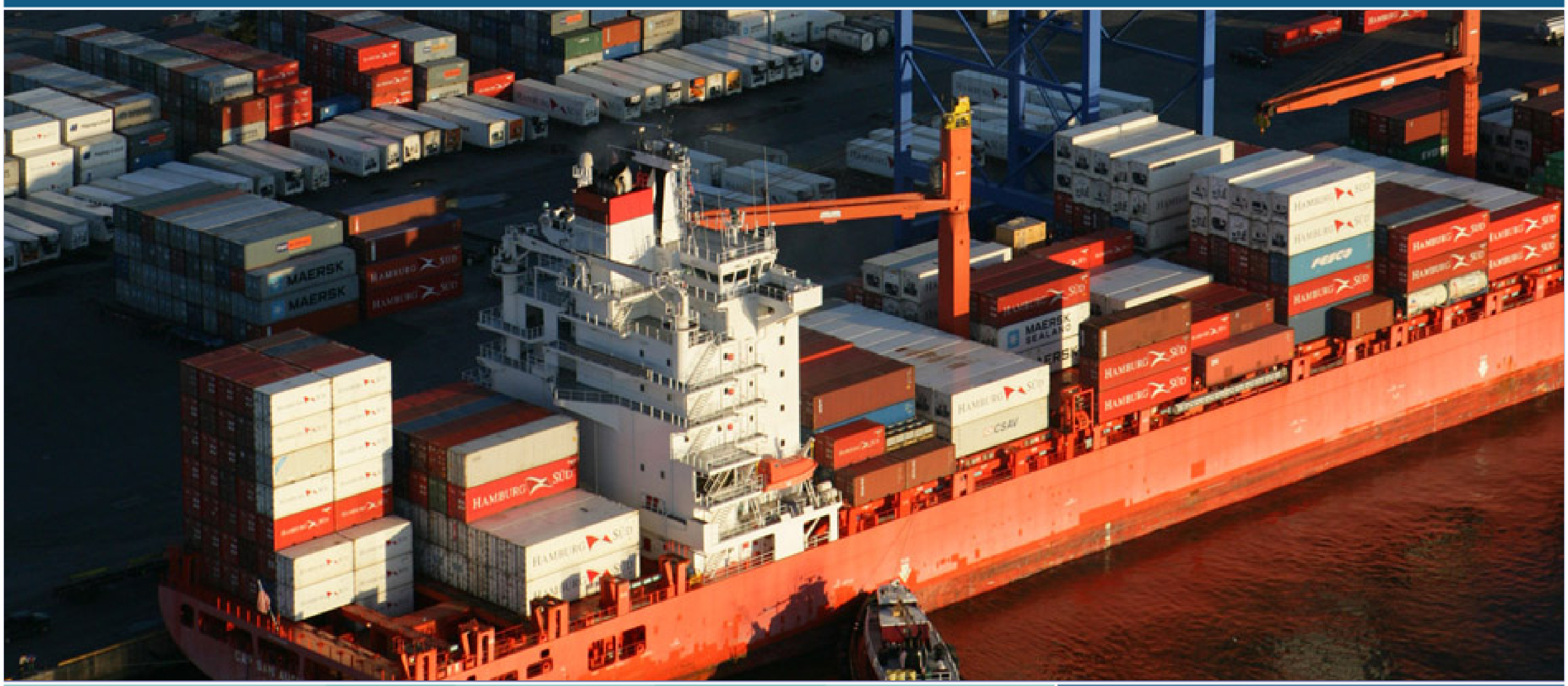
Making Freight Data More Accessible

SHRP2 C20: Local Freight Data Improvement



APRIL 2016

DELAWARE VALLEY
dvrpc
REGIONAL
PLANNING COMMISSION

The Delaware Valley Regional Planning Commission is dedicated to uniting the region's elected officials, planning professionals, and the public with a common vision of making a great region even greater. Shaping the way we live, work, and play, DVRPC builds consensus on improving transportation, promoting smart growth, protecting the environment, and enhancing the economy. We serve a diverse region of nine counties: Bucks, Chester, Delaware, Montgomery, and Philadelphia in Pennsylvania; and Burlington, Camden, Gloucester, and Mercer in New Jersey. DVRPC is the federally designated Metropolitan Planning Organization for the Greater Philadelphia Region — leading the way to a better future.

DELAWARE VALLEY dvrpc REGIONAL PLANNING COMMISSION

The symbol in our logo is adapted from the official DVRPC seal and is designed as a stylized image of the Delaware Valley. The outer ring symbolizes the region as a whole while the diagonal bar signifies the Delaware River. The two adjoining crescents represent the Commonwealth of Pennsylvania and the State of New Jersey.

DVRPC is funded by a variety of funding sources including federal grants from the U.S. Department of Transportation's Federal Highway Administration (FHWA) and Federal Transit Administration (FTA), the Pennsylvania and New Jersey departments of transportation, as well as by DVRPC's state and local member governments. The authors, however, are solely responsible for the findings and conclusions herein, which may not represent the official views or policies of the funding agencies.

The Delaware Valley Regional Planning Commission (DVRPC) fully complies with Title VI of the Civil Rights Act of 1964, the Civil Rights Restoration Act of 1987, Executive Order 12898 on Environmental Justice, and related nondiscrimination statutes and regulations in all programs and activities. DVRPC's website, www.dvrpc.org, may be translated into multiple languages. Publications and other public documents can be made available in alternative languages and formats, if requested. DVRPC public meetings are always held in ADA-accessible facilities and in transit-accessible locations when possible. Auxiliary services can be provided to individuals who submit a request at least seven days prior to a meeting. Requests made within seven days will be accommodated to the greatest extent possible. Any person who believes they have been aggrieved by an unlawful discriminatory practice by DVRPC under Title VI has a right to file a formal complaint. Any such complaint may be in writing and filed with DVRPC's Title VI Compliance Manager and/or the appropriate state or federal agency within 180 days of the alleged discriminatory occurrence. For more information on DVRPC's Title VI program, or to obtain a Title VI Complaint Form, please call (215) 592-1800 or email public_affairs@dvrpc.org.

Table of Contents

Figures

Tables

Appendices

Introduction

The Local Freight Data Improvement project is funded through the second Strategic Highway Research Program (SHRP2). SHRP2 offers a road map that will lead to improved freight data sets and freight modeling practices. This round of the SHRP2 Implementation Assistance Program helps state DOTs and MPOs deploy the SHRP2 solutions and develop innovative Freight Demand Modeling and Data Improvement (C20) products. The Delaware Valley Regional Planning Commission (DVRPC) Local Freight Data Improvement project was focused on enhancing the availability and visibility of freight data for regional planning. The core objectives of this project were to:

- Identify and adapt disparate sources of data;
- Refine current data sources for regional or sub-regional applications;
- Establish, pool, and standardize a portfolio of core freight data that supports planning, programming, and project prioritization; and
- Improve and update *PhillyFreightFinder*.

In order to achieve these objectives, the keys steps involved with the Local Freight Data project were to develop a data collection plan, develop a core freight database, update *PhillyFreightFinder*, and develop an open source version of *PhillyFreightFinder*. Throughout the course of the project, a group of public and private stakeholders were engaged as a project steering committee. These members were drawn from an existing Data Subcommittee of the DVRPC's freight advisory committee, the Goods Movement Task Force and included:

- Rick Crawford, Norfolk Southern Corporation
- David Harvey, Philadelphia Regional Port Authority*
- Nicholas Haynes, Philadelphia Regional Port Authority
- Paul Myhre, Maritime Exchange of the Delaware River and Bay
- Chris Swann, Select Greater Philadelphia*
- Michael Singer, UPS
- Brian Wall, Pennsylvania Department of Transportation
- James Mosca, Pennsylvania Department of Transportation
- Talvin Davis, New Jersey Department of Transportation*
- Thanh Le, New Jersey Department of Transportation
- Michael Cooper, Philadelphia Department of Commerce

*Individual left position and committee during course of project

Defining Data Needs

An increasing regard and emphasis on freight considerations within transportation planning is the result of a number of factors, such as increasing freight volumes, dramatic shifts in global and domestic supply chains, and evolving consumer preferences and purchasing habits. The growing importance of freight transportation is reflected in the significant freight planning provisions found within the current federal transportation legislation, *Fixing America's Surface Transportation Act* or *"FAST" Act,* and the number of SHRP2 initiatives that acknowledge and advance freight.

Historically, the acquisition and use of freight data by MPOs and state DOTs for planning purposes has been fraught with inherent challenges and obstacles. Among the complications has been competition within and across different modes, difficulty in obtaining comprehensive and current data, and a lack of continuity in data collection and analysis efforts.

Despite these challenges, seeking to establish a robust and well-maintained freight data program is highly desirable. Of primary importance, freight represents *the economy in motion* and has a direct correlation to employment. In a region like the Delaware Valley with close to 3 million jobs, an extensive and diverse freight network, and stiff competition for transportation funding among different types of projects, it is vital to tap into and cultivate available data sources to identify transportation investments that serve shippers and carriers under any future scenario and that leverage and encourage public-private partnerships.

Additionally, the movement of freight is an extremely significant part of the transportation landscape. Trucks, freight trains, ships, and airplanes are large, impactful forms of transport that are often co-mingled with passenger forms of transportation. Therefore, understanding the routing and behavior of all freight "vehicles" can ultimately help better integrate freight transportation operations and facilities with community goals (e.g., improved air quality).

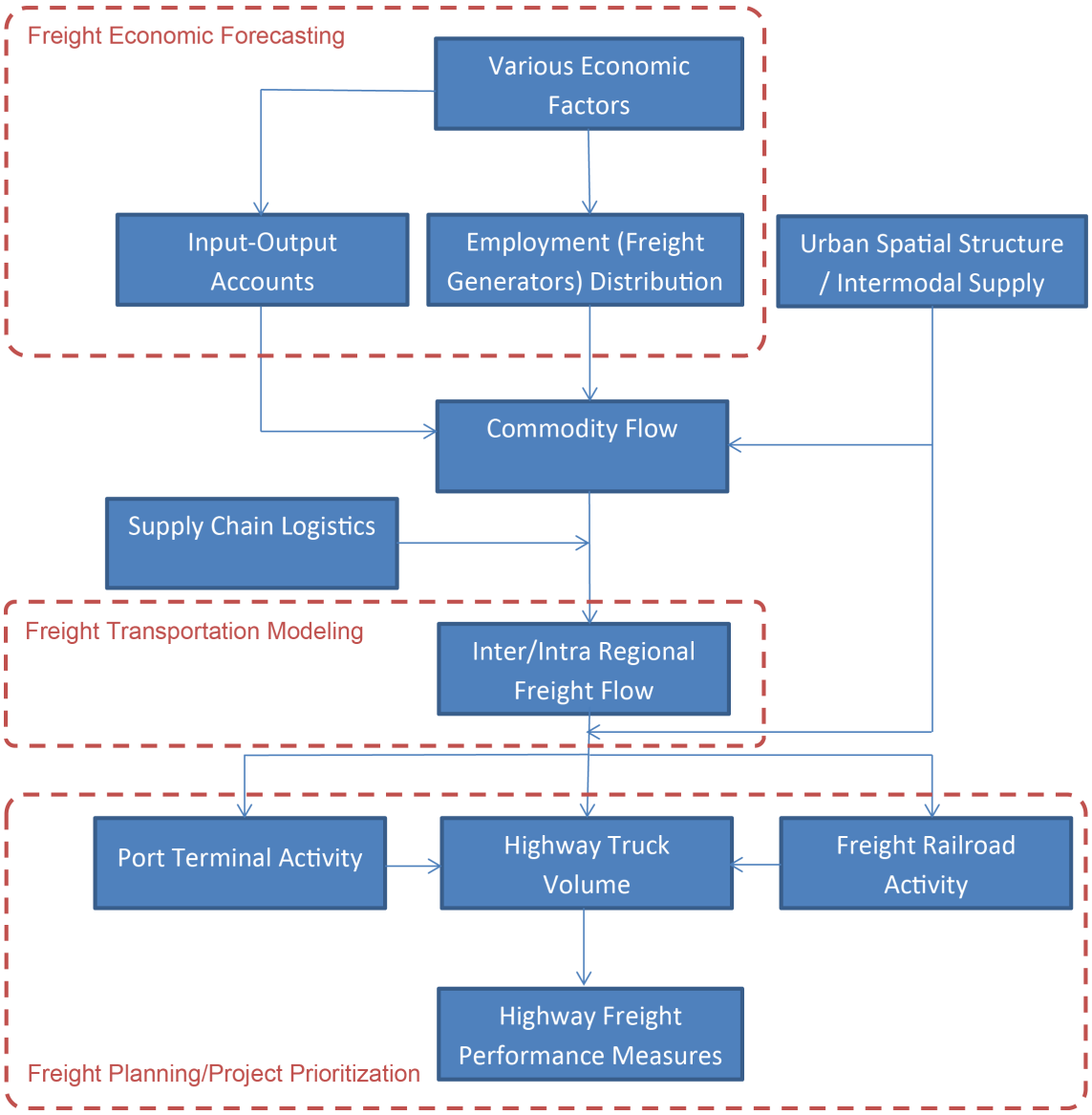
For the current SHRP2 freight data initiative, DVRPC has identified the following practical uses and outcomes:

- Enhancement of visualization tools and the portrayal of freight flows and trends for a variety of audiences (especially through the *PhillyFreightFinder* tool)
- Setting of a foundation for freight performance measures and the identification of freight-specific bottlenecks and problem locations for select commodities and industries
- Establishment of a process to prioritize candidate projects seeking federal funds (e.g., for the Long-Range Plan and Transportation Improvement Program) and grasp their relative benefit to improving freight systems and operations within the region
- Creation of an analytical framework to connect regional economic development with freight transportation system improvement needs by profiling freight generators, freight-related employment, and supported commodity movements within the region
- Data preparation for developing and calibrating an advanced regional freight demand model

In order to fully support these identified needs, this project explored the opportunity to collect data from a spectrum of economic and transportation sources. Overall, a regional freight data and analysis framework, as shown in the following flowchart, is proposed for this study and as a long-term guidance for freight data collection and analysis initiatives at DVRPC. Under this framework, the DVRPC project team identified

potential data sources that help illustrate different components that drive freight movements in the region and profile the transportation system performance to support freight planning, modeling, and project prioritization.

