

Map Matching and Real World Integrated Sensor Data Warehousing



www.nrel.gov/tsdc

www.nrel.gov/fleet_dna

Evan Burton – Data Engineer (Presenter)
Jeff Gonder – Vehicle System Analysis Team Lead
Adam Duran – Engineer/Analyst
Eric Wood – Engineer/Analyst

2013 Federal Committee on Statistical Methodology
(FCSM) Research Conference
November 4-6, 2013
Washington, DC

NREL/PR-5400-60893

NREL—Real World Vehicle Drive Cycle Database

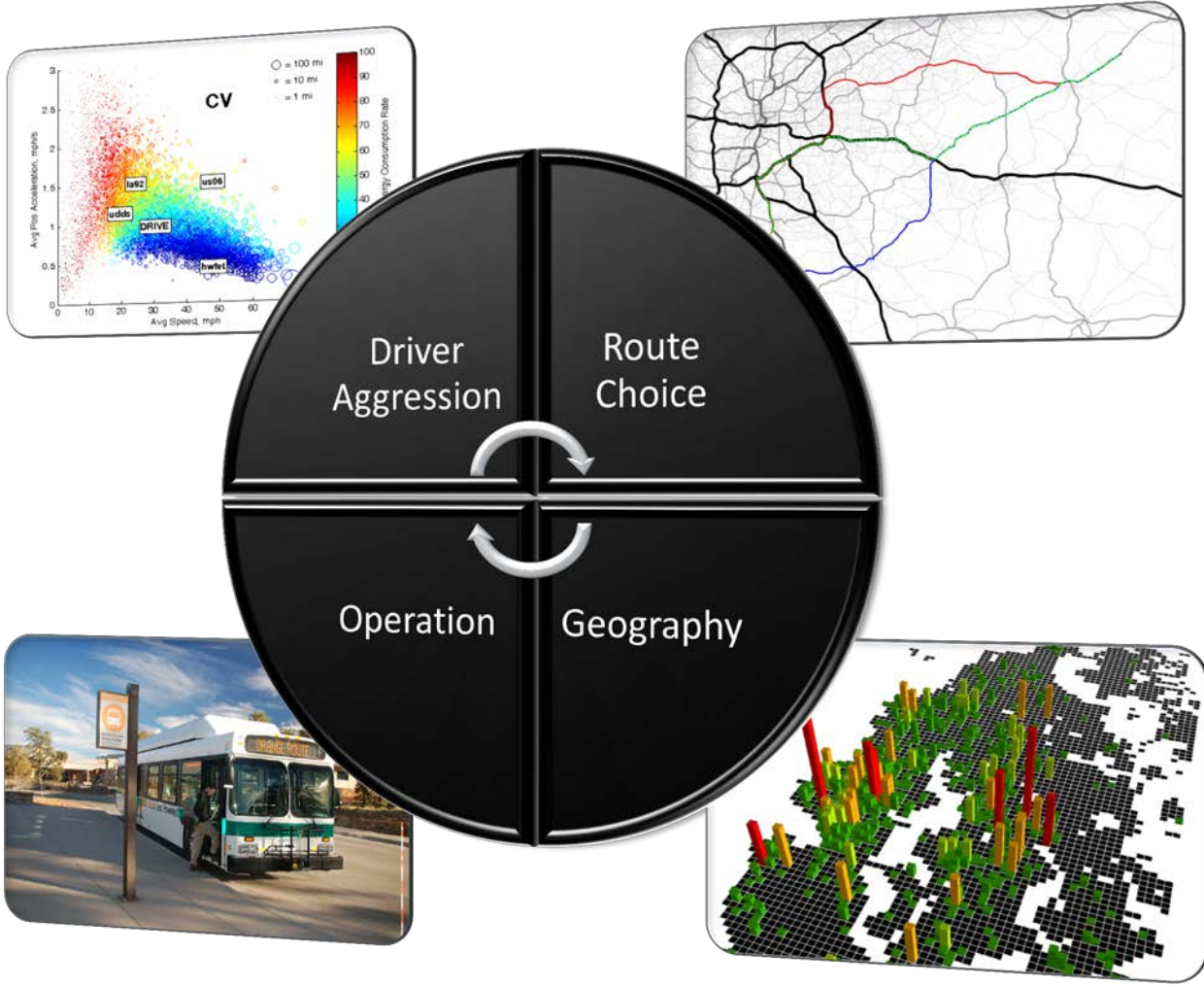
Quantifies otherwise overlooked variables in vehicle analysis.

Map matching sensor data appends additional variables at second-by-second increments :

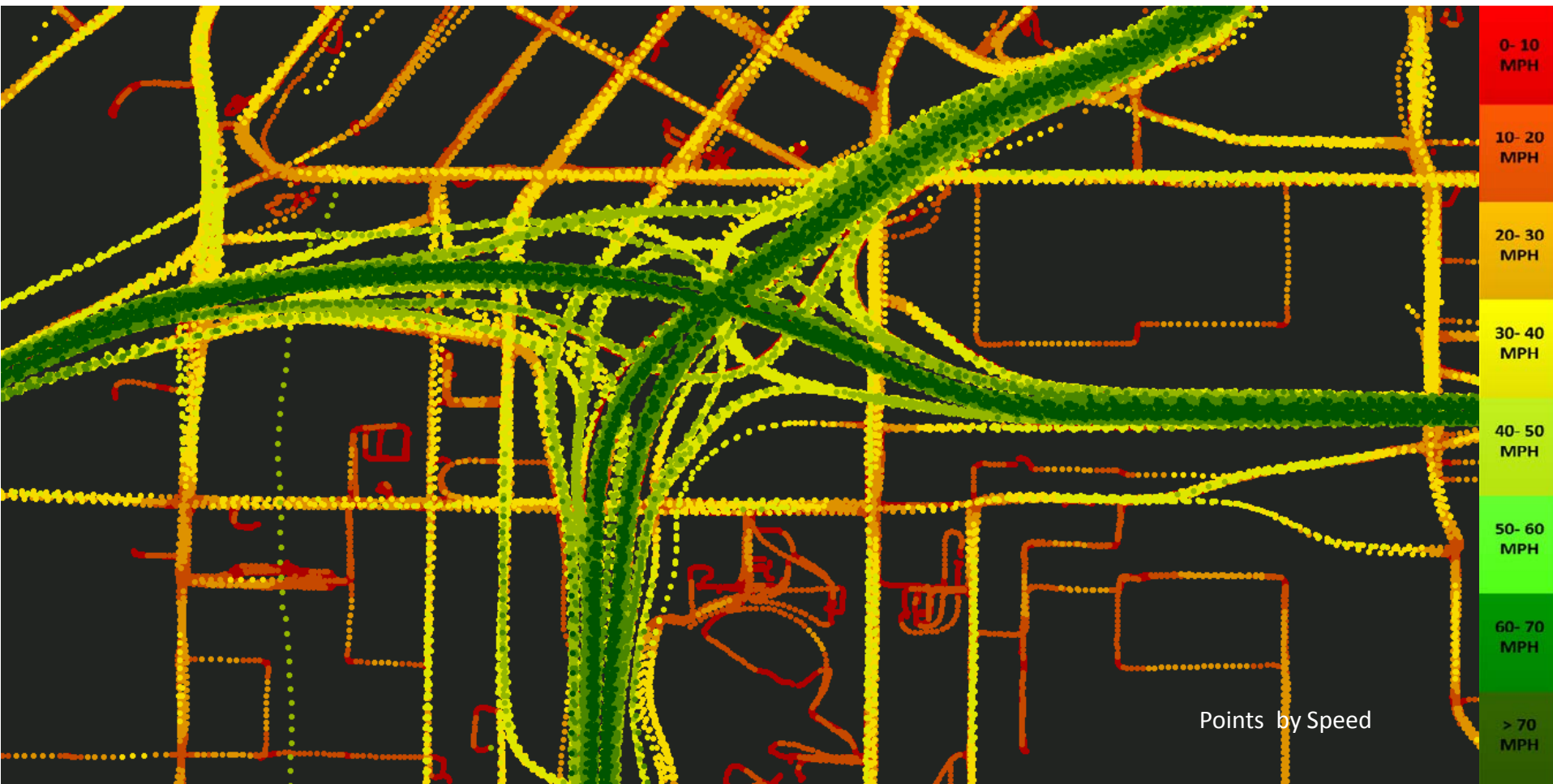
- Allows for scalability of analysis
- Incorporates route choice and geography into results

Uses a significant sample group of both personal and commercial vehicles to provide a resource for engineers and researchers focusing on:

