunities to improve the MFD. These improvements can be categorized as (a) data, (b) interface, and (c) reporting improvements.

In terms of data improvements, the MPO participants expressed a need for more disaggregated data—i.e., down to the city or region level. Most participants also expressed the need for more data to be included in the MFD. In terms of the interface, the participants recommended that the variables are defined when the mouse is moved over the variable name, as well as for relevant criteria to be interactively updated as selections are made. Most of the recommendations, however, related to the creation of the reports and how the data are displayed. For example, participants recommended that the data be displayed in a grid-like format to allow for sorting, filtering, and further analysis of the data. Ms. Prozzi remarked that most of the remaining effort will be on improving the data reporting structure.

Ms. Prozzi concluded by mentioning that the MFD was reviewed positively and many potential users seemed enthusiastic about the database. Ms. Prozzi also mentioned that the research team is sharing this information with the FHWA, because it is believed that other states may also benefit from this exercise.

Finally, Ms. Prozzi invited the participants to ask any questions or provide comments. A transcript of the questions and comments is provided in Section 3 of *Questions and Answers* (later in this report).

Questions and Answers

The following sections provide a transcript of the questions and comments provided by the Webinar participants (*in italics*) and the responses by the research team. Participant names are provided if the participants identified themselves.

Section 1

[Mr. Fauver] In conjunction with TxDOT, what future improvements to the software are in the works? Will these software improvements make the current software obsolete within the next few years? Has the CTR maintained a list of users so that they may be notified of any software improvements or upgrades?

[Ms. Jolanda Prozzi] Improvements are currently being made to the software in terms of updating the data sources and restructuring the query reporting structure. After these improvements, the current version of the software will not necessarily become obsolete, but the data sources would ultimately need to be updated and at this time there is no funding allocated to do that.

[Mr. Fauver] What types of studies and major transportation investments may this software be used for in terms of scope? Is it primarily geared toward statewide planning efforts (e.g., state rail plan or LRTP efforts, etc.), or can it also be used for major corridor studies and project-level NEPA studies? Has it been used for any MPO long-range transportation planning purposes?

[Ms. Prozzi] At this stage, the focus was on TxDOT planning studies, so the data in the database are too aggregate for corridor studies—mainly because that is the information available. MPOs have shown a lot of interest in the data product, but they have a need for more disaggregated data. Unfortunately, most of it needs to be collected.

[Dr. Walton] We are working with the North Central Texas Council of Governments on a freight study now to link emerging technologies with new data and current databases.

It was mentioned that the current implementation of the software does not assess the quality of the data, but does it include any methodology data that the creators of the data might have included and which could relate to the limitations of the data source?

[Ms. Prozzi] We included about three to four pages of information regarding each data source, including a website link to the data source, information on how the data is collected, the sample selection and size, and the quality control procedures implemented, but we did not evaluate the methodology and left it up to the user to assess the validity of the data.

[Mr. Onder] Do you have an example of how this database will be used?

[Ms. Prozzi] We have some feedback from potential users that mentioned that the Database could be useful for them. Some possible applications of this Database after distribution to TxDOT could be for freight planning studies and the state rail plan, among others. However, this particular Database has not been used for any particular application yet, because it is still in the development phase.

Section 2

[Ms. Quinn] Beyond the Federal level data, did you find state or local data that are useful?

[Ms. Prozzi] TxDOT has a large Commercial Vehicles Survey program that collects commercial vehicle data. At the time, the DOT was more interested in county-level data, and for this reason, the data from this survey could not be incorporated—mainly because it was more urban movements. However, after receiving the MPO feedback, this decision should be revisited.

In addition, the research team also reviewed data sources that were included in other research studies, but these sources were not incorporated mainly because the data were either outdated or the sample sizes were too small.

[Ms. Quinn] Is there any other department in the state, besides TxDOT, that is collecting data (e.g., the Commerce Department)?

[Ms. Prozzi] Not that I am aware of.

[Dr. Walton] The Commerce Department [information] we have here is economic development-related and does have some data and information but does not have a rigorous data set. Thus, we chose not to include it at this time.