Figure 1: Freight data and analysis framework



Ongoing, recurring data collection efforts are envisioned as the cornerstone of an effective freight data program. Based on the findings of this research, DVRPC will develop a multi-faceted data collection program that will augment and complement other travel monitoring programs (e.g., traffic, bicycle, and aviation counting) that are already in-place.

Data Inventory

A key component of the data collection effort was an inventory of freight data sources available for use in regional freight planning activities. This inventory was used as an early evaluation technique to establish the priority data sets that would be pursued in the full data collection effort. These data sets were evaluated for their practical uses as defined above, in addition to the cost to acquire, time/cost of processing, utility over

time, and potential value for similar efforts in other jurisdictions. The following section provides the inventory and evaluation of these data sources.

Freight Network Supply

The base, upon which all of the freight data is built, is the freight network supply. This network is a composition of modal networks including maritime, highway, rail, aviation, pipeline, and the intermodal nodes through which these networks interact. DVRPC has done extensive work in identifying various components of the freight network that constitute the supply of goods movement facilities. This effort will seek to refine the highway component of the network to be consistent with available freight data and core highway freight facilities.

Regional Primary Freight Highway Network

Source: DVRPC | Source date: 2014
Resolution: Segment (facility) | Cost: N/A

The Regional Primary Freight Highway Network is the highway network developed by DVRPC Freight Planning staff. This network is based on various state and national data sets and is geographically refined by internal GIS efforts. This network will require revision in order to include a more extensive coverage of facilities in the region. In addition to the existing DVRPC network, the HPMS network and HERE NHS+ network were identified for consideration.

Freight Activity by Mode

Classification Counts - Highway

Source: DVRPC | Source date: 2009-14
Resolution: Highway segment (point location) | Cost: $ (in-house)

Classification counts for the region are maintained by the Delaware Valley Regional Planning Commission. This point-location count data provides partial coverage of the regional highway network. Existing counts can be extracted using spatial conflation to the Regional Freight Network and applied to the appropriate segments in order to understand the current freight activity on the highway network. Additional counts can be acquired to fill existing gaps in the network.

WIM Truck Counts - Highway

Source: DOT | Source date: 2014
Resolution: Highway segment (point location) | Cost: free (public)

Weigh-in-motion truck counts are generated by Weigh-in-Motion (WIM) stations that provide total volume, as well as some classification of vehicle type. These counts would be capable of providing a regular update of freight activity on monitored facilities. This regular availability of data would contribute to the maintenance of a long-term freight data program.

Port Gate Turn Counts - Highway

Source: Terminal Owner/Operators | Source date: 2014
Resolution: Trip Generator/Intermodal Node (point location) | Cost: free (private)

Port Gate Turn counts provide the number of trucks entering and leaving individual terminal facilities. These counts are maintained by the individual facilities and can be requested in a standardized format through the data collection effort.

In order to establish a standardization technique for the data request, data could be collected from the two public port facility owners first. The data could then be requested from private terminal operators in the region. There is a risk that the data will not be shared by all or most parties, especially private terminal owners. This

data could be used to help disaggregate commodity flow values, such as tonnage, and help illustrate terminal activity. The gate turn counts also benefit future planning efforts related to NHS connector projects. These data sets along with detailed PIERS data could help DVRPC better understand the modal selection of various commodities being imported/exported through regional port terminals.

Annual Vessel Report - Maritime

Source: Maritime Exchange — Source date: 2014
Resolution: Port terminals (aggregated by terminal name) — Cost: free (private)

Annual Vessel Reports from the Maritime Exchange provide ship count data for each of the active port terminals on the navigable rivers in the region, the Delaware and Schuylkill. This data is maintained by the Maritime Exchange as a part of their larger role as a reporting and monitoring agency for port and river activity for the Delaware River and Bay. The primary purpose of the Maritime Exchange data collection is not to provide port terminal activity counts but rather to track ship activity on the river. The port terminal activity counts are a derivative of this process. As such, the details being provided by the Maritime Exchange are limited to ship type and count by port terminal. There is currently no information available about tonnage or value transferred at ports. This data does not include any barge activity.

This data is currently available to members of the Maritime Exchange and through this collection effort DVRPC worked with the Exchange to identify an appropriate user agreement and data format. This data is useful in disaggregating additional volume and value data available from other data sources. The ship count data is currently, and will continue to be used in *PhillyFreightFinder* to illustrate activity level at port facilities.

Waterborne Commerce Statistics - Maritime

Source: U.S. Army Corps of Engineers — Source date: 2012
Resolution: U.S. Army Corps Port — Cost: free (public)

The U.S. Army Corps of Engineers maintains several data sets through the Waterborne Commerce Statistics Center (WCSC). This data provides information such as total volume, value, and Twenty-Foot Equivalent Units (TEU's) for, at least, the top 50 Principal Ports in the United States. This data is available in various table formats and is not consistent with the U.S. Customs Port codes. In addition to high level summary data, detailed commodity data is available in aggregate for the ports on the Delaware River, though resolution is not consistent with other data sources or the high level statistics. The Waterborne Commerce Statistics data is useful for understanding overall port activity in the region. WCSC data can be integrated to the regional freight database for comparative statistics to other regions.

US Waterborne Foreign Container Trade – Maritime

Source: U.S. Maritime Administration — Source date: 2013
Resolution: U.S. Customs Port — Cost: free (public)

The data prepared by the U.S. Department of Transportation Maritime Administration provides multiple series of data for waterborne container trade by U.S. Customs Port. The data set includes government and non-government shipments through U.S. Ports, excluding Jones Act Trade and empty containers. The data is compiled from PIERS Data. The data provides annual TEU's and metric tons of containerized trade in total, import, and export series by port name.

The data series is available for each year back to 1997 and is usually made available for the preceding year by the end of the first quarter.

Train Counts - Railroad

Source: Class I Railroads | Source date: 2014
Resolution: Railroad segment | Cost: $

Train counts for the region can be requested from the Class I partner railroads. The resolution and type of counts (daily train, volume, etc.) that may be made available from these sources has yet to be defined. Cooperation on this effort has been limited and remains an identified challenge. Alternatives have been considered if railroads are not responsive, including manual video log counts and network mapping of road train service in the region.

Freight Economic Data

To support freight planning, modeling, and project prioritization, it is important to understand relevant economic factors that drive freight movements, and at the same time would be impacted by freight improvement projects. Important economic factors are the number of freight generators, size (such as employment, payrolls, and sales), and activities (production and attractions). The following are some data sources identified to be valuable for profiling freight related economic development, identifying freight generators, and forecasting freight demand.

National Establishment Time-Series (NETS)

Source: Walls & Associates | Source date: 2010
Resolution: Establishment (point location) | Cost: $$

The National Establishment Time-Series (NETS) database is a proprietary database produced by Walls and Associates. It is a revised version of the Dun and Bradstreet database. Using each company's unique Data Universal Number System number (or numbers, in cases where separate divisions within a company have unique DUNS numbers), Walls and Associates creates a time series for each business and then screens the data to eliminate duplicates and identify anomalies. In this establishment-based database, establishments are categorized by 6-digit NAICS code; all employment, sales, and other activities are reported at the actual facility—not the headquarters. In addition, the NETS database includes sole proprietors, part-time jobs, and farm operations, and has been found to be more accurate in reporting data for small, privately-owned firms and public sector employers such as post offices and public schools. The NETS database was purchased by DVRPC for employment forecasts. Municipal-level employment data from the 2010 NETS database was reviewed by DVRPC staff to eliminate obvious errors (such as duplications). The data was then reviewed by the region's county planning staffs and further refinements were made based on local knowledge (including errors in location and missing large employers).

Table 1: NETS Data Fields

COMPANY	Company and/or trade name
ZIP + ZIP4	Zip code + 4-digit
NAICS 2 DIGIT	NAICS code 2-digit
NAICS 6 DIGIT	NAICS code 6-digit
ADOPT_EMP	Establishment employment (see notes above)
TIM2SECTOR	Travel Improvement Model employment sectors
ADOPT_TAZ	TAZ code for modeling purposes
NAICS 6 DIGIT	NAICS code 6-digit

The NETS database can be utilized to disaggregate commodity flows, estimate freight activity in the region, and linked to DVRPC Freight Centers, to help further define the region's freight generators and attractors.

County and Zip Code Business Patterns

Source: U.S. Census Bureau | Source date: 2012
Resolution: County and Zip Code | Cost: free (public)

County Business Patterns (CBP) and Zip Code Business Patterns (ZBP) are Census products that provide annual statistics for businesses with paid employees by industry. This series includes the number of establishments, employment during the week of March 12, first quarter payroll, and annual payroll. This data is useful for studying the economic activity at the regional or sub-regional level; analyzing economic changes over time; and as a benchmark for other statistical series, surveys, and databases. The included annual payroll information is useful for profiling freight generation as a supplement to other freight employment data to account for seasonal employment and overtime hours at freight generators.

IHS Global Insights Freight Finder

Source: IHS Global | Source date: 2011
Resolution: Establishment (point location) | Cost: $$$

Global Insight's Freight Finder is a companion to its TRANSEARCH data service, adding specific facilities to the freight shipment information TRANSEARCH provides. The Freight Finder database contains 10,000 establishments in the Commonwealth. The data set is not nearly as comprehensive as the NETS data but does provide information related to the freight generation and attraction of businesses through the provision of inbound/outbound tons. The data purchased by the Commonwealth is limited to Pennsylvania establishments.

Table 2: IHS Freight Finder Data Fields

Business Information	Name, Contact, Address, Phone, Website
County	County location
SIC Code	Standard Industrial Classification
NAICS Code	North American Industry Classification System
Employment/Range	Number of employees/9 categories
Sales/Range	Amount of sales/9 categories
Inbound STCC/Tonnage	STCC4/tonnage 5 largest commodities
Outbound STCC/Tonnage	STCC4/tonnage 5 largest commodities
Total Inbound Tonnage	Total inbound tonnage
Total Outbound Tonnage	Total outbound tonnage

Torto Wheaton Database

Source: CB Richard Ellis | Source date: 2014
Resolution: Establishment (point location) | Cost: $$

Torto Wheaton is a database of commercial and industrial property, and associated market data. The data set includes all buildings with square footage greater than 10,000. It contains building square-footage data and other attributes about usage. The database can be purchased for various time series and with varying levels of aggregation. DVRPC has decided not to purchase this data as a part of this project.

CoStar

Source: CoStar | Source date: 2014
Resolution: Establishment (point location) | Cost: $$

CoStar is an online database of commercial and multifamily (rental apartments) property and associated market data, sales, and tenant information. It contains building square-footage data and other attributes about properties. It is primarily used as a tool for real estate brokers, but can be useful for research purposes. This

data may be useful for estimating freight generation at a finer geographic resolution. However, the data is known to have issues in terms of completeness and reliability due to brokerage reporting practices on available properties.

Additional Considerations

On the economic data side, the Make/Use tables from the Input-Output Accounts Data published by the Bureau of Economic Analysis may be useful to understand commodity flow and freight movements behind the scenes. However, since several sources of commodity flow data are already available to DVRPC in this study, this data was not integrated.

Commodity Flow

Commodity flow is a direct representation of freight movements and measure of freight demand in the region. A few data sources are available from both public and proprietary sources with different geographic resolutions and focus of mode. They are discussed in the following section. To support local freight planning, commodity flow at the county level would be ideal.

Freight Analysis Framework

Source: Federal Highway Administration — Source date: 2007
Resolution: FAF Zones- MSA/CSA by State — Cost: free (public)

The Freight Analysis Framework (FAF) is a FHWA funded and managed data and analysis program that provides estimates of the total freight flow in to, out of and within the United States; between individual states, major metropolitan areas, sub-state regions, and major international gateways. The main data source of the FAF is the Commodity Flow Survey (CFS). The current FAF data includes 123 domestic freight analysis regions or “zones”, consisting of 74 metropolitan regions, 33 regions (representing a state‘s territory outside these metropolitan regions), and 16 regions identified as entire states (within which no FAF3 metropolitan regions exist). The DVRPC region is covered by two FAF zones. The latest version is FAF3, which provides estimates for tonnage, value, and domestic ton-miles by region of origin and destination, commodity type, and mode for the base year of 2007 and forecasts through 2040. The next update of FAF based on the 2012 CFS is scheduled to be released in late 2015. FAF is a valuable tool to convey freight profile and statistics to the states and the public to support economic analyses. However, FAF, given its current aggregate format, is insufficient to support freight planning at the sub-regional (e.g., county) level. Thus, the FAF data will be not used directly (other than potentially benchmarking other data) for this study.