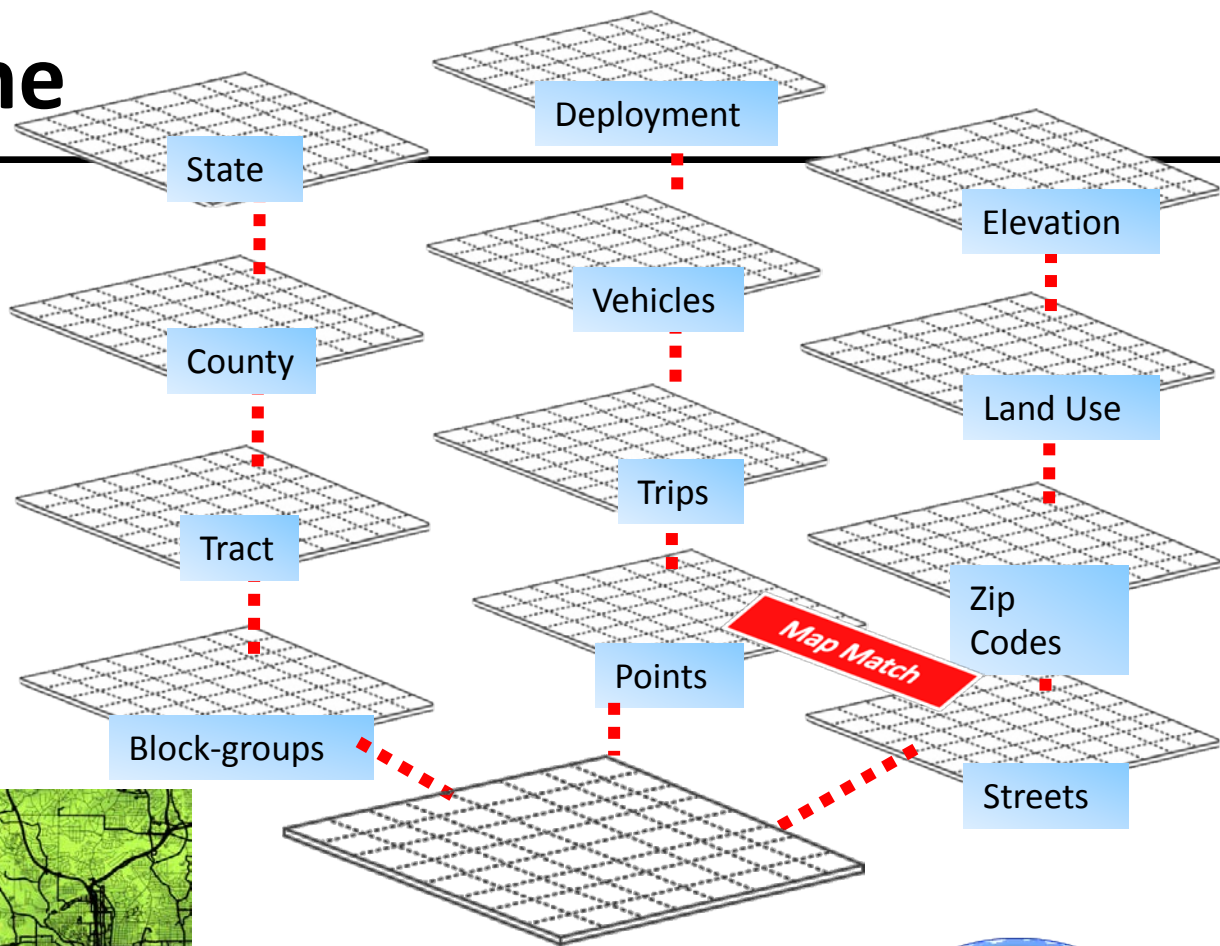
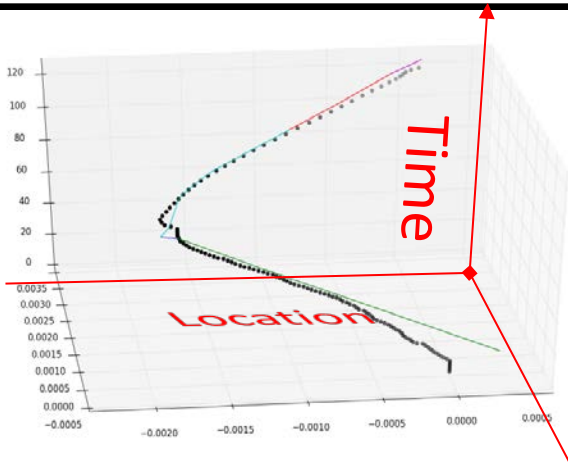
- Regional vehicle emissions modeling
- Alternative vehicle application identification
- Vehicle-to-grid electrification
- Regional transportation demand modeling
- Green routing
- Intelligent route-based control strategies for hybrid electric vehicles
- Long-term battery degradation analysis



The Data – Location/Speed/Time



Location and Time



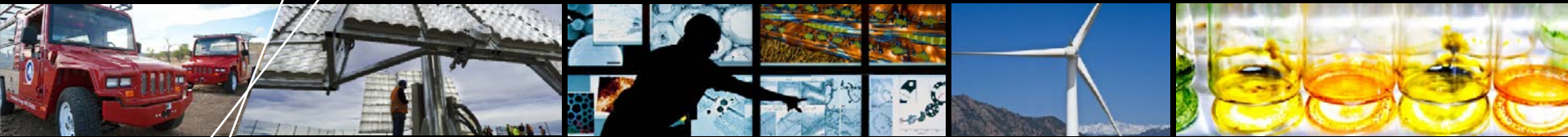
Spatial Reference Data:

- 2010 Census
- 2011 ACS
- USGS 1/3 DEM
- NAVTEQ Streets



- All data sets are linked using an XY coordinate system (Earth)
- Links established by passing primary keys through records
 - Trip 1 ended at tract 3
- Time provides the final link for moving objects
 - Vehicle 1 enters road 1 at time A and exits at time B





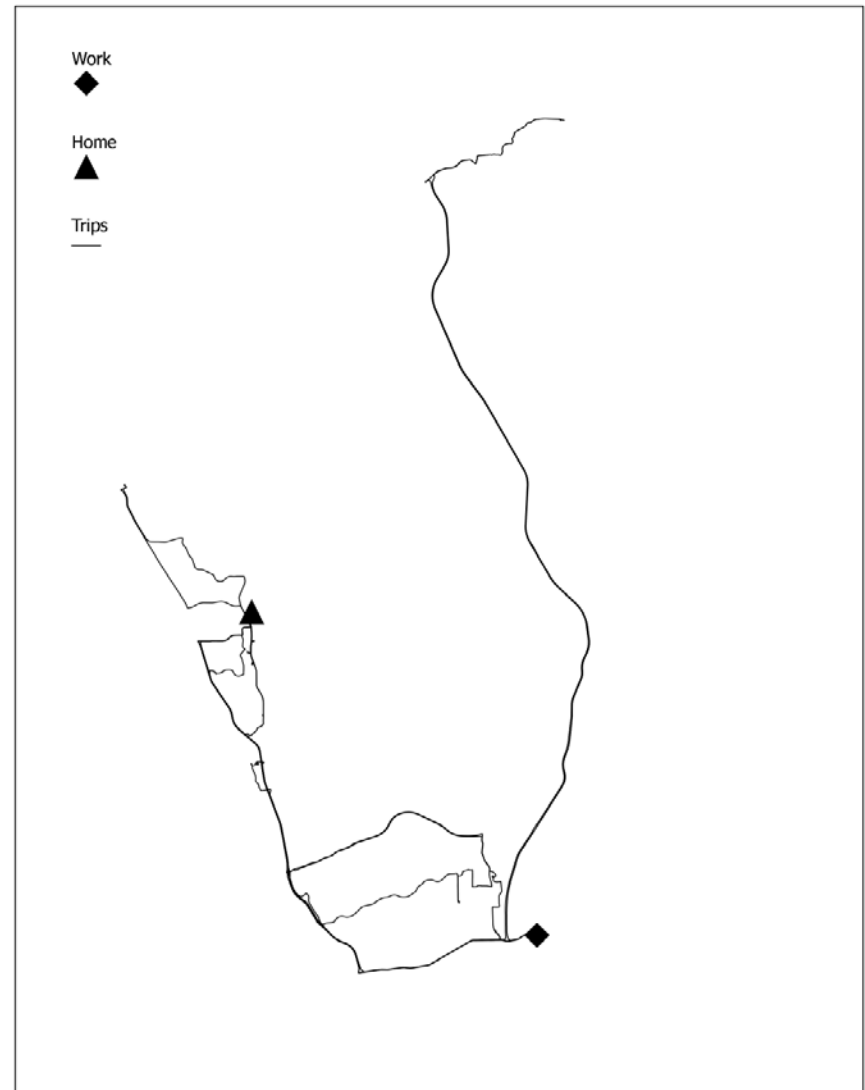
Map Matching

Vehicle Processing – End Product

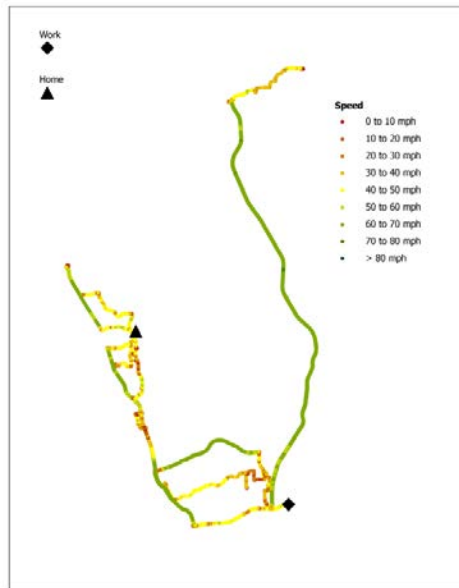
- **Second-by-second data are summarized at the trip level**
 - Trips are also categorized by end and start type (home/work/school)
- **Vehicle processing of the second-by-second data includes:**
 - Filtering the recorded data
 - Appending data from additional data layers to second-by-second data
- **Summary statistics are generated at the vehicle, day, and trip levels**
- **Trip classifiers are grouped to vehicle and day levels**

Vehicle Processing (Next Slide)

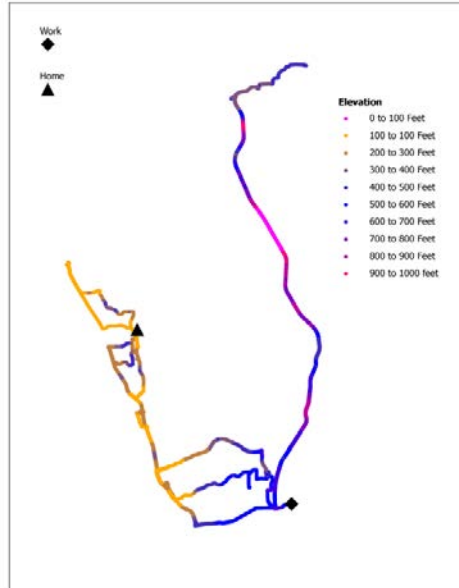
1. Speed – Filtered
2. Location – Filtered
3. Elevation – Appended
4. Grade – Appended
5. Street ID – Appended
6. Street Functional Class – Appended
7. Street Speed Category – Appended