[Ms. Quinn] We are always interested in other data sets that people have collected because we get asked this by other states all the time.

[Mr. Frkonja] What zone structure did you use to code the Origin-Destination data?

We combined the Federal state code—i.e., the first two digits—and the TxDOT county level codes—i.e., the last three digits—to form a unique code for each origin and destination.

Did you talk about how the original research was funded and what would you do if additional planning research funds were available for this model? My understanding is that SPR helped fund this original study?

[Ms. Prozzi] The original funding came from TxDOT's Research Technology and Implementation Office (RTI). The second component, which is the outreach to the TxDOT Districts and the MPOs and sharing of this product with potential users, also came from RTI. At this moment, there is no funding allocated to maintain this product. In other words, there is no funding to update the data annually or to improve the usability of the software. If more funds become available, we would have the opportunity to improve the software and address some of the suggested improvements that we will not be able to incorporate within this project. It will also give us the opportunity to look for ways to incorporate more freight data into this product.

[Dr. Walton]: Yes, I think it is the SPR program that the state administers through its R&D group and it is a rigorous process that they go through to identify potential areas and then they go through a competitive process. At this time, we do not know that there is any effort to support us. The implementation phase was an add-on after some of the products from the previous year were identified and selected to go into a more near-deployment phase, so those underwent an implementation phase.

Section 3

[Eric Phil] Is the software capable of writing out flow tables based on Origin-Destination geography?

[Mr. Seedah] At this time, we are limited in time and budget so we could not include this feature. However, it is definitely a possible feature that can be added to future versions of the product.

What is the status of this product? Are you close to wrapping it up?

[Ms. Prozzi] We are in the process wrapping it up. The main objective of the implementation phase was to demonstrate this product to six TxDOT districts and collect feedback and suggestions. Our final report for this implementation phase is due at the end of January 2012, so we are in the wrapping-up phase.

There are other states that will probably be interested in what you are doing here, is there someone in particular that they can contact?

[Ms. Prozzi] Yes, they can contact either Jolanda Prozzi or Dr. Walton and we will gladly share what we have developed as part of this effort.

[Jennifer Moczygemba] If I remember correctly, the software is structured to filter for everything that is related to Texas, but the coding structure exists for all the other states, so is it only a matter of having a server that can contain all the data from the other states?

[Ms. Prozzi] That is right. In reviewing all of the public data sources, we extracted all of the data that pertained to Texas, but all the codes are there and we acquired all of the data for the other states. It is only a matter of coding the data and packaging it in a different way.

After the last question, participants provided input regarding potential agencies that could be interested in the MFD. One participant suggested presenting the MFD to one of the AASHTO committees—possibly the intermodal or the planning committee. The participant added that it is always important to present projects such as this one to the leadership. Other possible interested parties could be the participants of the mid-year summer meetings of the freight and economics committees, including Mr. Ralph Schmidt, who works with freight data. Other possible interested individuals include Ms. Ann Purdue, liaison for most of the freight committees, and Ms. Joedy Cambridge, liaison for the port and waterway committees.

[Kirk Fauver] What are the costs involved to get other state's data into the product?

[Ms. Prozzi] I do not know. We need to analyze the costs involved. A lot of the groundwork has been done already but we have not considered this possibility.

Suggested to think about it, because other states might ask.

[Ms. Prozzi] The most important thing is to fund the maintenance of the database and continued implementation to keep it a very valid tool and to keep it updated.

Cost to update it?

[Ms. Prozzi] Not a large cost. The data structure is there. Resources are only required to upload the data and get it in the required structure. It is not a major cost to update it and disseminate it to the relevant parties.

If TxDOT would want to keep it going they need to know the costs.

Could it be done every other year?

[Ms. Prozzi] Yes, some of the data sources are only updated every 5 years, so not a lot of resources are needed every year to update the data sources.

Try to cost it out.

Good effort.

Final Remarks

After all questions were answered, Mr. Fauver thanked Ms. Prozzi, the CTR team, and all the attendees for participating in the Webinar. The webinar ended at 2:00 p.m.