Freight USA

Source: Oak Ridge National Laboratory — Source date: 2007
Resolution: County — Cost: free (public)

Various research and modeling efforts have been undertaken to integrate the CFS with other data sources to develop commodity origin-destination (OD) flows at a finer geographic resolution. Freight USA is one of these freight data improvement efforts from the Oak Ridge National Laboratory. This data provides estimates of national freight OD flow (weight and value) by commodity type and mode at the county-to-county level. The commodity is categorized by 2-digit Standard Classification of Transported Goods (SCTG) codes. The mode of transportation follows that defined in the CFS. The majority of the development efforts for the Freight USA was in filling suppressed cells in CFS matrices. A freight generation and attraction model based on economic indicators, such as County Business Patterns and Make/Use Tables from Input-Output Accounts Data published by Bureau of Economic Analysis, was applied. All data was distributed to state, then to county level by the developed freight generation and attraction models. The final original-destination matrix was balanced using Interactive Proportional Fitting (IPF) methodology. This data is potentially valuable for sub-

regional freight analysis and it is free of charge to government agency users. With TRANSEARCH data readily available, this data will not be integrated (other than as a potential source for comparison).

IHS Global Insights TRANSEARCH

Source: IHS Global | Source date: 2011
Resolution: County (BEA outside region) | Cost: $$$

TRANSEARCH is a proprietary data product from Global Insight Inc. that provides estimates of freight flows across the United State at the state, business economic area (BEA), and county levels by commodity type and mode of transportation. The commodity is categorized by 4-digit Standard Transportation Commodity (STCC) codes. The geographies include 172 BEAs and over 3,000 counties. TRANSEARCH data has been purchased and processed by PennDOT as a part of the Commonwealth's freight plan updates. TRANSEARCH data is available for the entire DVRPC region, including member counties in New Jersey. The base year of data is 2011 and forecasts include 2015, 2020, 2030 and 2040.

Table 3: IHS TRANSEARCH Data Fields

Year	Base or scenario year
Origin Region	Origin region number
Destination Region	Destination region number
STCC	STCC code as alpha-numeric code (XX_XX)
Mode	Mode number (15 modes in 6 mode groups)
Tonnage	Decimal tons
Value	Decimal value
Unit	Decimal units (by commodity)
STCC4	4-digit STCC node (XXXX)

The county-to-county freight data illustrates freight transportation demand at the regional and county level, supports local freight planning, and prioritizes freight improvement projects at key locations. Using TRANSEARCH data will also build a direct connection to freight planning and analysis at the state level. In this project, TRANSEARCH data might be further evaluated at the regional level. Currently, Global Insight Inc. benchmarks the flows against reported freight volume data using two primary sources: 1) private carrier information that they acquire as part of a data exchange program with railroads and truck carriers; and 2) truck count information released by the state Departments of Transportation.

Public Use Waybill Data - Railroad

Source: Surface Transportation Board | Source date: 2012
Resolution: National Sample | Cost: free (public)

Public Use Waybill data is a sampling of rail transportation shipment data from waybills. The data contains origins and destinations, as well as key shipment details including STCC commodity codes. This data is available as a 2% sample file through the state Departments of Transportation. Data can be used to identify carloads, commodity, freight revenue, and tonnage.

The detail fields of the data provide information about the routing and car types that can be useful for assigning commodity flow data to the regional network.

PIERS - Maritime

Source: PIERS | Source date: 2014

Resolution: U.S. Customs Port | Cost: $$$

PIERS data is transcribed and standardized import/export data for U.S. port activity. This proprietary dataset is maintained as a database with individual bills of lading stored with standardized values allowing for better quantification of volume and OD information for various commodities at a port level of granularity.

Table 4: PIERS Data Fields

Cargo Type	Containerized	Non-Containerized
Direction	Imports	Exports
Vessel Arrival Dates	Month/Quarter/Year/Last 3 months	
Country/Geography	Country	Country name & code
	Foreign port	Foreign load or discharge port
	Origin/Destination city	Origin or destination city of cargo
	Ultimate port	Ultimate origin or final destination
	U.S. port	U.S. load or arrival port
Commodity	Jock (TSUSA) code/description	Seven-digit Jock (TSUSA) code/description
	Harmonized code	Six-digit Harmonized Tariff code/description
	Key commodity	Description of a product group
Shipping Line/Vessel/NVO	Shipper/Consignee name	Exporter/shipper or importer/consignee name
	Shipper/Consignee zip code	Shipper/Consignee zip code
Container	Container size	Two-character field denoting container size
Packaging /Cargo Flags	Hazardous	Hazardous material flag
	Refrigerated	Refrigerated cargo flag
	RORO	Roll on/Roll off cargo flag
	Unit of measure	Type of packaging

The Philadelphia Regional Port Authority (PRPA) and South Jersey Port Corporation (SJPC) purchase this data for research and marketing efforts. DVRPC can leverage this maritime commodity flow data to fill existing gaps in the import/export of commodities from the TRANSEARCH database. In addition, the PIERS database would allow DVRPC to better illustrate the network demand and trip generation created by various types of port traffic (ship types, commodities). The data once integrated to the freight database, if aggregated properly, could be represented in *PhillyFreightFinder* to help outside users better understand the movement of various commodities through the regional port system.

DVRPC identified one risk with using this data and it is related to the existing user agreement. Before making use of this data, DVRPC will work with the PRPA staff to determine if the user agreement allows for sharing of aggregated data extracted from the PIERS database.

USA Trade Online – Maritime & Air

Source: U.S. Census Bureau | Source date: 2014
Resolution: U.S. Customs Port | Cost: $

The USA Trade data tool provided by the U.S. Census Bureau's Foreign Trade division is a subscription based product providing access to international trade data. The tool provides access to import/export data in various formats. Reports can be run at the district, port, state, and national level. For each geographic resolution, data can be extracted by select commodity classification codes, by value and shipping weight. For a local freight data program, the port resolution was determined as the most viable for providing detailed regional activity and statistics.

The port resolution USA Trade Online data is provided for two-, four- and six-digit Harmonized System (HS) level commodities. Three modes of transportation are provided, including air, vessel, and containerized vessel. The data is available from 2003 to current date with a month delay on release. Custom reports can be requested for each of these variables and the ports and trade countries of interest. Each report is limited to imports or exports.

US EIA Company Level Imports – Maritime (petroleum)

Source: U.S. Energy Information Administration | Source date: 2014
Resolution: U.S. Customs Port | Cost: free (public)

U.S. EIA Company Level Imports data are provided on a monthly basis, sourced from Form EIA-814. This data covers petroleum imports to U.S. territories.

Table 5: Company Imports Data Fields

RPT_PERIOD	Report period in YYMM format
R_S_NAME	Importing company name
PROD_NAME	Product name
PORT_CITY	Port of entry
PORT_STATE	State abbreviation
CNTRY_NAME	Country of origin name
QUANTITY	Import quantity (thousand barrels)
PCOMP_RNAM	Processing company name
PCOMP_SNAM	Processing facility name
STATE_NAME	Processing company state

This data provides detailed import data for bulk petroleum products for the region. The data can be used to disaggregate and validate commodity flow data sets that are available at higher resolutions. In addition, the Company Level Imports data sets provide detailed information about petroleum sourcing by refining units and can help to illustrate and communicate changes in the supply change over time.

Freight Performance Measures

Highway freight performance is the focus of this project. Performance measures for other modes are either difficult to define or infeasible to collect data, such as on the rail side performance. For the highway mode, truck travel times in freight-significant corridors (ideally for the entire regional primary freight network) are direct performance measures. Truck travel times (and reliability) indicate how well the regional highway network is being used in supporting freight movements and in understanding activity by time of day. Two data sources are identified here.

National Performance Management Research Data Set

Source: FHWA (HERE) Source date: 2014
Resolution: NHS by Traffic Message Channel (TMC) Cost: free (public)

National Performance Management Research Data Set (NPMRDS) is a national data set of average travel times made available through FHWA to MPOs for performance measurement. This data includes the HERE probe data for cars and the American Transportation Research Institute (ATRI) probe data for trucks. The data covers the National Highway System (NHS)—the same base for the regional primary freight network and is updated on a monthly basis starting in July 2013. Only observed field probe data points are included in this data. If no field data is available, no travel time is reported and no data modeling, filtering, blending or smoothing is applied to the raw observed values reported in this data set. As a result, the sample size may be limited, but the data quality may be more (or less) accurate than other data sets. This data can be utilized for performance measures, since it is the only data source for truck travel time.

I-95 Vehicle Probe Project

Source: I-95 Corridor Coalition (INRIX) Source date: 2014
Resolution: Highway Down to Arterial by Traffic Message Channel (TMC) Cost: free (I-95 members)

I-95 Vehicle Probe Project (VPP) is another historical travel time data set available to DVRPC and participating public agencies through the I-95 Corridor Coalition. The included travel time data is provided through INRIX as best estimates of speed and travel time. If density of probe data is insufficient, modeling and historical data may be used in addition to field data to estimate traffic conditions. The data has more coverage, including all interstates and major arterials (archived back to January 2009) when compared to the NPMRDS data, and more data validation has been undertaken in the DVRPC region. However, the VPP data does not separate trucks' speed and travel time from cars. The VPP data may be used to supplement the NPMRDS data as performance measures for freight-significant corridors not covered in the NPMRDS data. The truck speed and travel time reliability may be inferred from car speed, based on observations of trucks versus cars in similar corridors, if necessary. Using the VPP data also involves mapping TMC segments (the TMC codes could be different between the VPP and NPMRDS data) and calculating performance measures from the archived data. Thus, comparisons of these two data sets will be limited.

Data Collection Priorities

The data collection plan developed from the above inventory identified a priority set of data that would be collected in the project. Throughout the course of the project several issues arose in the collection of data resulting in a shift to alternate sources, as noted below. The following data sets were collected throughout the SHRP2 C20 project.

1. Primary Regional Freight Network – NHS and NHS connectors
2. TRANSEARCH - commodity flow
3. USA Trade – maritime commodity flow (international)
4. US EIA Company Level Imports – maritime commodity flow (petroleum products)
5. NETS – freight economic and generators
6. CoStar – freight econonmic and generators
7. NPMRDS – truck speed/performance measures
8. Classification counts – truck activity
9. WIM truck counts – truck activity
10. Port gate turn counts – port/truck activity

11. Maritime Exchange vessel counts – maritime activity
12. Waterbourne Commerce Statistics [Principal Ports & NDC Links] – maritime activity
13. US Waterbourne Foreign Container Trade – maritime activity (containerized cargo)
14. Train counts – rail activity

Data Processing and Analysis

Database Design and Implementation

Following the identification of data sets that would best fit the practical needs of freight planning in the region, the project team undertook a database design process. The database design was created to guide the integration process and serve as a tool for future data collection and sharing efforts. The project team was able to leverage an ongoing agency effort to modernize the existing ArcGIS database and server systems during this effort. The resulting database design utilizes a dedicated, centralized freight tablespace in the Oracle database. All network data is georeferenced and available for use in ArcGIS Desktop across the agency through an ArcSDE connection.

The database schema is available in Appendix A. Regional Commodity Flow data was stored in the same tablespace using its native schema.

Data Processing

The data collection and integration effort undertaken in this project, was conducted jointly by the Office of Freight and Aviation Planning and the Office of Modeling and Analysis. The project team, with the help of key stakeholders and partners undertook this effort over the course of four months. The following documentation provides key details about the processing that was done through this effort.

Network Improvement

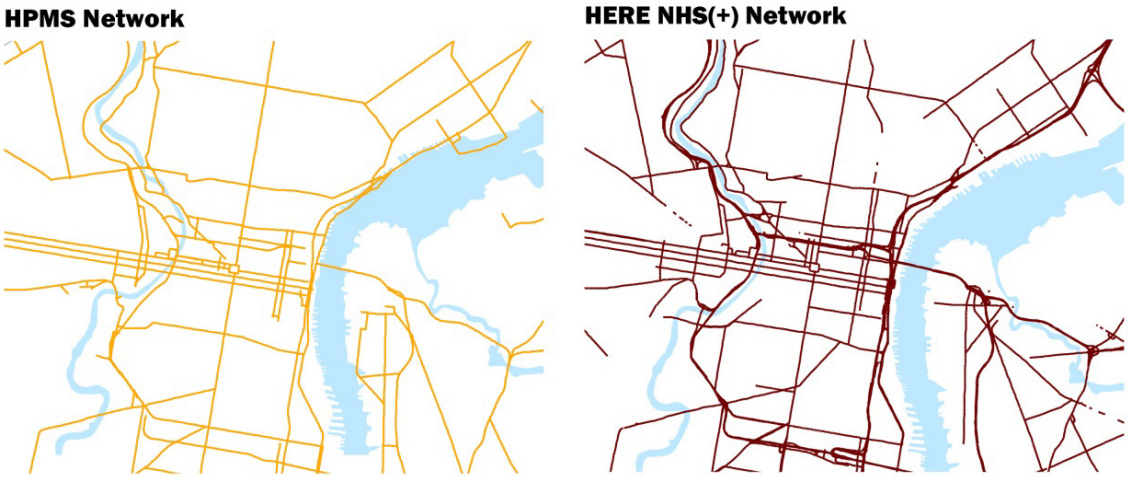
A major component of the data processing conducted through this project was related to the refinement of the existing *PhillyFreightFinder* network components. Throughout the effort, all data sets that included a “network” component were geocoded or matched to their related facilities. For some this also involved improvements to the geometry or segmentation of these facilities. In this project, the refinement of the Primary Regional Freight Network was the most comprehensive data processing done related to network improvement. The process is documented in this section.

Primary Regional Freight Network

The data collection plan included the identification of a Primary Regional Freight Network for the region. The existing *PhillyFreightFinder* network designates Interstate routes and limited-access highways as the “Freight Highway Network”. These facilities were deemed too limited to be a comprehensive representation of the freight network for truck mode movements in the region. An evaluation of alternatives was conducted utilizing networks identified in the data inventory.