Vehicle Processing – End Product



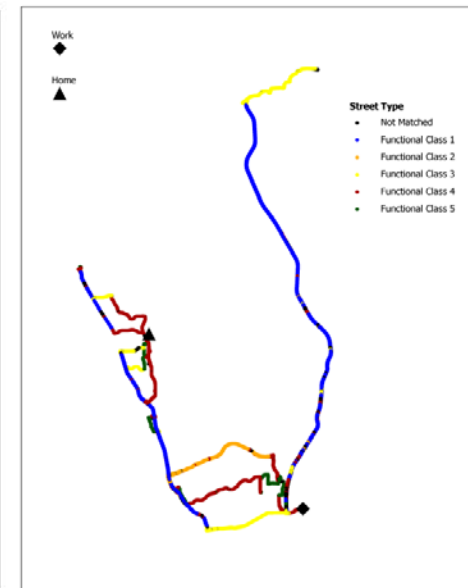
Speed – Filtered
Location – Filtered



Elevation – Appended



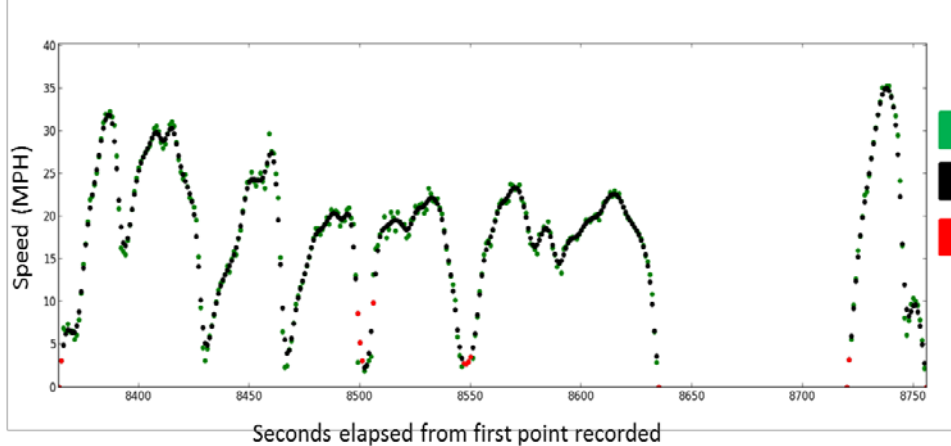
Grade – Appended



Street ID – Appended
Street Functional Class
– Appended
Street Speed Category
– Appended

Speed & Location Filters

1

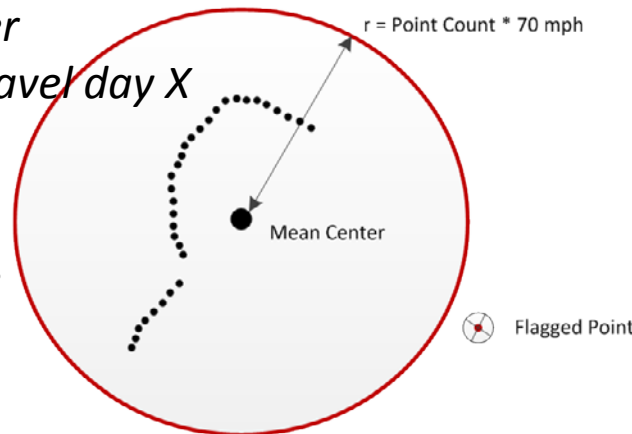


Speed Filtration

- *Flags unrealistic speed*
- *Backfills missing points*
- *Interpolates missing and flagged data*

Severe outlier filter

- *Removes points falling outside a data defined buffer*
- *Radius = Hours of travel day X 70 mph*



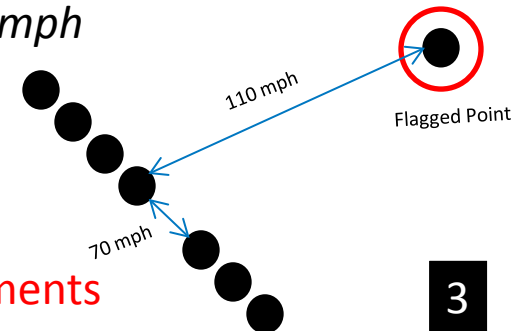
2

Mean Center

- *The mean latitude and longitude independently*

GPS Jumps Filter

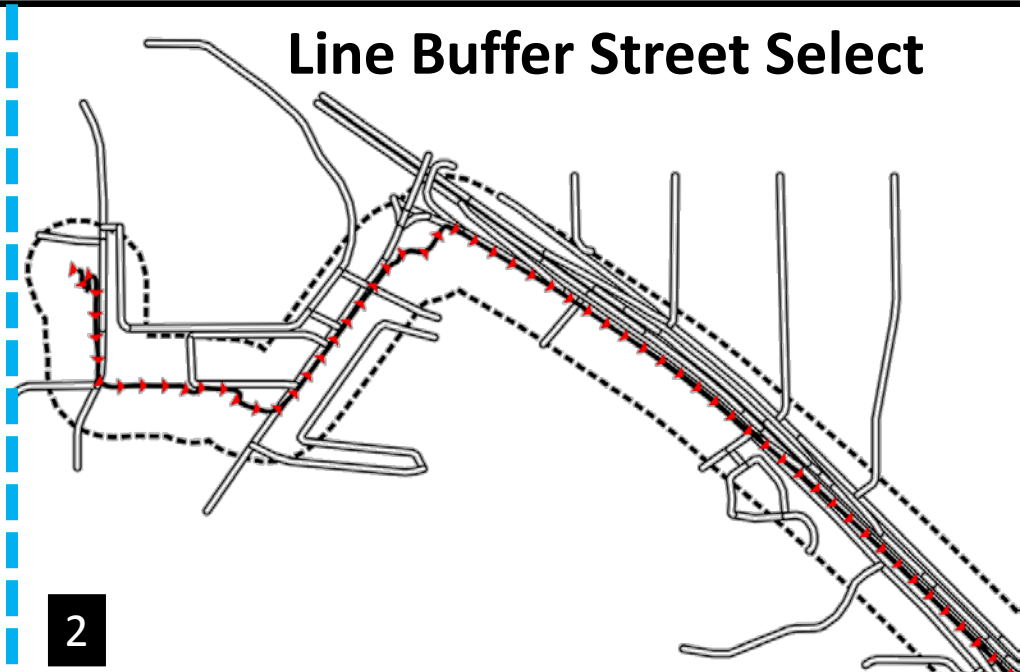
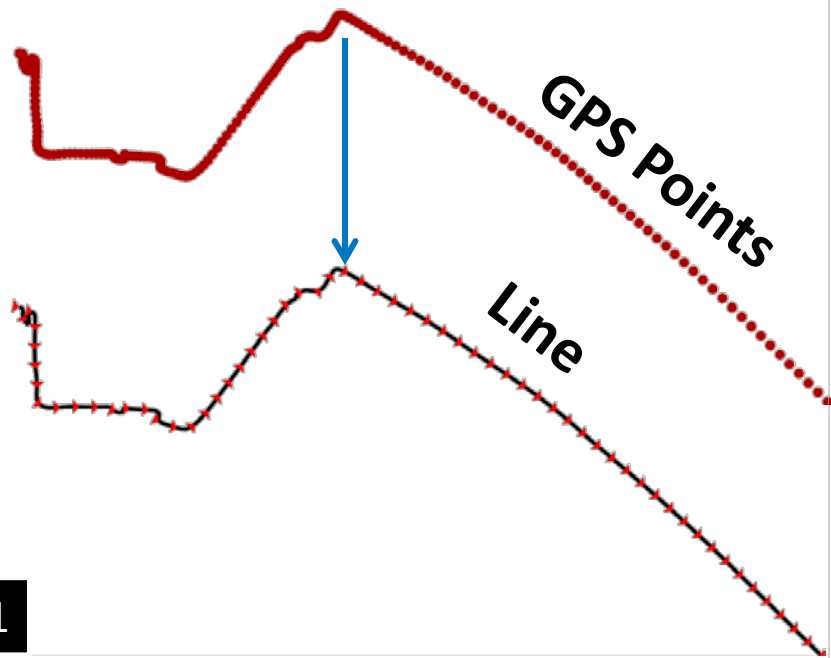
- *Calculates a rough speed value from distance and time*
- *Points where the calculated speed is > 100 mph are flagged until calculated speed falls under 100 mph*



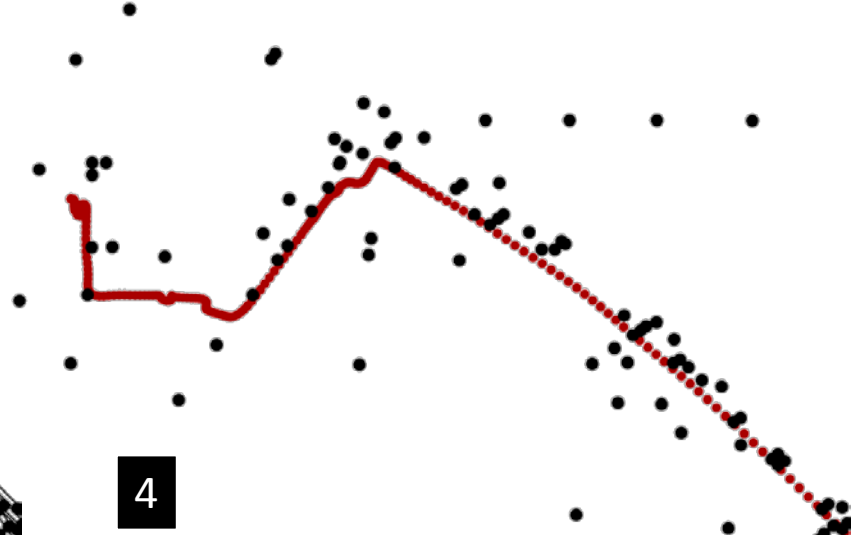
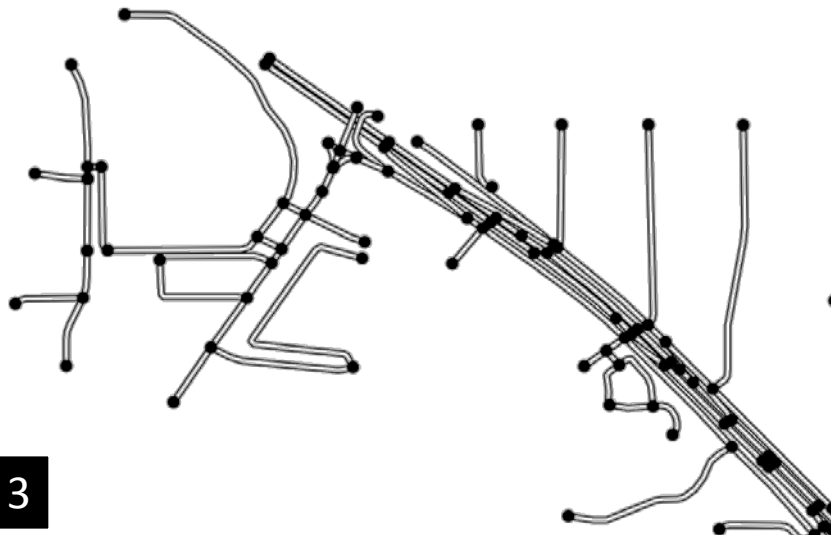
3

Typically less than 2% of GPS points require these processing adjustments

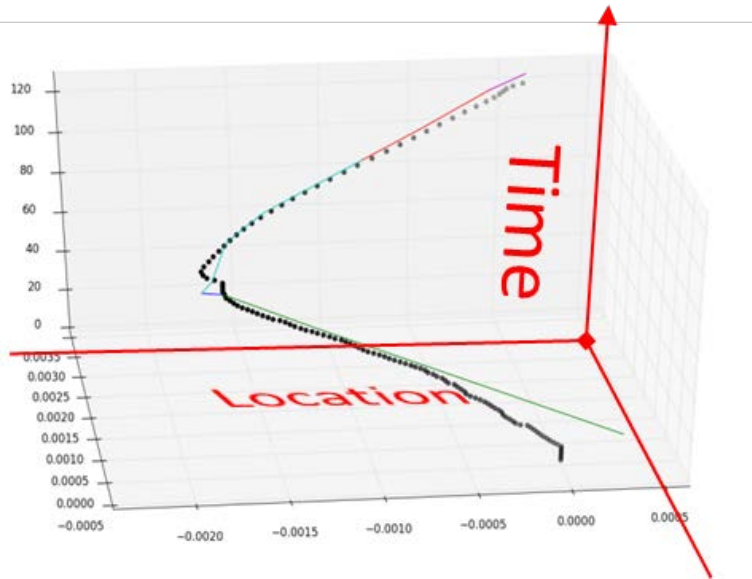
Map Matching: Dynamic Node Selection



Unique Node
Options



Map Matching: Solve



Point-to-node links are established using time

- *Each point is linked to the nearest node within 500 feet of a road*
- *Link is established by appending the node identifier to the point data (time)*
- *Points are down sampled to a discrete time at which the vehicle passed the node*
- *Vehicle x passed this node at this second*

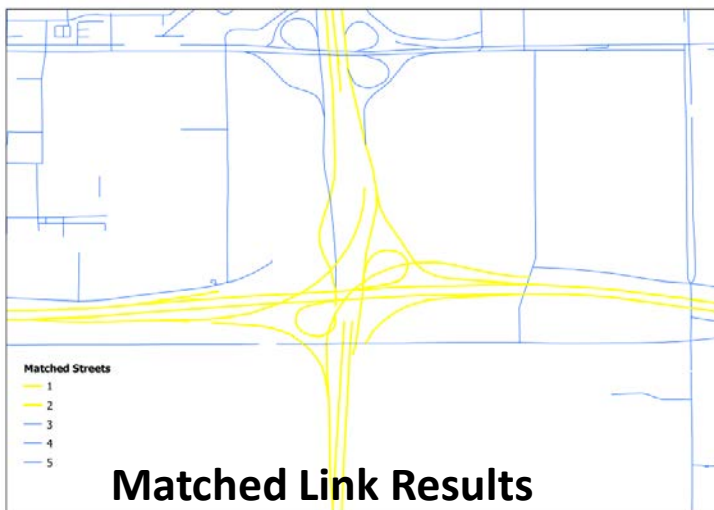
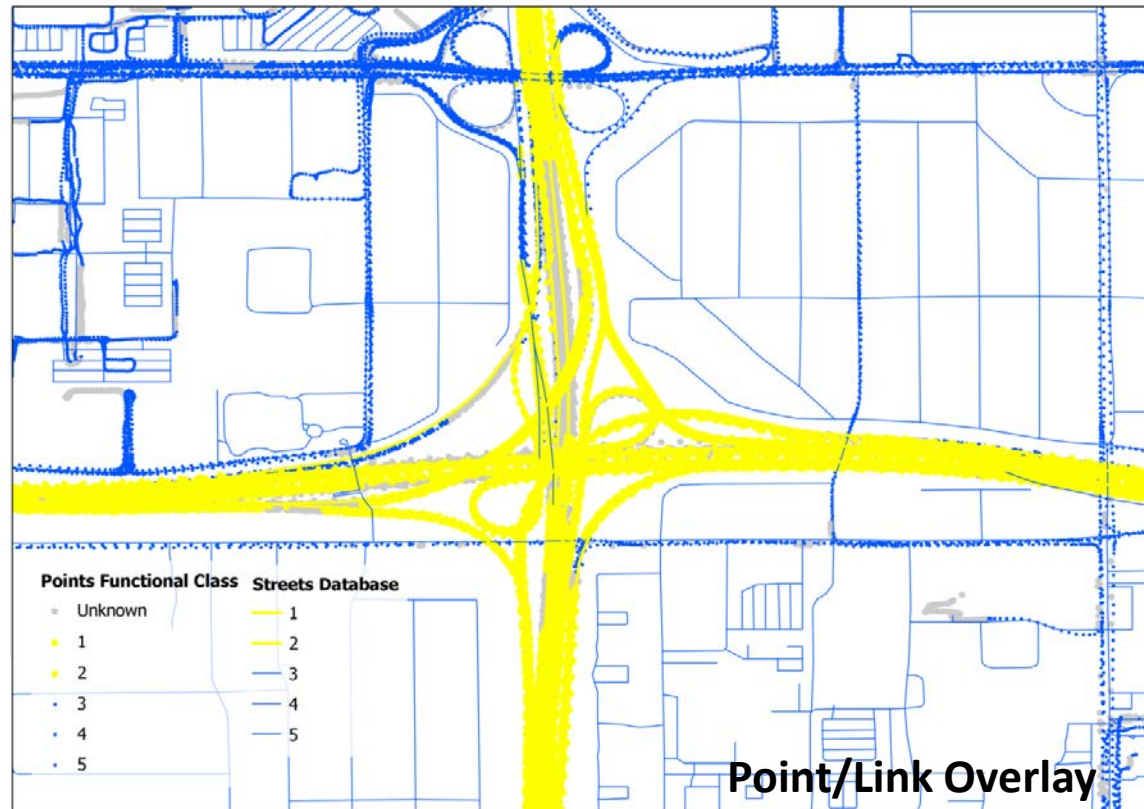
Solving connections logically (topology)

- *The segments connected to the node and the possible exit node options for each entry node are identified independent of previous or next known results*
- *Options are then tested logically in order using multi-step decision tree*
 - *Identifies and removes latent links*
 - *Fill gaps in connection of consecutive segment links*
- *Appends attributes from the linked street segment to the data points falling within the entry and exit time of the matched segment*