Figure 2 is a comparison of the HPMS network versus the HERE NHS+ network, highlighting the difference in coverage. This project utilized the HERE NHS+ network for performance measures as it is the base network for the NPMRDS dataset. The project team chose not to create a crosswalk between the two data sets, and instead utilized the HERE NHS+ system as a base from which to work. In order to provide a better representation of the freight highway network in the region, the project team improved the HERE NHS+ network by including all NHS connectors, routes known to have higher than average truck percentage, and any segments that improved the connectivity of the network.

Figure 2: Highway network comparison



While the project team chose to keep the new network independent of other internal highway networks; creating a crosswalk with other key networks was identified as a future effort. The project team did not feel that enough data analysis could be conducted in this effort to define a more refined Primary Regional Freight network. This has been identified as a priority next step.

County Trade and Economic Data

The data collection plan had identified several datasets related to economic trade, employment, and commodity flows. These data sets in their raw form provide value to various efforts that are conducted by DVRPC. In an effort to make this data more usable on a sub-regional level, the project team processed each of the data sets for inclusion in county and economic analysis.

TRANSEARCH

The IHS TRANSEARCH database was identified as the primary commodity flow data source for the purposes of this study. The county level commodity flow data provided both current and future flows in the region. This data could be integrated directly into the *PhillyFreightFinder* database. The project team, working with county partners and other stakeholders, identified a series of reports and views of this data that would provide county planners with key data metrics.

In addition, the dataset was utilized to support multiple planning efforts in the region over the course of the project. These projects included a regional port impact study, a marine highway market analysis, and preliminary data analysis for a port market study. This commodity flow data set will be utilized for future efforts as is, and could be further disaggregated to better illustrate and calibrate freight models in the region.

NETS

DVRPC conducted a thorough processing of the NETS database in earlier efforts as a part of regional forecasting work. This work modified and cleaned the NETS database and was reviewed at the municipal level by county planners. The Local Freight Data Improvement project team chose not to replicate this effort or further enhance the accuracy of this data. This decision was made in order to keep the freight planning and employment work at the regional level consistent with other agency employment studies.

The project team chose to utilize this employment dataset to better understand the distribution of freight related employment. The georeferenced establishment data was utilized to refine the current geographic

extents of the DVRPC Freight Centers. This data was also aggregated at the county and freight center level to better communicate the role of freight employment at these geographic resolutions.

CoStar

DVRPC had also previously subscribed to the CoStar dataset for the purpose of employment and economic impact studies. Through the Local Freight Data Improvement project, this dataset was aggregated in a similar fashion to the NETS data. Facility attributes such as type, square footage, and occupancy status were aggregated in order to better illustrate the physical capacity of freight generation and attraction at the Freight Center and County levels. This point data set was used with the NETS data to further refine the geographic extent of the DVRPC Freight Centers.

The CoStar facility point data was not included at the individual facility level in the *PhillyFreightFinder* database due to the rate of change that occurs in this database. Instead the CoStar data was included in aggregate summary format at the freight center and county level. In the event that the agency should chose to discontinue its subscription to CoStar, the full point data set for the region will be extracted and included in the *PhillyFreightFinder* database.

Highway Performance

The evaluation and processing of data related to highway performance was a critical component of the data processing effort. The project team utilized several datasets to create the foundation for a highway performance analysis. The processing of this data is documented in this section.

NPMRDS

The DVRPC project team utilized the NPMRDS truck data for the region's network, as defined by the provided HERE NHS+ base network. The development of this network was covered in the Primary Regional Freight Network section of this report.

The NPMRDS data was accessed for the purpose of creating a series of truck performance measures for the highway network. The project team downloaded the necessary network shapefiles along with several years of corresponding data. Using GIS, the project team extracted the DVRPC network and identified the corresponding TMC codes that existed in the region. Utilizing the matched TMC codes for approximately 3,500 segments, a process was developed to import matched activity records to a database that would contain only records for the DVRPC NPMRDS network.

The resulting database served as the basis of further analysis in order to quantify the respective variables for several highway performance measures. The project team, utilizing R, conducted statistical analysis to eliminate outliers and calculate average speed, travel time index, and highway performance index. Due to a desire to identify multiple time bins throughout the day and issues with low sample sizes; the project team used 18 months of weekday data in order to provide necessary outputs with an acceptable level of confidence. The average speed and Travel Time Index (TTI) for nine distinct time bins were illustrated from this data. These time bins included:

- midnight to 5:00 AM,
- 5:00 AM to 7:00 AM,
- 7:00 AM to 9:00 AM,
- 9:00 AM to 11:00 AM,
- 11:00 AM to 1:00 PM,
- 1:00 PM to 3:00 PM,
- 3:00 PM to 5:00 PM,
- 5:00 PM to 7:00 PM,
- 7:00 PM to 9:00 PM, and
- 9:00 PM to midnight.

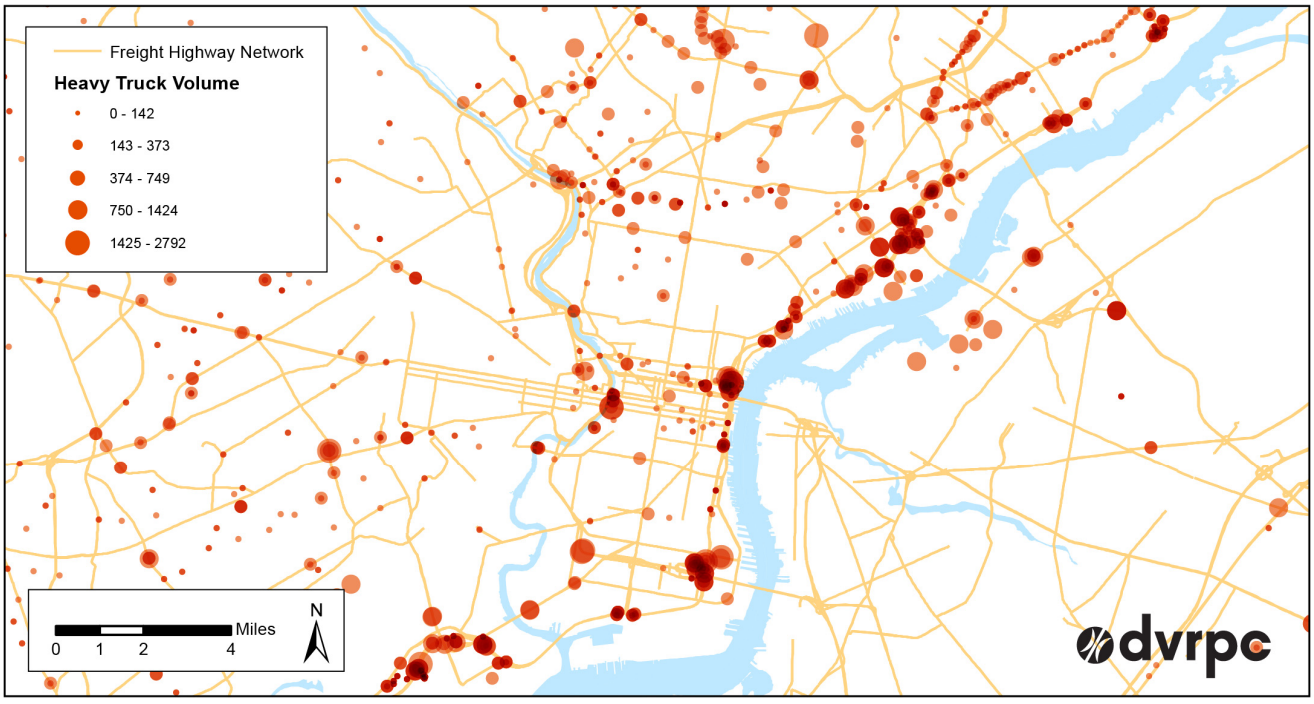
The resulting indices for each of the TMC during each time bin were joined to the geocoded TMC network. The project team, utilizing guidance from comparable analysis created a scale for each index to categorize the resulting scores from good to bad in a color coded sequence. This allowed the project team to create a map product that demonstrates performance for the network throughout the day. The NPMRDS network and its related indices table were added to the *PhillyFreightFinder* database. The project team determined that based on the data that supported each index, average speed and travel time index would the priorities for communicating performance in *PhillyFreightFinder*. These were included in the "Highway Performance" tool as described in the next section of this report.

Appendix B contains a complete documentation of the methodology utilized for processing the NPMRDS data.

Classification counts

DVRPC conducted an analysis of the manual and machine class counts that existed in the Travel Monitoring count database. It was found that a high proportion of the existing counts were located in the Pennsylvania counties, a result of an ongoing count program with the Pennsylvania Department of Transportation. Notable gaps in the system were the interstate system, which traditionally have not had classification counts conducted due to technical restrictions. In addition, the project team identified gaps on the existing NHS Connectors in the region and compiled a collection plan for future counts. Below is a sample of the current coverage in the DVRPC region for class counts. The highway network is the DVRPC modified HERE-NHS network.

Figure 3: Classification count coverage in DVRPC region



The counts were not integrated into the *PhillyFreightFinder* but were integrated internally for analysis of total delay. Future efforts will be necessary to provide enough coverage for a region wide analysis of delay. Further evaluation will be conducted over the coming years to identify opportunities to fill gaps in the classification

count coverage in both New Jersey and existing freight routes that do not have proper coverage in Pennsylvania.

WIM truck counts

The state Departments of Transportation for Pennsylvania and New Jersey were identified as sources for WIM data. While the Pennsylvania Department of Transportation has only one active WIM station in the region, New Jersey has twenty stations. Data from these stations was available online as PDFs. Through the project DVRPC was able to access additional formats and reports from the New Jersey Department of Transportation. The project team determined that it would not be appropriate to become a secondary source for these counts but should instead work with the NJ DOT to find a better method for accessing the data created by the WIM stations. During the course of the project an arrangement was not able to be developed. However, this was due to personnel and time constraints, and not a lack of willingness to share the data. DVRPC will continue to pursue this effort and create a streamlined method of linking NJ DOT's WIM data with the DVRPC highway network.

As a temporary measure, the project team did integrate the last two years of WIM data at a summary level for each of the region's WIM stations. This data is linked to the matching highway segment in the internal *PhillyFreightFinder* database.

Maritime Activity and Indicators

The maritime freight mode was the mode with the single highest quantity of data sources available to the project team. Each data set had unique geometry and measurement types. The project team utilized each data set to create a series of indicators that could be used for tracking maritime activity over time. This data was also integrated into the internal *PhillyFreightFinder* database for use in port related planning efforts. The following provides documentation on the processing of the maritime data.

Maritime Exchange Vessel Counts

The Maritime Exchange had been informally sharing its annual vessel report with DVRPC for planning purposes for several years. In an effort to ensure that this data would be consistently available to DVRPC for planning, the project team explored the process through which this data can be attained in the future. As a result, the data will be purchased for a small fee as a part of the DVRPC's membership with the Maritime Exchange. The project team identified a custom report that will be run once annually by the Maritime Exchange and delivered in Comma-Separated Values format to DVRPC staff. A custom built script was developed to process this digital file to append the data to the existing *PhillyFreightFinder* "Ship Calls" data table.

Waterborne Commerce Statistics

Waterborne Commerce Statistics covers a variety of data from the U.S. Army Corps of Engineers. This data improvement project identified the Principal Ports file as a key data source for illustrating the performance of regional port terminals and trade versus national maritime trade performance. The project team worked closely with the DVRPC Office of Long-Range Planning and Economic Coordination to determine how this data along with other indicators would fit into the "Regional Indicators" program. It was determined that ranking the MSA by port activity versus other MSAs nationwide, as well as change in import and export activity versus national averages would be an optimal output of this dataset.

The project team developed a script for processing the Principal Ports data and appending this to the *PhillyFreightFinder* database as a summary table. As a high level aggregate dataset, this data is intended primarily as an indicator of port trade change and performance over time.

U.S. Waterborne Foreign Container Trade

In order to expand on the WCSC data for total volume and USA Trade data for HS level import/exports, the project team identified the U.S. Waterborne Foreign Container Trade dataset as a reliable source for measuring container activity by TEU. This dataset is made available at the U.S. Customs Port level which was consistent with USA Trade data but not with the WCSC data which utilizes a U.S. Army Corps port definition.

For the purposes of refined analysis of individual container port performance, port terminals in the existing DVRPC network were matched to the appropriate U.S. Customs Port. The region contains five unique U.S. Customs Maritime Port designations and a total of six container handling terminals. The only Customs Port requiring disaggregation was the Philadelphia Customs Port which could be accurately determined with the assistance of public reports provided by the Philadelphia Regional Port Authority, the public authority that leases all three of the container terminals in the Philadelphia Customs Port. The resulting data was stored as a linked "Container Activity" table in the *PhillyFreightFinder* database.

USA Trade

Early in the data collection process the project team and the PRPA had identified a possible shared user agreement for access to PIERS data. However, given the complexities of such an arrangement and personnel changes, the project team utilized the USA Trade product from the U.S. Census Bureau as an alternative. This data is available at various resolutions and formats but was accessed for the two-digit HS code for all terminals in the DVRPC region as well as for the national total. The project team developed a Python tool for data intake that streamlines and partially automates the process, though manual report downloads are still necessary.

This data set had been identified for the purposes of assessing the international trade performance of the region's ports versus the nation as a whole. A data view was developed to illustrate this data that aggregates all regional terminals and commodities by direction as well as the comparative national statistics.

The use of HS codes for commodity identification creates compatibility issues for more detailed analysis linked to domestic trade patterns. Future efforts could explore linking domestic and international trade patterns as well as work to disaggregate data from the U.S. Customs Port level to the DVRPC Port Terminal level.

U.S. EIA Company Level Imports

The U.S. EIA Company Level Import data was included with this effort as an extension of a previously undertaken commodity profile of crude oil in the Delaware Valley. Company Level import records provide a key metric to measure shifts in petroleum crude oil trade over time. The critical role of petroleum products to the regional port system and economy, justified including this data for tracking purposes to better understand key shifts in the petroleum supply chain. No mode is provided for this, however, using port of entry, a mode was determined (rail or maritime) for all shipments destined for regional refineries.

The project team developed a Python script to automatically download and append monthly data records to the *PhillyFreightFinder* database.

Communication and Sharing

The primary outcomes of the Local Freight Data Improvement project went beyond the delivery of a new database and data collection program. Core to the project was the enhancement of communication and visualization tools. The primary tool for DVRPC's communication of freight data is *PhillyFreightFinder*, an online application and data portal. The project team had two goals in this project related to the *PhillyFreightFinder* application: create new tools and visualizations for the enhanced data, and create an open-source version of the application for replication by other agencies. In addition to these primary outcomes, the project team identified several enhancements to the planning process.

PhillyFreightFinder Tool Enhancements

Enhancements to the *PhillyFreightFinder* application were a critical outcome of this project, as it serves as the primary vessel for communicating freight data to DVRPC's freight stakeholders. Throughout the project, the team collected input from the steering committee and planning partners at the county level. This input helped to guide the type of tools that would be most useful to build into the application in this project. These tools made use of the data collected in this project and communicated it through visualizations that illustrate freight activity and/or performance in the region. The following provides a brief overview of each of the three new tools created in this project.

Maritime Indicators Tool

The port system of the Delaware Valley is a critical component of the economy and freight transportation system. Measuring and communicating the performance of this system is a priority for the Office of Freight and Aviation Planning. While the agency has not yet fully defined a set of performance measures, the project team felt that indicators that track activity through the port would be a valuable tool. The maritime activity tool can be broken out into three distinct components; annual indicators, vessel activity by terminal, and maritime activity over time.

The annual indicators widget was designed to communicate performance of the port year over year. The project team identified five indicators that help to explain the regional port system's performance. Several of the indicators are based on a change in the region versus a change at the national level to better demonstrate the region's performance versus current national trends.

- **Regional Port Rank by Volume:** This ranks the regional ports versus all other MSA regions in the nation with ports based on the total volume of shipments. This indicator utilizes U.S. Army Corps Principal Port data aggregated to the MSA level to calculate rank.
- **Total Trade (tons):** This is an indicator that assesses the change in total trade, measured in tons, from the previous year. The indicator itself is determined based on the change relative to the national change. The indicator utilizes domestic trade date from the Principal Ports dataset and international trade from the USA Trade database to estimate a total trade for the region. The national change was calculated utilizing all Principal Ports activity and USA Trade for all U.S. Ports.
- **Total Foreign Trade:** This indicator explores the value of maritime foreign trade and the change year over year. The indicator is determined based on the change of foreign trade value in the region versus the change for all U.S. maritime ports. The indicator utilizes USA Trade for the value of maritime foreign trade.
- **Maritime Exports (tons):** This indicator assesses the change in export trade, measured in tons, from the previous year. The indicator is determined based on the change in export tonnage in the region versus change in exports for all U.S. maritime ports. The indicator utilizes USA Trade for the tonnage of exports.

- **Export Share of Foreign Trade:** This indicator shows the percentage of foreign trade that is export, for the region. The indicator utilizes USA Trade for the calculation of exports as a share of total foreign trade.

The second component of the Maritime Indicators Tool is vessel activity by terminal. This simple diagram, as shown in Figure 4, illustrates the 31 marine terminals in the DVRPC region, color coded by facility type and scaled based on the number of ship calls the terminal received that year. The user is able to change the year to see how terminal activity varies over time. Mousing over an individual terminal provides details on the number of ship calls and the percent change from the previous year of counts.

The third component of the Maritime Indicators Tool is maritime activity over time. This interactive graph, charts the volume of domestic, import, and export shipments over time from 2003 through the last year with full data available. The domestic and foreign trade data is based on the same sources as the maritime indicators. In addition, a second series charting a single line represents containerized cargo activity, measured in twenty-foot equivalents (TEUs). Mousing over the chart provides a pop-up that lists the details for each data point.

Below is a screenshot of the Maritime Indicators Tool that is available at **www.dvrpc.org/phillyfreightfinder#maritime-indicators**.

Figure 4: Screenshot of Maritime Indicators Tool

Highway Performance Tool

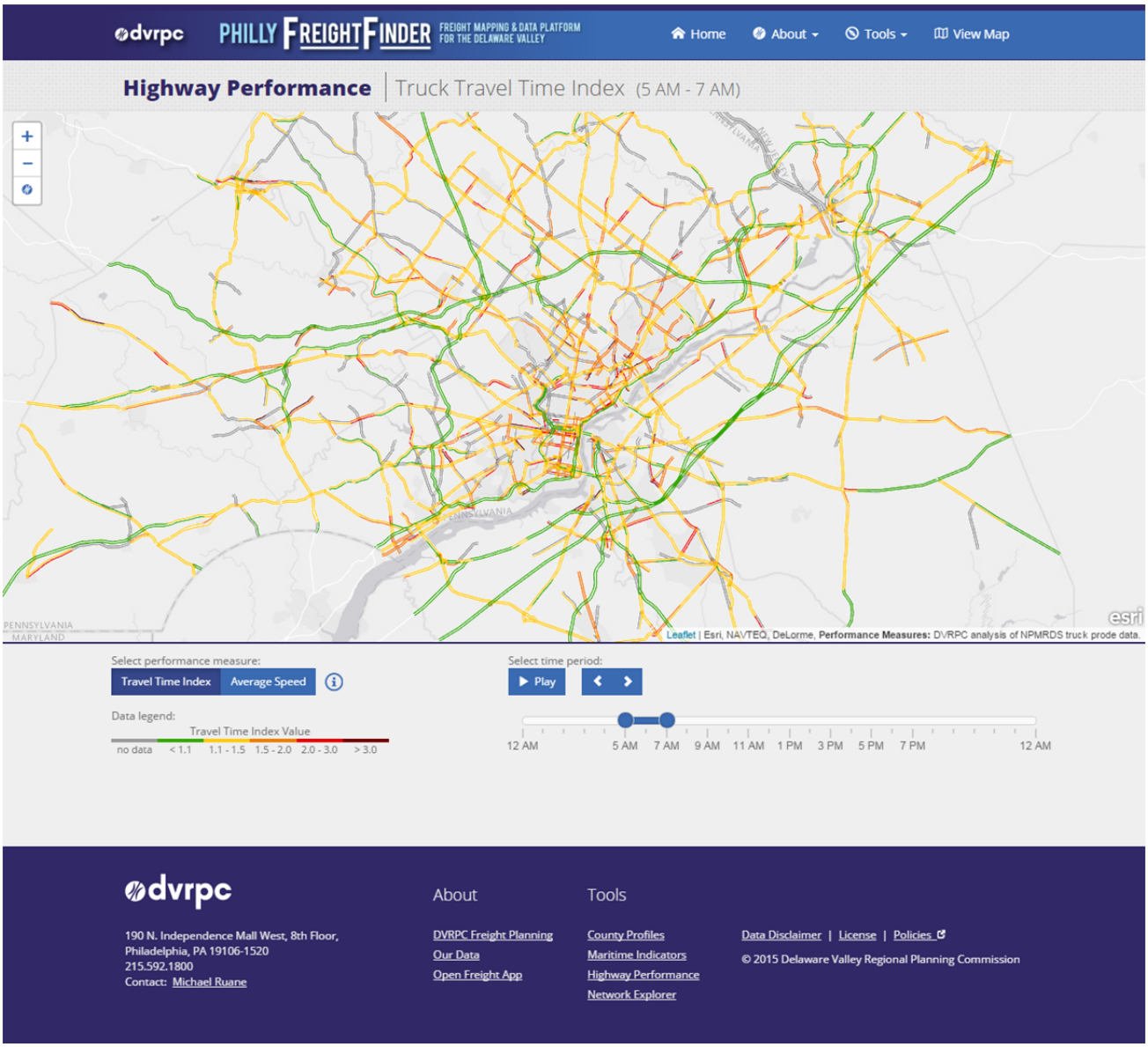
An efficient highway system is integral to the freight economy and transportation system. While many tools currently exist to planners, such as the VPP Suite and maps produced by various agencies illustrating congestion, few provide a truck focus. In addition, many are available only through subscription or are provided as static graphics.

This tool was designed to make full use of the NPMRDS dataset to illustrate two performance measures for the network throughout the day. The indicators are Truck Travel Time Index and Truck Average Speed. The tool loads a map service of pre-rendered tiles for each of the indicators for each of nine time bins throughout the day. Users can page through the time bins or play an animation that automatically pages through each of the time bins, showing the shifts in performance across the network throughout the day. Additionally, users have the ability to zoom in on select regions or routes to view in more detail the network elements of interest to them.

Future enhancements will require further data processing, in order to provide aggregate data at the route and/or county level.

Below is a screen shot of the Highway Performance Tool that is available at **www.dvrpc.org/phillyfreightfinder#highway-performance**.

Figure 5: Screenshot of Highway Performance Tool



County Profile Tool

DVRPC's primary planning partners are our county members. To better facilitate the sharing of freight activity and economic activity data with these partners, the project team developed the County Profile Tool. This tool provides two distinct components for each of the counties; network and trade patterns.

The network component of the profile tool provides the user with key statistics, images, and a map of the county freight system. Freight network statistics are compiled based on the network capacity and activity levels documented in the *PhillyFreightFinder* database. These are compiled and presented in a simple table. In addition, a static map provides an overview of the county's key freight facilities and images of freight in the region help the user to understand in a visual way, what freight on the ground looks like.

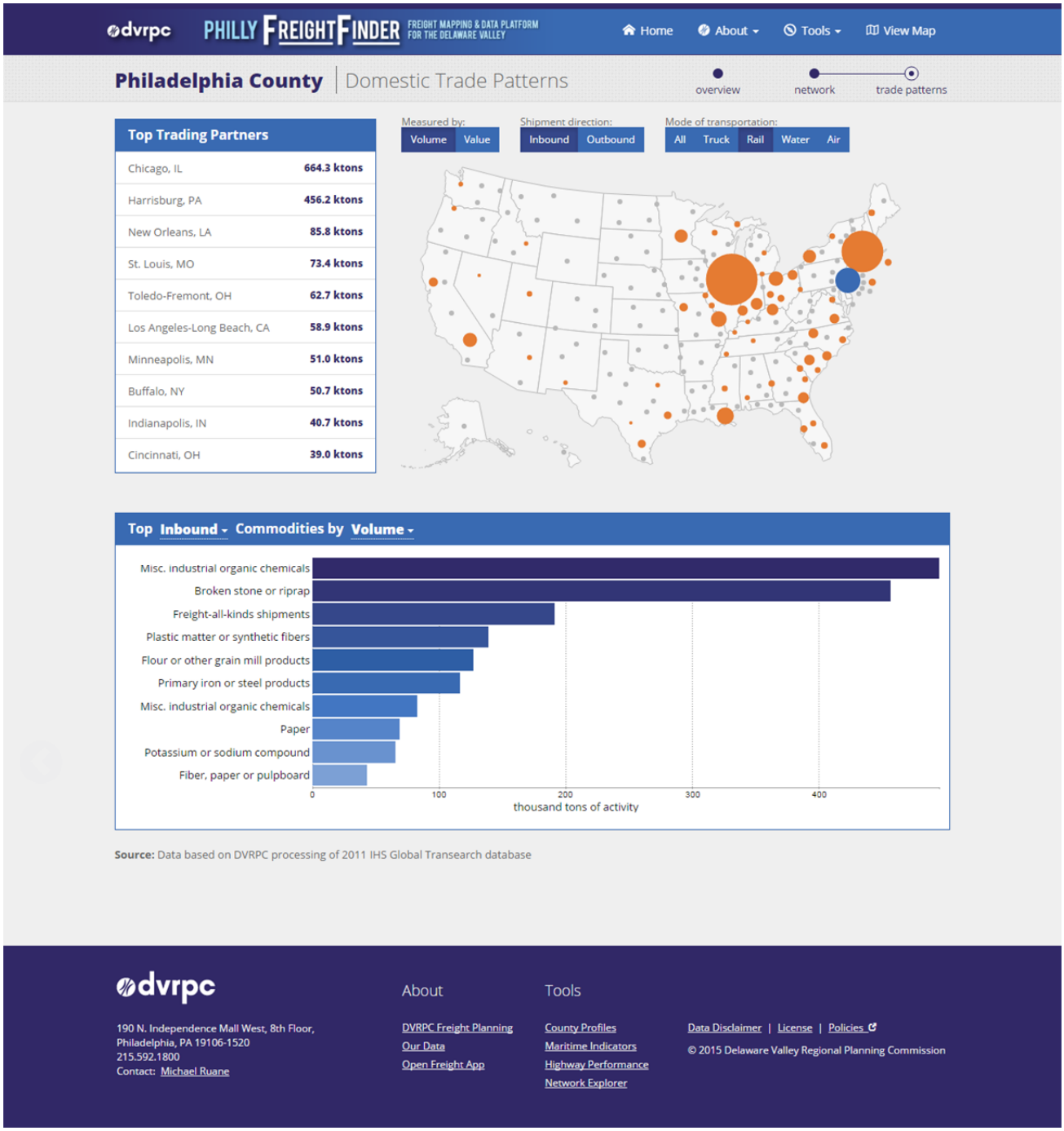
Figure 6: County Profile Tool – network view

The trade patterns component of the profile tool provides users with a way to explore domestic trade for the county. The data is based on 2011 TRANSEARCH commodity flow data. This component summarizes the data for the county partners, providing outputs that they can use for planning purposes. The data can be filtered and viewed in several ways:

- **by measure**: volume or value;
- **by shipment direction**: inbound or outbound; and
- **by mode of transportation**: all combined, truck, rail, water, or air.