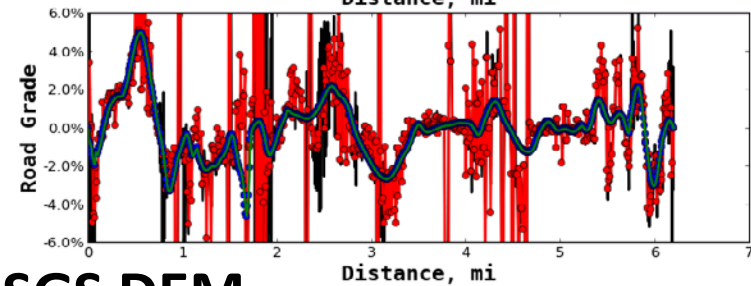
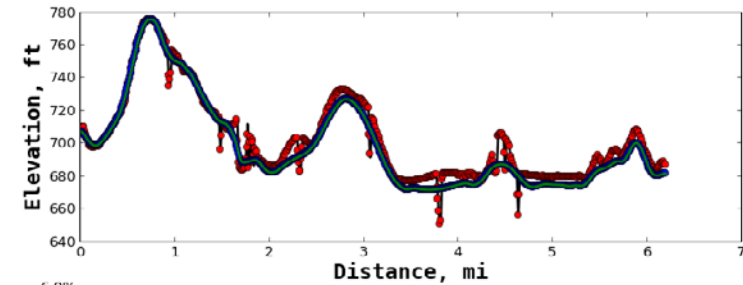
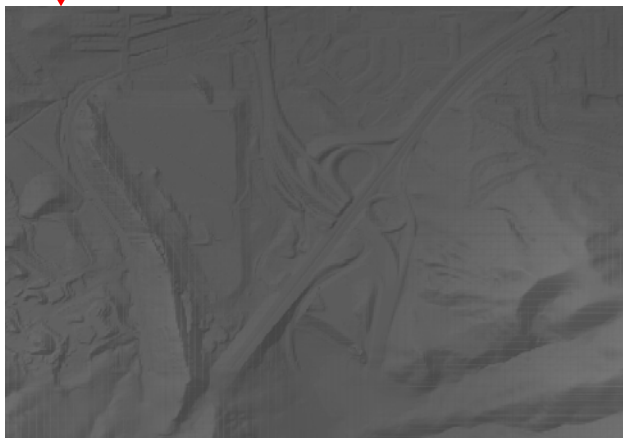
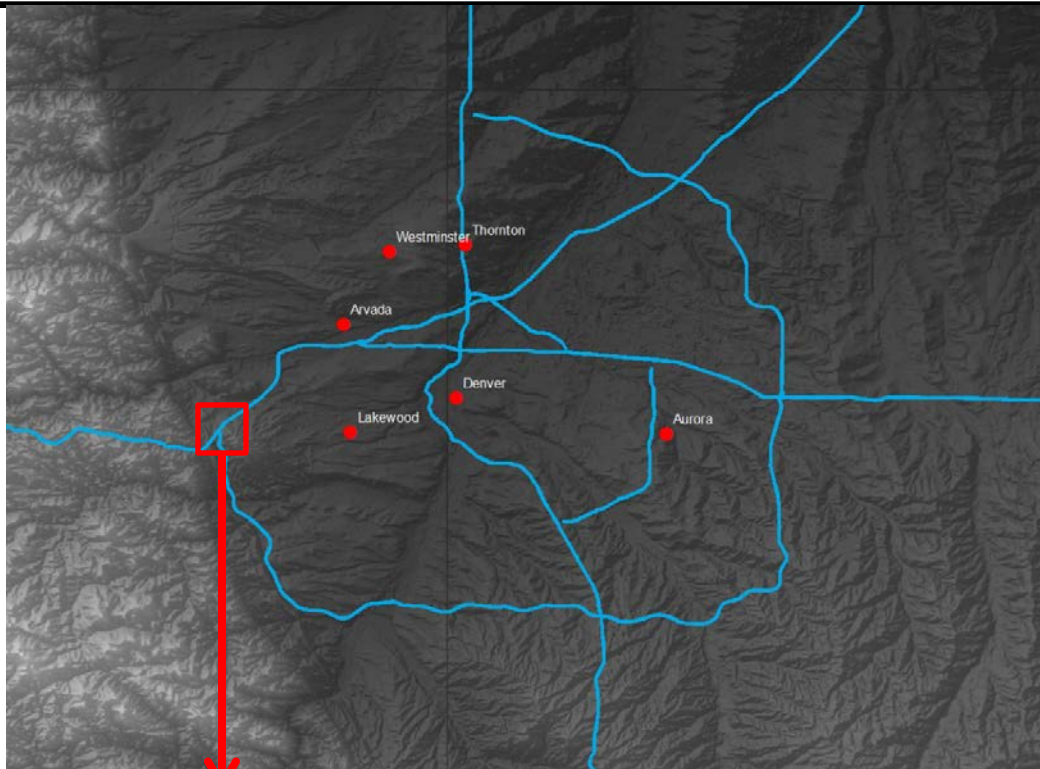
Map Matching: Accuracy

Complex overpasses

- *Connectivity can become ambiguous when so many options are available*
- *95% of distance matched across all data sets*
- *Cleaned up post processing during road based analysis*

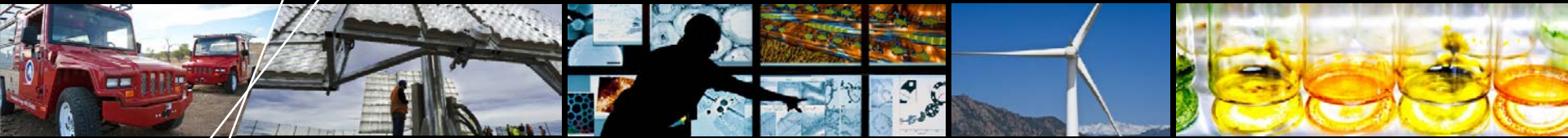


Elevation/Grade



USGS DEM

- *The elevation for each location recorded is looked up from the USGS Digital Elevation Model (DEM)*
- *The elevation profile returned is smoothed*
- *Down sampled to normalize for distance to calculate grade*
- *The processing accuracy has been validated against roads with available ADAS grade measurements*

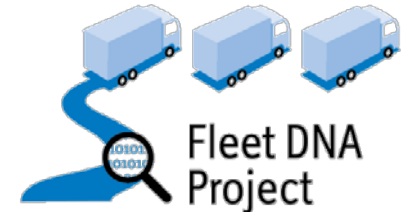


Integrated Sensor Data Warehousing

NREL - Real World Vehicle Data Warehousing

- **Fleet DNA – DOE**

- Warehouses medium- and heavy-duty fleet vehicle operational data across vocations, drivetrains, fuel type, and vehicle types
 - Has the ability to crowd source data collection, allowing for large-scale comparative analysis
- Removes identifying information and provides data to public for vehicle performance comparisons



- **Transportation Secure Data Center (TSDC) – DOT and DOE**

- Securely archives and provides public access to detailed regional transportation survey data
- Speed/time/location data for personal vehicle samples

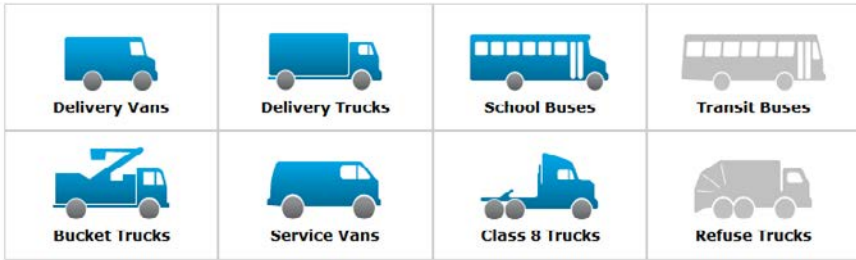
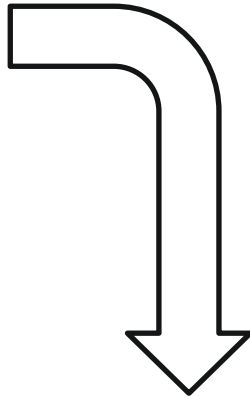
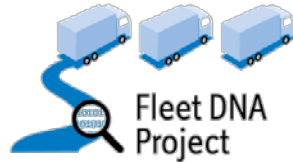
- **Reporting is specific to project using the same source data, but different ways of looking at it**

- TSDC – Trip starts and ends are classified relative to the home, work, school location
- Fleet DNA – Trip starts and ends are classified relative to the depot location

Geography: Sample Groupings

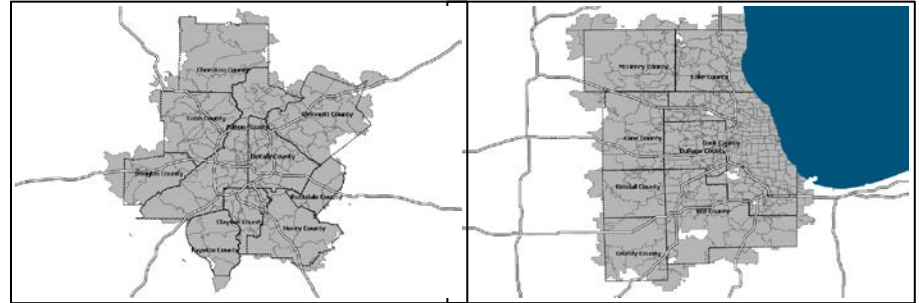
Fleet DNA categorizations (vehicle specs):

- *Make/Model/Year*
- Class
- Vocation
- Type
- Drivetrain
- Fuel Type
- City
- State



TSDC categorizations (geographic)

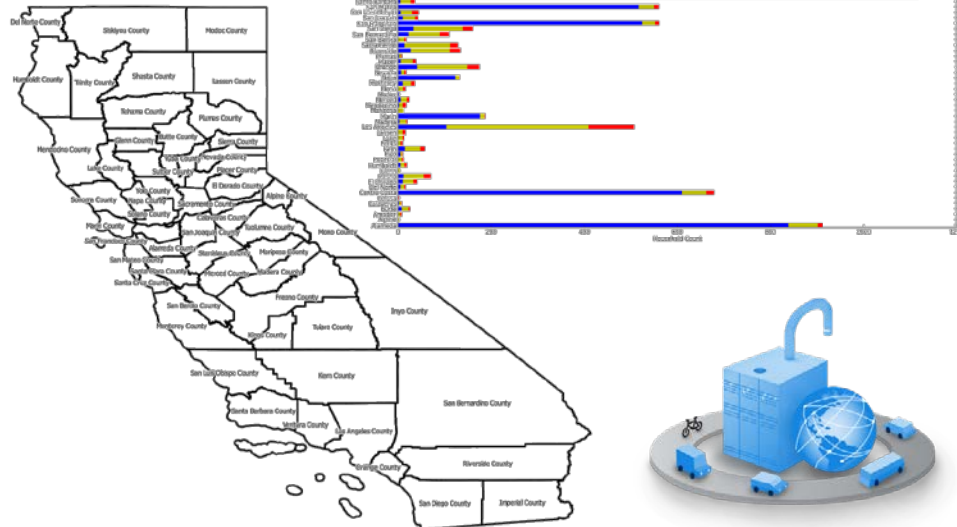
- Regions for data collection
- *Make/Model/Year*