For each of the filter combinations the viewer provides the top ten trading partners and their corresponding trade, the top ten commodities traded, and a map illustrating the geographic distribution of this trade across the United States.

Figure 7: County Profile Tool – trade patterns view



Open Source Template

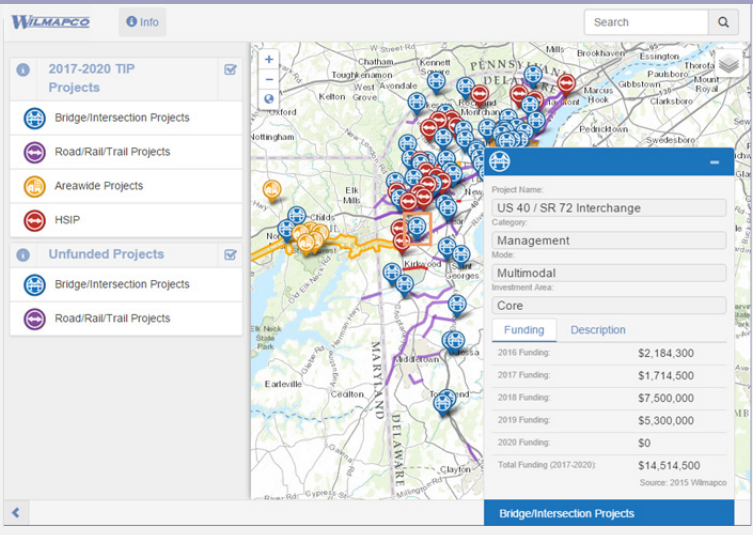
A key outcome of the project was to develop a way for other agencies to replicate the *PhillyFreightFinder* interface. Through the provision of an open source template and documentation, known as Open Freight App, the DVRPC project allows other agencies to leverage the basic template to easily make their freight network data visible to the public. By utilizing the Open Freight App template, rather than building from scratch, planning agencies can focus their time and resources on the development of data sets that can serve to improve the access to information on freight facilities and their role in economic development and transportation across the country.

The Open Freight App was developed to provide the basic functionality of the *PhillyFreightFinder* application. This functionality includes the mapping and identification of network facilities in a web based portal that can utilize both services and simple web file formats. The goal was to create a tool that could be replicated without a need for high cost infrastructure such as ArcServer capabilities to serve map feature services. This template can utilize simple JSON files created in any GIS software suite or, for those with higher capacity infrastructure, make use of feature services.

The documentation included with the project helps users step through the process of customizing and loading data into the tool. The tool has been built with comments throughout to help guide users in the HTML and JS code. In addition, helper functions have been developed to streamline the development process. While the tool does not make full use of a MVC framework, the simple process established in the documentation makes this product able to be utilized by most users that have basic understanding of HTML, JS, and CSS languages.

Case Study: WILMAPCO TIP Map

During the project, the DVRPC team conducted a limited release of the Open Freight App to MPO's that expressed an interest in replicating PhillyFreightFinder. Wilmington Area Planning Council (WILMAPCO) was an early adopter that was able to launch a product during the timeline of the project based on this template. WILMAPCO chose to utilize the application template to create a tool that visualizes and provides information about their Transportation Improvement Program (TIP). The staff at WILMAPCO, with the assistance of an intern, were able modify the template and launch a product for use by regional stakeholders.



WILMAPCO TIP Map available at **www.wilmapco.org/Tip/fy2017/**

The fully documented and commented code project is available on GitHub, a web-based Git repository service. Through this project repository, users can clone and customize the project for their own purposes. In addition, changes and improvements can be contributed back to the template project, or posted as improvements and bugs are resolved. The project is provided with an MIT License. The full project includes:

- Full documentation and guide on:
  - Adding map functionality for layers and features
  - Creating info windows for feature data
  - Creating and modifying legend and layer controls
  - Populating search criteria
  - Converting ArcGIS Shapefiles to lightweight GeoJSONs for use in the map
- Fully commented versions of all necessary HTML, JavaScript, and CSS files
- A starter set of freight facility icons
- License and credit materials

The full project and further details are available at **www.dvrpcfreight.github.io/open-freight-app**.

For those interested in the more advanced *PhillyFreightFinder* project, the full open-source code is available, without documentation, at **www.github.com/DVRPCfreight/phillyfreightfinder**.

Appendix A

Appendix A: *PhillyFreightFinder* Database Schema

River [river]
- ***pff_id**
- name
- townships
- start (mile from ocean)
- end (mile from ocean)
- length
- width
- depth
- clearance
- source

River Channel [river_channel]
- ***pff_id**
- river_id
- up_vessels

Maintained Channel [maintained_channel]
- ***pff_id**
- name
- river_id

Anchorages [anchorage]
- ***pff_id**
- name
- township_s
- start
- end
- length
- river_id [river]

Anchorages (pts) [achorage_pts]
- ***pff_id**
- *anch_id [anchorage]

Good neighbor [good_neighbor_pts]
- ***pff_id**
- description
- municipality

Port Terminal [ports]
- ***pff_id**
- name
- type
- owner
- operator_s
- township
- cargo
- cap_1 [berth qty]
- cap_2 [berth length]
- cap_3 [berth depth]
- cap_4 [cranes]
- cap_5 [warehouse sq ft]
- fc_id [freight center]
- river_id [river]

Port Terminal (pts) [port_pts]
- ***pff_id**
- *port_id [port terminal]

Ship Calls [ship_calls]
- ***pff_id**
- *link_id [port/anch]
- count year
- ship calls
- source

River Crossings (pts) [major_river_crossings]
- ***pff_id**
- name
- type
- owner
- operator
- use
- town_1
- town_2
- cap_1 [tracks/lanes]
- cap_2 [dbl stack/vert clear]
- cap_3 [286k/river clear]
- cap_4 [plate]
- source
- link_id [rail/hwy]

Freight Center [freight_centers]
- ***pff_id**
- name
- center_typ
- township_s
- acres
- rail_id

Freight Center (pts) [freight_centers_pts]
- ***pff_id**
- *fc_id [freight center]

NDC River Links [river_ndc_links]
- ***pff_id**
- ndc_id

NDC Link Tons [ndc_link_tons]
- ***pff_id**
- year
- tot_up
- tot_dwn
- f10up
- f10dwn
- f20up
- f20dwn
- f30up
- f30dwn
- f40up
- f40dwn
- f50up
- f50dwn
- f60up
- f60dwn
- f70up
- f70dwn
- f80up
- f80dwn
- f90up
- f90dwn

Freight Center Employment [fc_employment]
- ***pff_id**
- *link_id [fc_id]
- data year
- type (NAICS)
- employers
- employees

Freight Center Facilities [fc_facilities]
- ***pff_id**
- *link_id [fc_id]
- data year
- type (CoStar groupings)
- space (GLA)

Highways [highway]
- ***pff_id**
- shield
- hwy_num
- state
- name
- type
- owner
- section_miles
- shield_id
- dir_1
- dir_2
- start
- end

Truck Parking [truck_parking]
- ***pff_id**
- name
- owner
- operator
- *cargo*
- township_s
- capacity
- activity
- utilization
- source
- hwy_id [highway]

Freight Rail Lines [rail_lines]
- ***pff_id**
- name
- type
- owner
- operator
- length
- start
- end
- miles
- trk_main
- trk_side
- trk_multi
- trk_dbl
- trk_weight
- trk_plate
- trk_activity

NHS Connectors [nhs_connectors]
- ***pff_id**
- name
- township_s
- cap_1 [lanes]
- aadt
- aadtt
- link_id
- hwy_id

NHS Connectors Pts [nhs_connectors_pts]
- ***pff_id**
- *nhs_id [nhs connector]

Truck Parking (pts) [truck_parking_pts]
- ***pff_id**
- *tpark_id [truck parking]

Count Lookup [count_lookup]
- ***pff_id**
- hwy_id [highway link]
- dvrpc_recnum

Mile Posts (pts) [mile_posts]
- ***pff_id**
- *rail_id [rail line]
- milepost

Control Points (pts) [control_pts]
- ***pff_id**
- *rail_id [rail line]
- control_pt

Rail Yards [rail_yards]
- ***pff_id**
- name
- owner
- operator
- township
- acres
- source
- rail_id [rail line]

Rail Yards (pts) [rail_yards_pts]
- ***pff_id**
- *yard_id [rail yard]

Intermodal Rail [intermodal_rail]
- ***pff_id**
- name
- type
- owner
- operator
- cargo
- acres
- capacity
- source
- fc_id [freight center]
- nhs_id [nhs connector]
- rail_id [rail line]

Intermodal Rail (pts) [intermodal_rail_pts]
- ***pff_id**
- *intermodal_id [intermodal]

PFF Grade Crossings [pff_grade_xings]
- ***pff_id**
- fra_id
- name
- rail_id [rail line]
- owner
- operator_s
- township
- aadt
- tracks
- source

Airports [airports]
- ***pff_id**
- navid (FAA)
- name
- owner
- type
- towered
- township
- cap_1 [runways]
- cap_2 [runway length]
- cap_3 [acres]
- activity [annual ops]
- fc_id

Airports (pts) [airport_pts]
- ***pff_id**
- *air_id [airport]

Heliports (pts) [heliport]
- ***pff_id**
- name
- owner
- municipality
- size

Pipelines [pipelines]
- ***pff_id**
- type
- county

Energy Terminal [energy_terminals]
- ***pff_id**
- name
- operator
- type
- port_link
- ref_capacity
- pipeline_link
- tank_link

Energy Storage [energy_storage]
- ***pff_id**
- port_id
- energy_id
- amount
- unit
- unit_abrv
- descr_

Energy Storage [energy_storage]
- ***pff_id**
- port_id
- energy_id
- amount
- unit
- unit_abrv
- descr_

Product Pipelines [product_pipelines]
- ***pff_id**
- name
- owner
- product
- capacity
- origin
- destination

geographic data

data table

relationship class

Appendix B

Appendix B: NPMRDS Data Process Memo

Overview

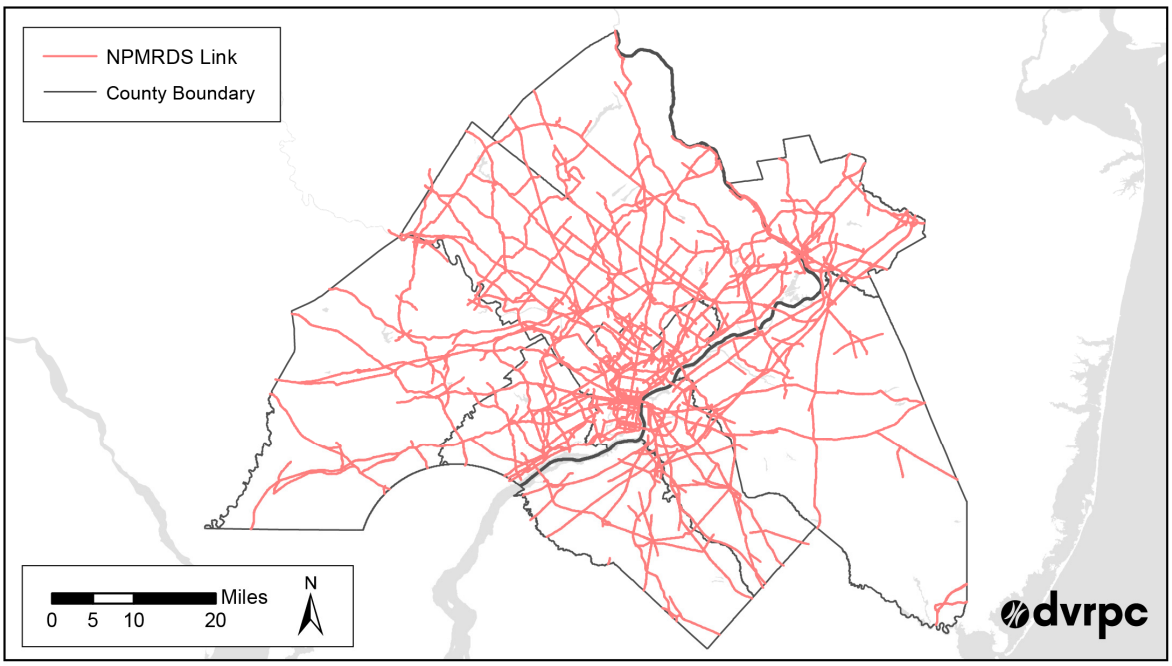
Through the SHRP2 C20 Local Freight Data project, the DVRPC team processed and analyzed the National Performance Management Research Data Set (NPMRDS) to develop truck performance measures. The following summarizes the data processing, analysis, and visualization process conducted by the project team and lessons learned in this effort.

The NPMRDS data is a national dataset of average travel times made available through FHWA to states and MPOs for use in developing performance measures. This data includes the passenger probe data from HERE and the freight/truck probe data from the American Transportation Research Institute (ATRI). The ATRI's truck probe data is based on the GPS devices embedded in the fleet systems to which ATRI has access. The NPMRDS data covers the National Highway System (NHS)—the same base for the regional primary freight network and is updated on a monthly basis. Compared to other travel time data, only observed probe data points are included in the NPMRDS data. If no observations are reported, replacement data is not synthesized. As a result, the sample size is smaller and the data quality may be less accurate than other sources (due to significant outliers and noise). In order to obtain a sample size that was statistically sufficient for deriving performance measures, the NPMRDS data released from 2014 Q1 to 2015 Q2 (one and half of years of data), was downloaded and used for this study.