Region of Deployment (Atlanta)

Region of Deployment (Chicago)

Deployment (Caltrans)



Privacy & Data Security



- Both TSDC and Fleet DNA data sets contain extreme detail regarding the travel of the vehicles
 - Vehicles equipped with GPS sensors require certain precautions to be taken to ensure privacy and data security
- Each public distribution portal has specific controls in place

Fleet DNA

- Reports and downloads are grouped by vocation
- Allows comparative analysis between deployments
- Deployments can be identified within the data through an identifier, but the lookup is not made available
- NREL processing is applied to the vehicle GPS samples, and 360+vehicle travel statistics are provided by vocation



TSDC – Public

- Personally Identifying information is removed from source data
- Spatial reference identifying any geographic identification below the census tract is removed
- Vehicle model is removed
- Original study data are provided for download
- NREL processing is applied to the vehicle GPS samples, and 360+vehicle travel statistics are provided by vehicle



Transportation
Secure Data Center
(TSDC)

TSDC – Controlled

- Following a approval process, researchers gain access to the full TSDC data sets through a virtual desktop (VM – Virtual Machine)
- An NREL administrator acts as the gatekeeper for all data in and out of the environment
- Users can work within NREL's databases within the VM using software provided
- Data may only be removed by the NREL administrator

Data Center Infrastructure

Data Center Servers



TSDC Website

Transportation Secure Data Center

Cleansed Data Access
Users may quickly access data from the following studies to support applications that require detailed travel distance and speed information, but that do not require detailed vehicle and longitude spatial information. Before viewing this data for the first time, you will be required to fill out a short registration form.

- California Department of Transportation - California Household Transportation Survey
- Chicago Metropolitan Agency for Planning - Regional Household Travel Survey
- Florida Regional Commission - Regional Travel Survey with GPS Data
- Texas Department of Transportation - Transportation Studies with GPS Travel Data with data for Abilene, Austin, El Paso, Houston/Galveston, Laredo, Rio Grande Valley, San Antonio, Tyler/Longview, and Wichita Falls
- West Regional Council - Traffic Choices Study

Anonymous fleet data

Fleet DNA Website

Fleet DNA: Vehicle Drive Cycle Analysis
Fleet DNA helps vehicle manufacturers and fleet managers understand the broad operational range for many of today's commercial vehicles. This tool offers access to vehicle fleet data summaries and visualizations similar to real-world "genetics" for medium- and heavy-duty commercial fleet vehicles operating within a variety of vocations.

Click on the vehicle type to access composite data, summaries, and visualizations. Use these tools to analyze and sort transportation data: **FAST** and **SRV**. Find more resources on the **Alternative Fuels Data Center**.

Contribute Data
To contribute your fleet data anonymously to help other fleets analyze and improve their drive cycle metrics, contact Sevin.Walton@nrel.gov, 303-275-4492.

For more information, refer to the [Fleet DNA fact sheet](#).

Delivery Vans, Delivery Trucks, School Buses, Transit Buses

Delivery Trucks
Fleet DNA has data for 9 delivery trucks operating for 128 days in the United States.

- Composite Data
- Data Summary and Visualizations

Related Studies
the operation of delivery trucks through these

Vehicle Dynamometer Evaluation and Comparison

NREL Firewall

TSDC Virtual Machines

Processing Server

Visualization and Reporting Server

Data Server

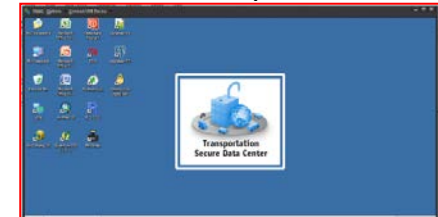
Secure Connection

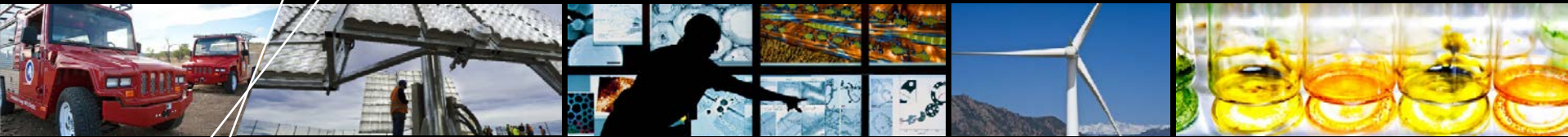
Virtual Desktop User's PC

VM Server

File sharing over secure network

Secure file transfer





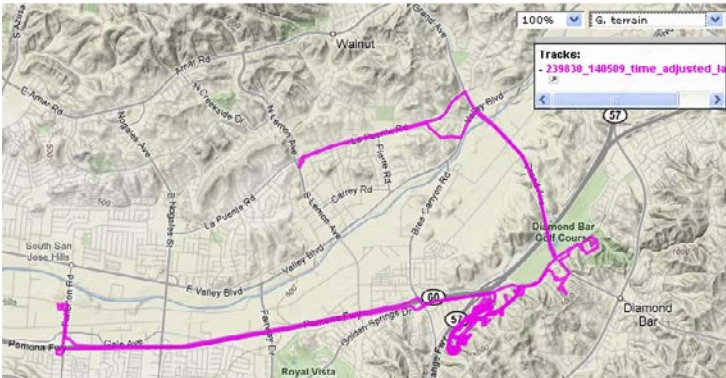
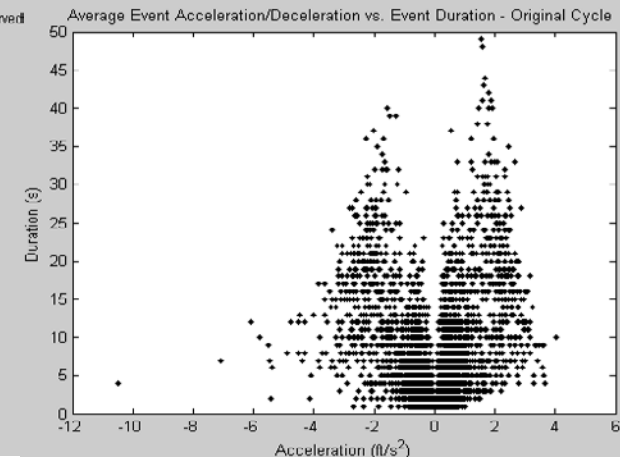
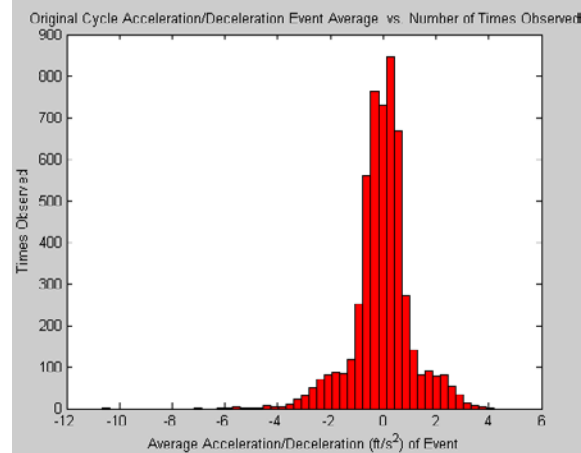
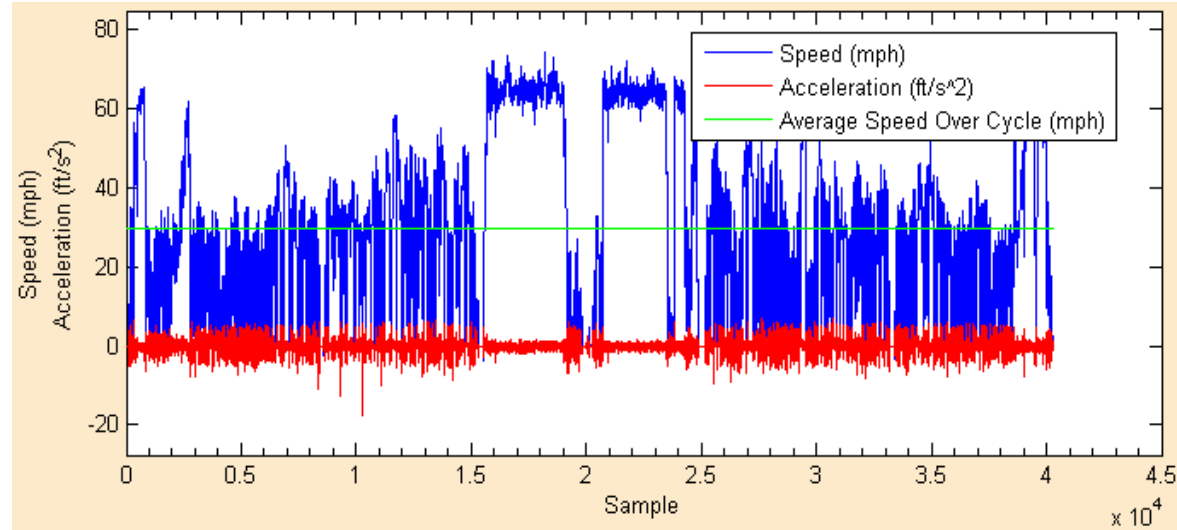
Analysis Examples

Vehicle Drive Cycle Characterization

Analysis Level: Vehicle

Drive Cycle Visualization

- Generate graphs of speed vs. time to observe vehicle operation
- Display histograms of useful statistics such as acceleration for analysis of performance envelopes
- Output latitude and longitude data for route visualization and mapping to ensure route/data consistency