Data Coverage

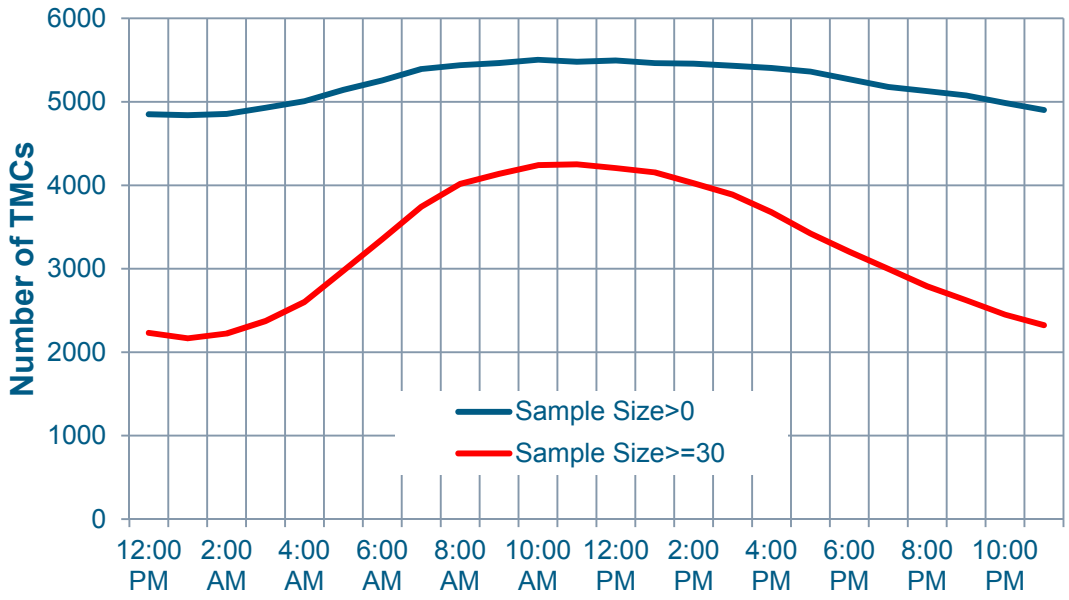
For the DVRPC nine-county region (including Bucks, Chester, Delaware, Montgomery, and Philadelphia in Pennsylvania; and Burlington, Camden, Gloucester, and Mercer in New Jersey), the NPMRDS data includes average truck travel times reported on 5,707 Traffic Message Channel (TMC) segments by direction every five minutes. The total length of the TMC segments is 5,473 miles; the length of individual TMC segments varies from 0.003 to 13 miles with an average length of 0.96 miles. Figure B-1 shows the TMC segments defined in the NPMRDS in the DVRPC region.

Figure B-1: NPMRDS data network coverage for DVRPC region



As expected, the sample size of the travel time observations is varied by TMC and time of day. Although the average sample size seemed to be statistically adequate, the sample size could be very small for certain TMCs and at certain time periods (e.g. early morning and late night). As shown in Figure B-2, the number of "valid" TMC segments was much less if a minimum sample size (e.g. 30) was defined in order for the derived performance measurements to be statically meaningful. The number of "valid" TMC segments that meet minimum sample size requirement drops more significantly during the off-peak periods.

Figure B-2: Distribution of valid TMC segments based on sample size criteria



Data Processing

The NPMRDS data process involved assembling a very large set of downloaded data into a database, extracting the observations in the DVRPC region, analyzing data (e.g. filtering outliers), mapping TMC segments in GIS, calculating performance measures (using Python and R) for designed time periods, and displaying the calculation results in GIS.

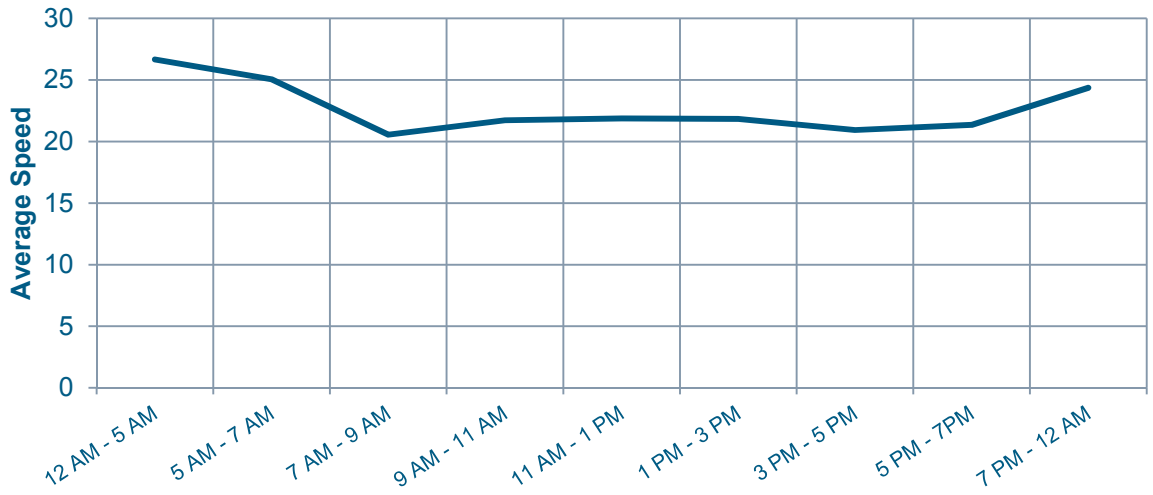
Data Aggregation

In the raw NPMRDS data, the average travel times are reported in each 5-minute bin. The 5-minute averages may be further aggregated to larger time intervals (e.g. by hour of day) if the calculated performance measures (or travel characteristics) are relatively consistent within the aggregation intervals. Using the availability of truck travel time data as a proxy for volume, truck travel characteristics (e.g. average speed by hour of day) show a clear single peak around noon – contrary to passenger vehicles and 'normal' commute behavior. A similar temporal variation is presented by other performance measures such as delay and on various TMC types (e.g., urban vs. rural). Therefore, larger bins (analysis intervals) may be sufficient to profile the variation of truck performance by time of day.

In addition, aggregating data into larger bins is desired to improve the yield rate of data points, given the requirement of a minimum sample size.

Data was aggregated into larger bins to improve yield rate of data points. A minimum sample size of 30 travel time readings per time bin was selected; this requirement equates to a margin of error of 15% and a confidence level of 90%. Hours of the day were grouped into time bins by considering the hourly speed and matching alike hours. The final time bins are five-hour long bins at the start and end of the day, and two-hour long bins throughout the day – nine time bins in total, as shown in the figure below.

Figure B-3: Average speed by designated time bins



Outlier Handing

Given the nature of its data source, the NPMRDS data exhibits significant outliers and noise. The outliers had to be appropriately identified and handled to obtain meaningful aggregated performance measures. Initial data cleaning was performed to remove any illogical data points. Data was considered invalid if the speed was greater than 100 miles per hour or if the travel time on a single link was greater than 2 hours.

In an attempt to further remove outliers, only data that resided within the 'whiskers' of a Tukey boxplot were considered valid. This is calculated by determining the first and third quartile. Data was considered valid if it fell between (Q1 – 1.5*IQR) and (Q3 + 1.5*IQR), where the IQR was the inter-quartile range (Q3 – Q1).

Other methods of a data cleaning were investigated: a flat speed cut-off, percentile based filtering, and classified filtering based on TMC attributes. Both the minimum speed specification and percentile filtering methods were unsatisfactory since the distribution characteristics were not taken into account. Applying filters based on TMC attributes is relatively more effective at data cleaning, but still produced unsatisfactory results.

After the data cleaning, the average TTIs and PTIs were more reasonable (i.e. decreased), while still being plausible given the road classification, type, etc., as shown in Figure B-4 and Figure B-5.

Figure B-4: Comparison of raw versus cleaned data average TTI

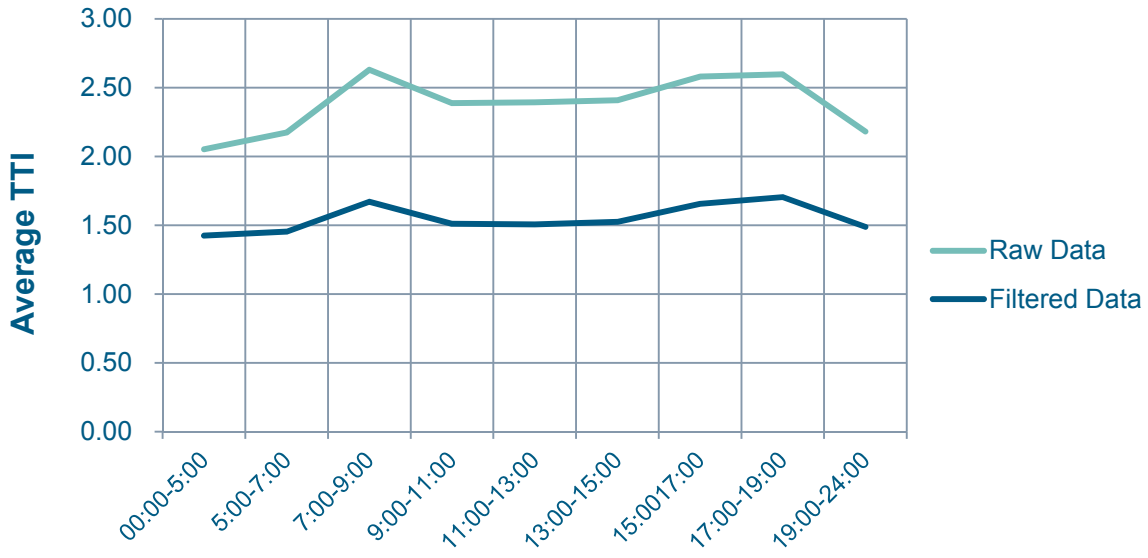
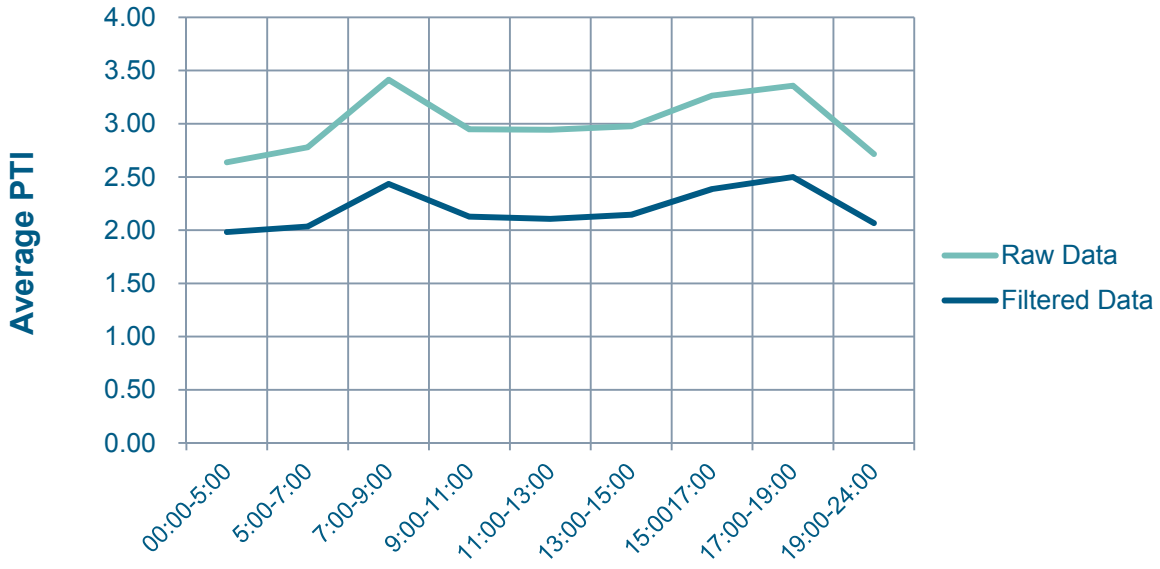


Figure B-5: Comparison of raw versus cleaned data average PTI



TMC Mapping and Visualization

NPMRDS data for individual TMC segments was spatially visualized by joining the processed travel time data to the road network in the DVRPC region. While initial attempts were made to match TMC segments to the road network used in DVRPC's regional travel demand model, the effort to accurately match TMC segments to the OpenStreetMap-based network was too time consuming for this study. Instead, the HERE road network shapefile was used, which already includes a lookup table to match TMC segments with individual links in the road network.

Some pre-processing was performed on the HERE network in order to facilitate improved data matching and visualization. TMC data is joined to individual links using a combination of an identifier unique to each link (LINK_ID field in the HERE network) and the travel direction (DIR_TRAVEL)—F, T, or B. To reduce the size of the HERE network dataset, bidirectional links (DIR_TRAVEL "B") are used to represent both travel directions on a single roadway. However, TMC data is defined in a single travel direction, "F" ("From-To") or "T" ("To-From"). To simplify the process of joining TMC data to the road network, bidirectional links were duplicated, retaining the original link in the "F" direction and creating a duplicate link to represent the "T" direction. The digitized direction of all "T" links were then reversed using ArcGIS to simplify the parameters required to clearly display both travel directions of bidirectional roads.

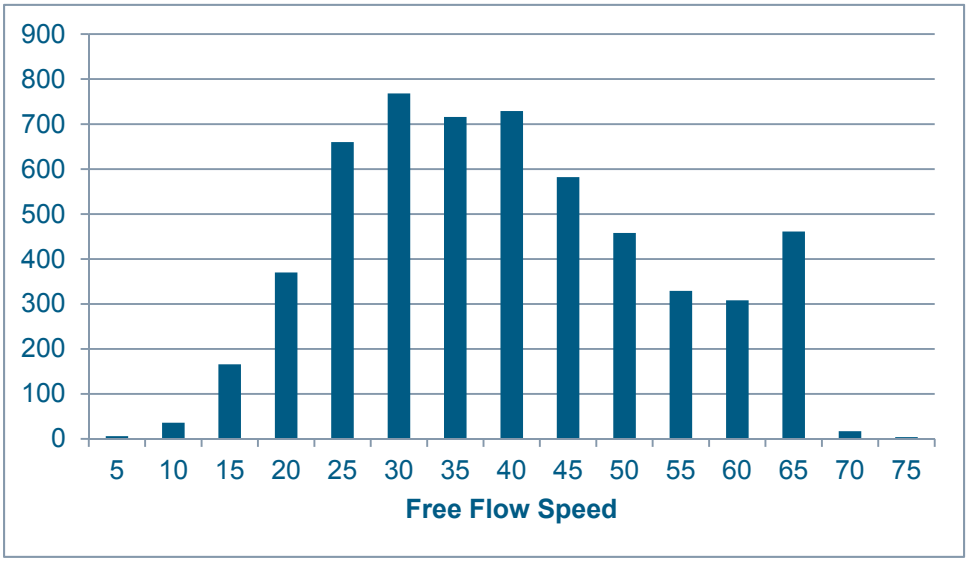
Performance Measures

After filtering outliers and aggregating the observations into the defined time bins, a few truck performance measures were calculated for each TMC and time bin. Average truck speeds and travel time delay measures were calculated directly using the NPMRDS data. Additional performance metrics can be calculated by combing the travel time data and truck counts stored in the *PhillyFreightFinder* and DVRPC databases.

Truck Free-Flow Speed

To understand truck delay, the first step was to determine the free-flow speed on each TMC so that free-flow travel time could be calculated and compared to the actual travel time by different times of day and year. The free flow speed was calculated by using the TMC segment length and the 15th percentile travel time for each TMC. The figure below is the histogram of free flow speeds for all TMCs in the DVRPC region.

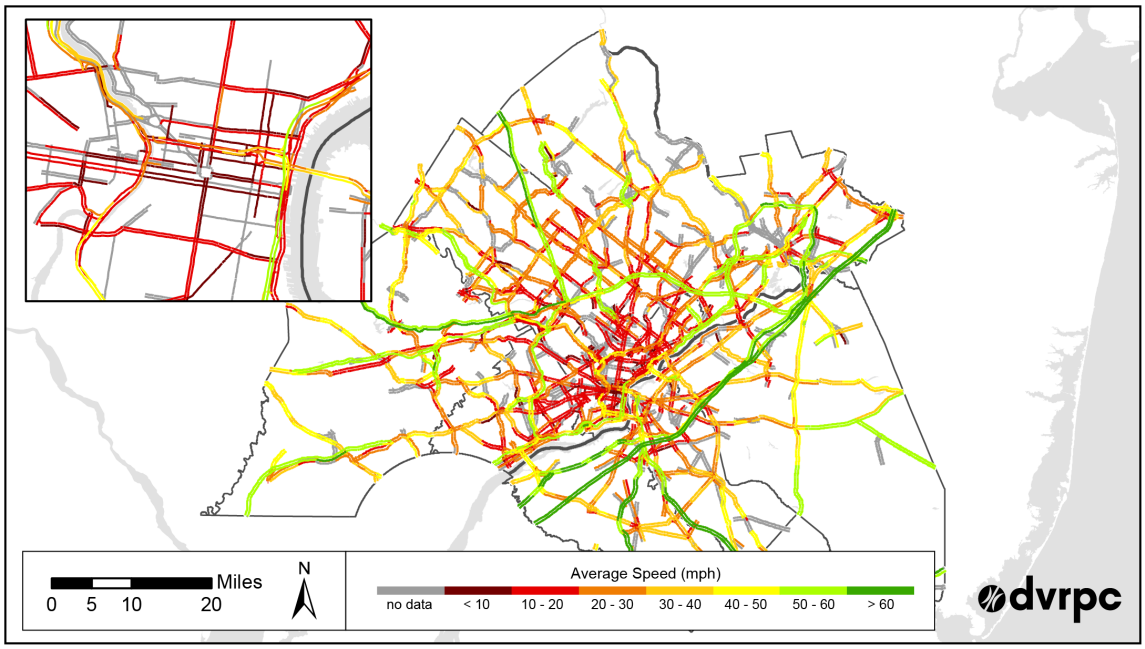
Figure B-6: Free speeds for all TMCs



Average Truck Travel Speed

For each time bin, the average travel speed was calculated by using the average travel time of all data samples within the bin and the TMC segment length.

Figure B-7: Average Truck Speed for 5PM to 7PM time bin



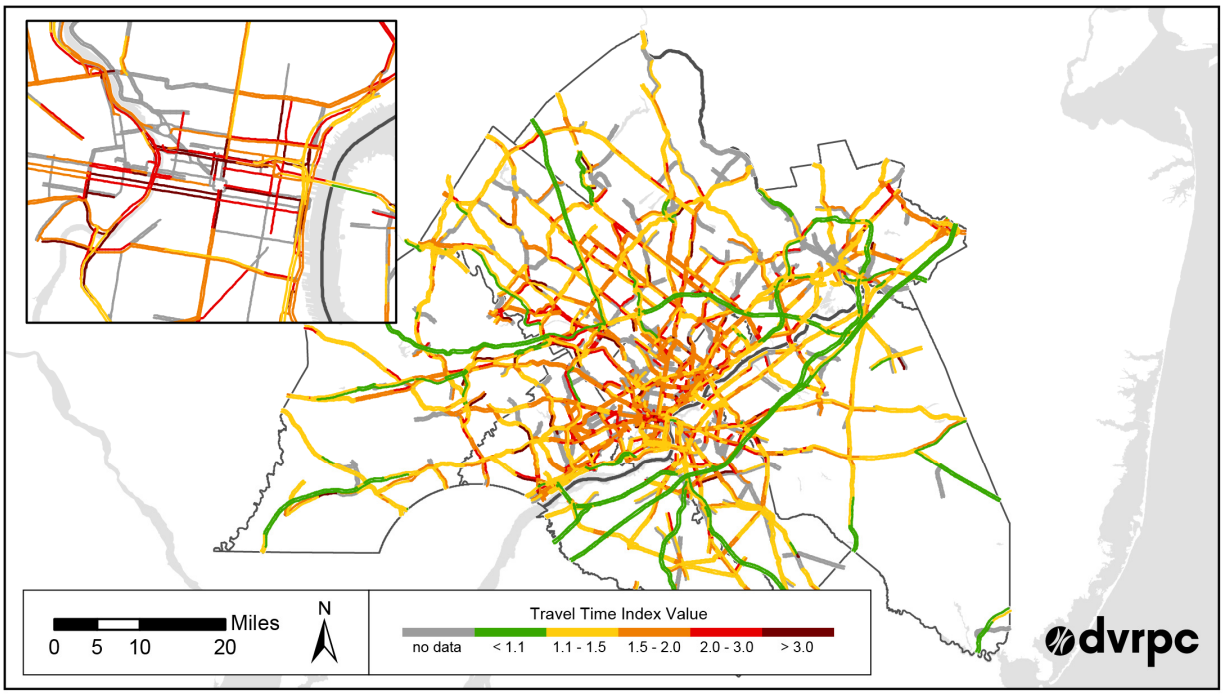
Travel Time Reliability

A few performance measures that describe car travel time reliability[1] were calculated for trucks:

- Travel Time Index (TTI): travel time represented as a percentage of the ideal travel time (= Average Travel Time / Free-Flow Travel Time).
- Planning Time Index (PTI): the total travel time that should be planned when an adequate buffer time is included (= 95% Travel Time / Free-flow Travel Time). The planning time index differs from the buffer index in that it includes typical delay as well as unexpected delay. Thus, the planning time index compares near-worst case travel time to a travel time in light or free-flow traffic. For example, a planning time index of 1.60 means that, for a 15-minute trip in light traffic, the total time that should be planned for the trip is 24 minutes (15 minutes × 1.60 = 24 minutes).

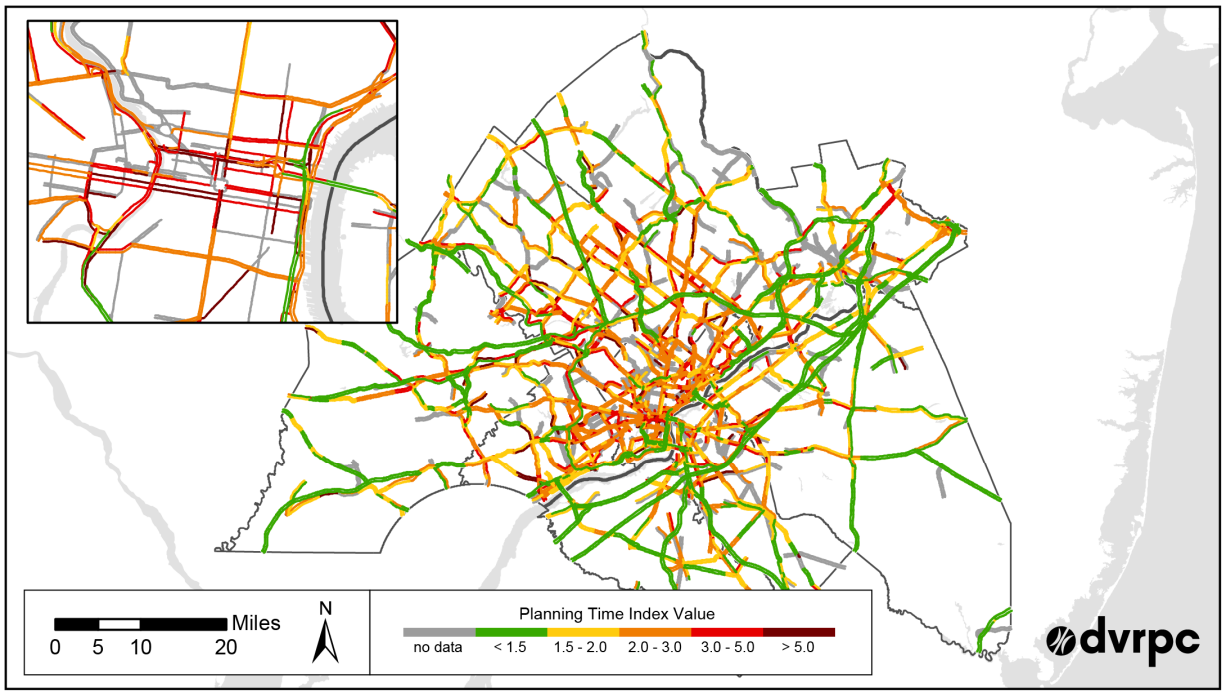
For each time interval, the TTI and PTI were calculated for each TMC with a minimum sample size of 30 within each time bin, as shown in Figure B-8 and Figure B-9. Comparatively, TTI may be more reliable to represent travel delay and travel time reliability under the constraint of sample size and large number of outliers in the raw data. PTI may be varied more significantly as the sample size changes.

Figure B-8: Travel Time Index for 5PM to 7PM time bin



[1] Federal Highway Administration, *Travel Time Reliability: Making It There On Time, All The Time,* 2013

Figure B-9: Planning Time Index for 5PM to 7PM time bin



Truck Delay

Total truck delay can be calculated based on truck travel time delay and truck volumes at selected TMC segments, designated routes or subareas. Truck volumes in terms of truck AADT were estimated by geo-referencing the class count locations to the TMC segment locations.

Conclusions

In this project, the NPMRDS data was used for calculating truck performance measures. Overall, the results appeared to be plausible in revealing the temporal and spatial variations and patterns of highway performance in support of truck movements, as compared to other probe data sources such as Google and INRIX. Further investigation of the data fidelity, outlier handling methods, and appropriate truck performance measures is necessary.

Making Freight Data More Accessible

SHRP2 C20: Local Freight Data Improvement

Publication Number: 15061

Date Published: April 2016

Geographic Area Covered:

The nine-county DVRPC planning area, which covers the counties of Bucks, Chester, Delaware, Montgomery, and Philadelphia in Pennsylvania; and Burlington, Camden, Gloucester, and Mercer in New Jersey.

Key Words:

Freight planning, Freight, Data, *PhillyFreightFinder,* SHRP2 C20, Performance Measures, County Data

Abstract:

With support of the Strategic Highway Research Program (SHRP2) C20 Local Freight Data Improvement program, the Delaware Valley Regional Planning Commission (DVRPC) established and advanced a comprehensive freight data improvement program for the DVRPC Region. This project sought to improve access and usability of a variety of data sets for regional freight planning through innovative analysis and visualization. The development of a collection plan, internal databases, and improvements to the *PhillyFreightFinder* web application were integral components of this comprehensive freight data strategy. A series of internal tools as well as an open-source, replicable template based on *PhillyFreightFinder* were developed and made available for use by other agencies seeking to create similar products in other regions.

Staff Contact:

Fang Yuan
Principal Transportation Engineer
(215) 238-2885
fyuan@dvrpc.org

Michael Ruane
Transportation Planner
(215) 238-2888
mruane@dvrpc.org

DELAWARE VALLEY dvrpc REGIONAL PLANNING COMMISSION

190 N. Independence Mall West, 8th Floor
Philadelphia, PA 19106-1520
Phone: (215) 592-1800
Fax: (215) 592-9125
www.dvrpc.org

DVRPC, 8TH FLOOR
190 N. INDEPENDENCE MALL WEST
PHILADELPHIA, PA 19106-1520
(215) 592-1800
WWW.DVRPC.ORG