Fleet Drive Cycle Characterization

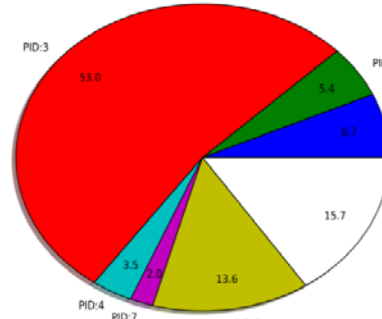
Analysis Level: Deployment

Delivery Van Vocational Database:

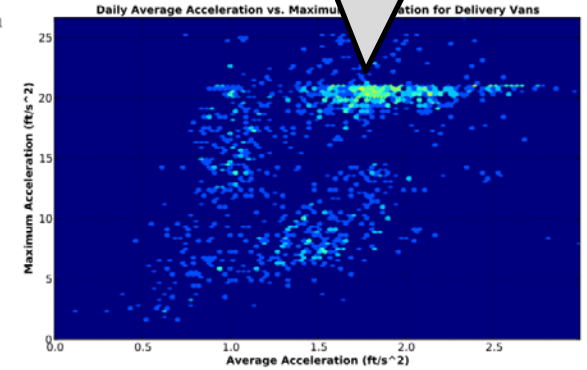
- Class 3–6 delivery van vehicles
 - Parcel, Linen, Food and Product vocations
- 92 unique vehicles-
 - Two conventional fuels (diesel/gas)
 - Four hybrid configurations (PHEV, GHEV, HEV, HHV)
 - ~600 EVs to be added
- 1,282 individual operating days
- Over a dozen unique geographic operating locations



Percent of Total Miles Collected by Project for Delivery Vans

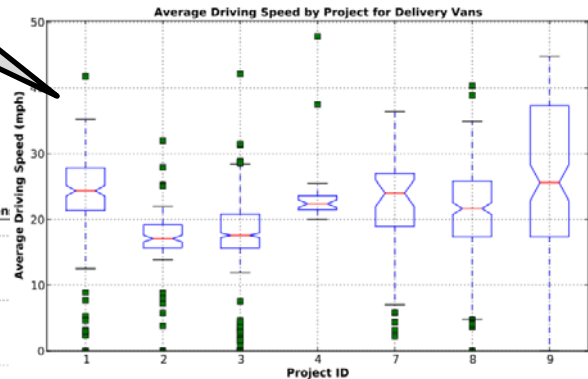
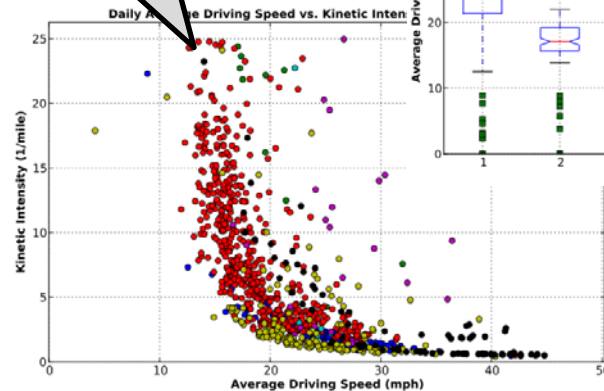


Common performance threshold



Wide range of operating speeds depending on location

Strong exponential relationship between speed and intensity



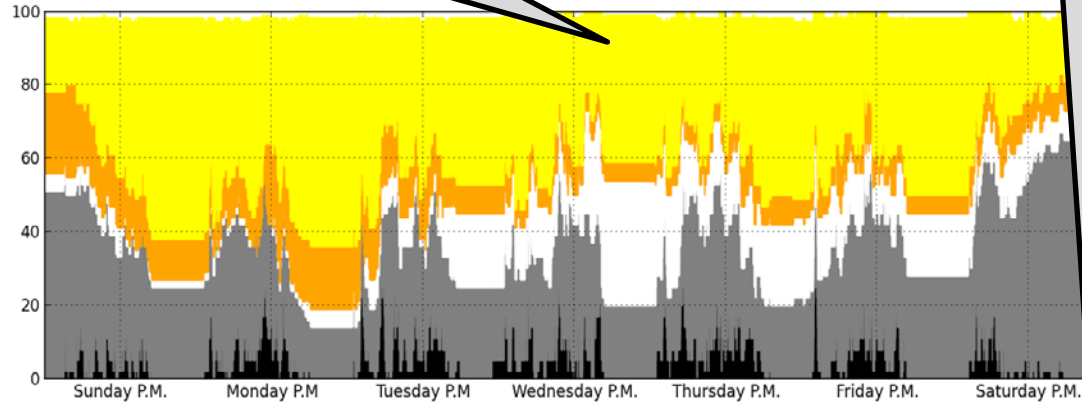
Charger Utilization: PHEV

Analysis Level: Vehicle relative to population

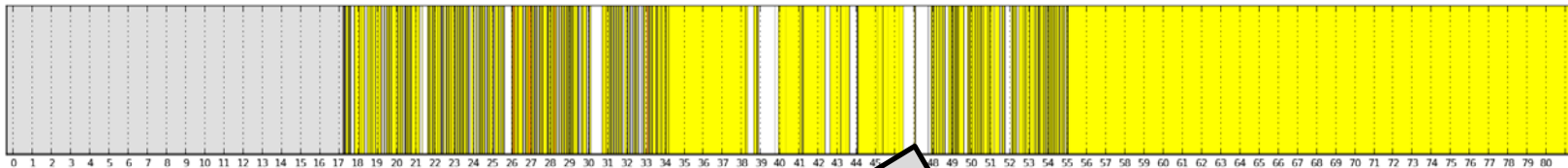
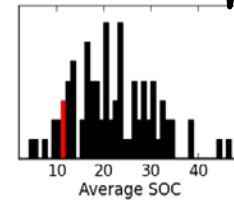
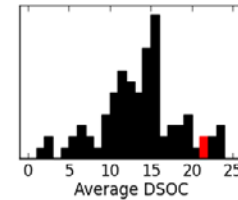
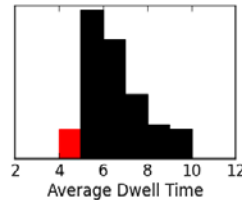
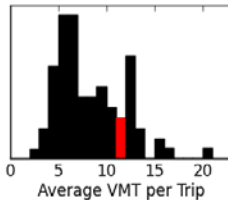
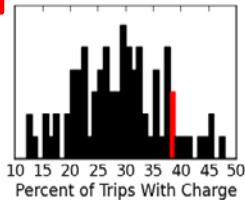
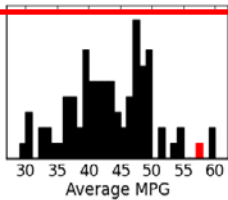
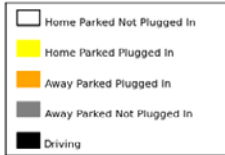
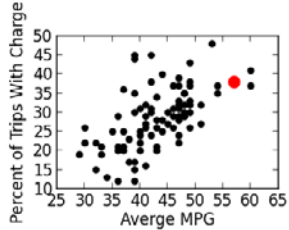
Vehicle statistics

Vehicle ID:	484	Trip Count:	1013
Charge Count:	382	Trips wCharge:	38%
Weeks:	35	Trips wHome Charge:	30%
Charges Home Charge:	80%	Home Locations:	1
	AVG	STD	Totals
MPG	57	49.10	149
Trip Distance	11.27	13.38	11419
Dwell Time	4.00	6.58	
SOC	11.34	20.72	
Delta SOC	21.35	36.26	

Vehicle state profile (% time by hour of week)



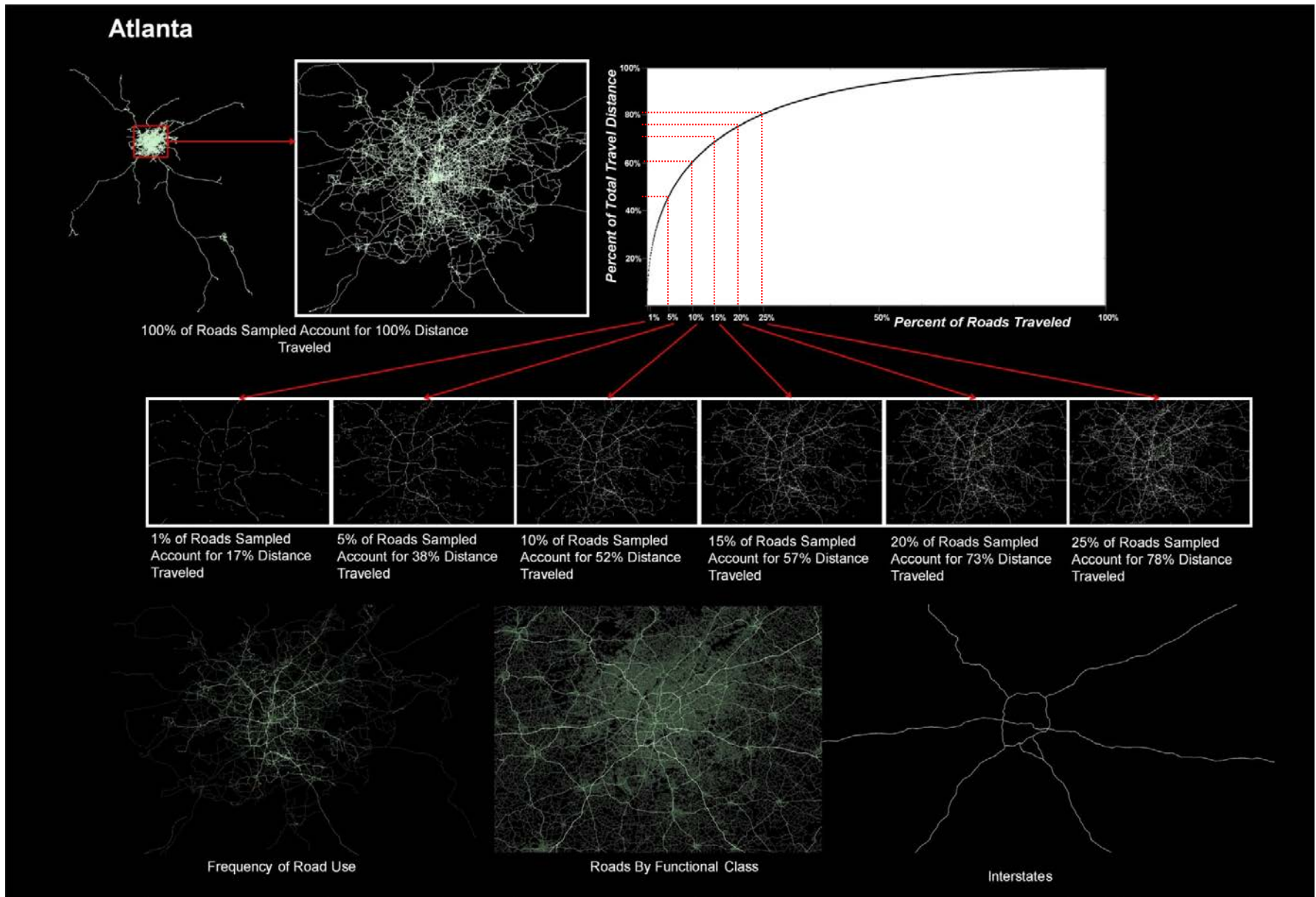
Vehicle statistics relative to population



State of operation by day of operation

Regional Road Usage

Analysis Level: Regional Infrastructure Utilization



Videos not available in this PDF

Cumulative Daily distance vs. Distance on Functional Class

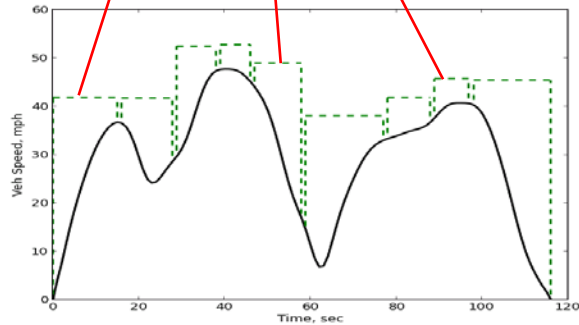
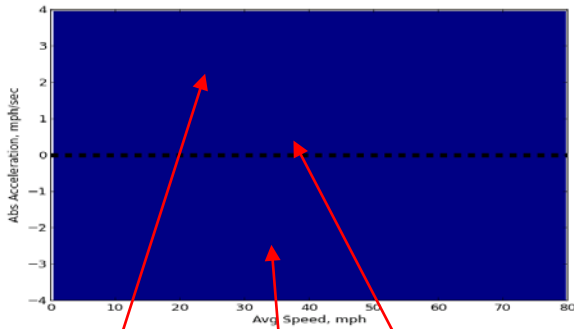
- *X axis is cumulative distance*
- *Y axis is cumulative distance functional class 1*

Vehicle Sub-sample Road Use

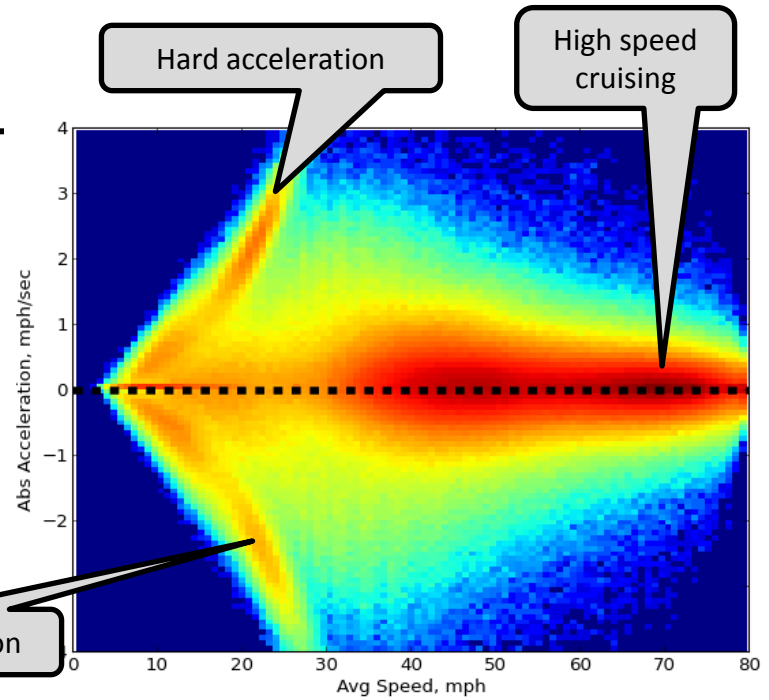
- *A small sample of vehicle days where the daily distance traveled exceeds 50 miles*

Drive Cycle Prediction

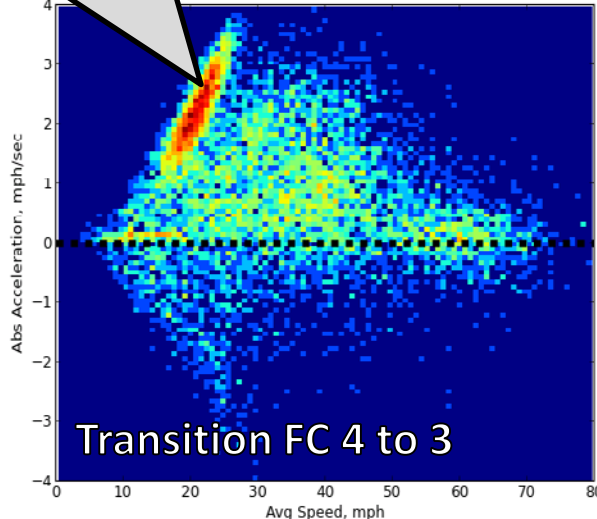
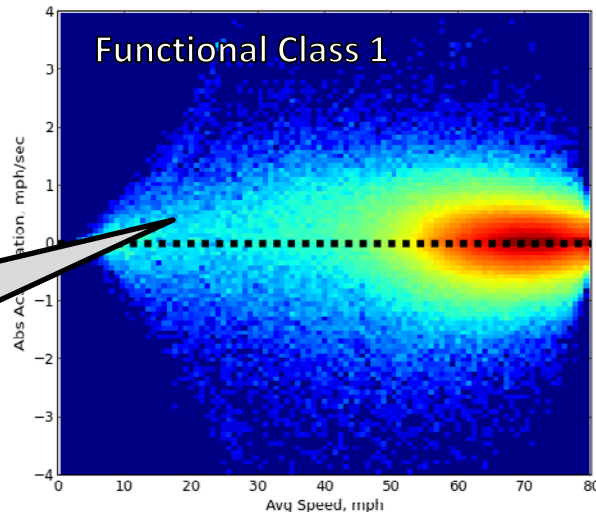
Analysis Level: Vehicles and infrastructure interaction



Low-speed data on FC1 assumed to be symptomatic of real-time road conditions (congestion, weather, etc.)

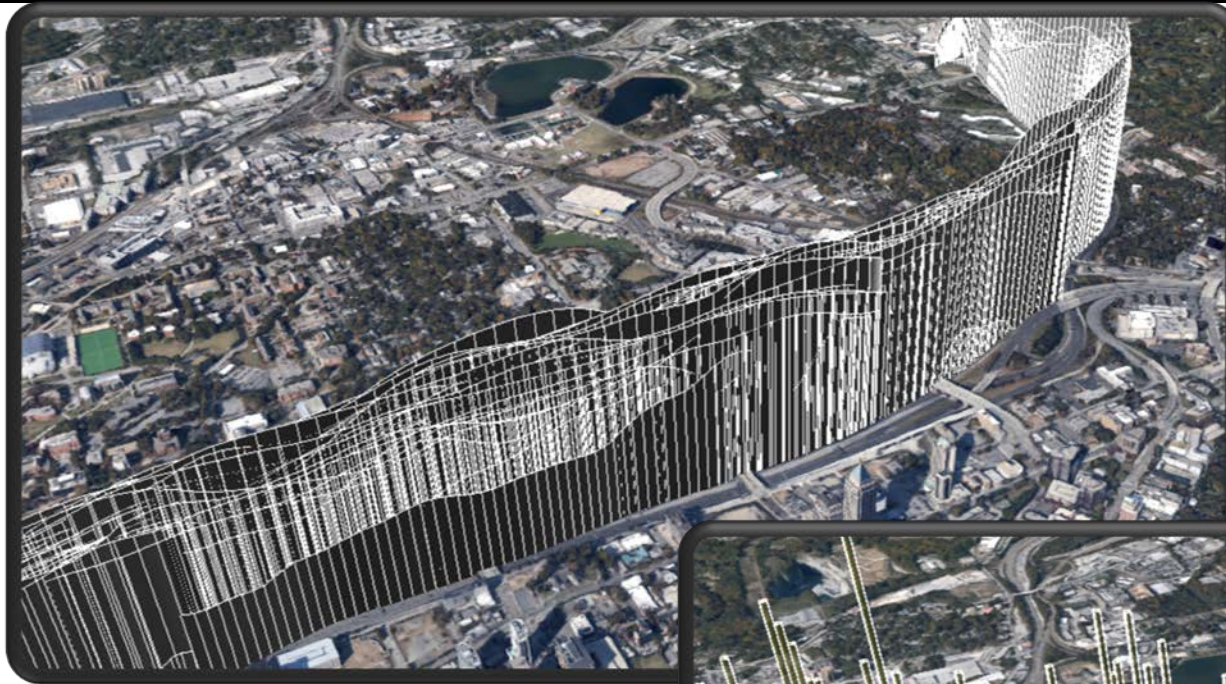


Near exclusive existence of data at positive acceleration



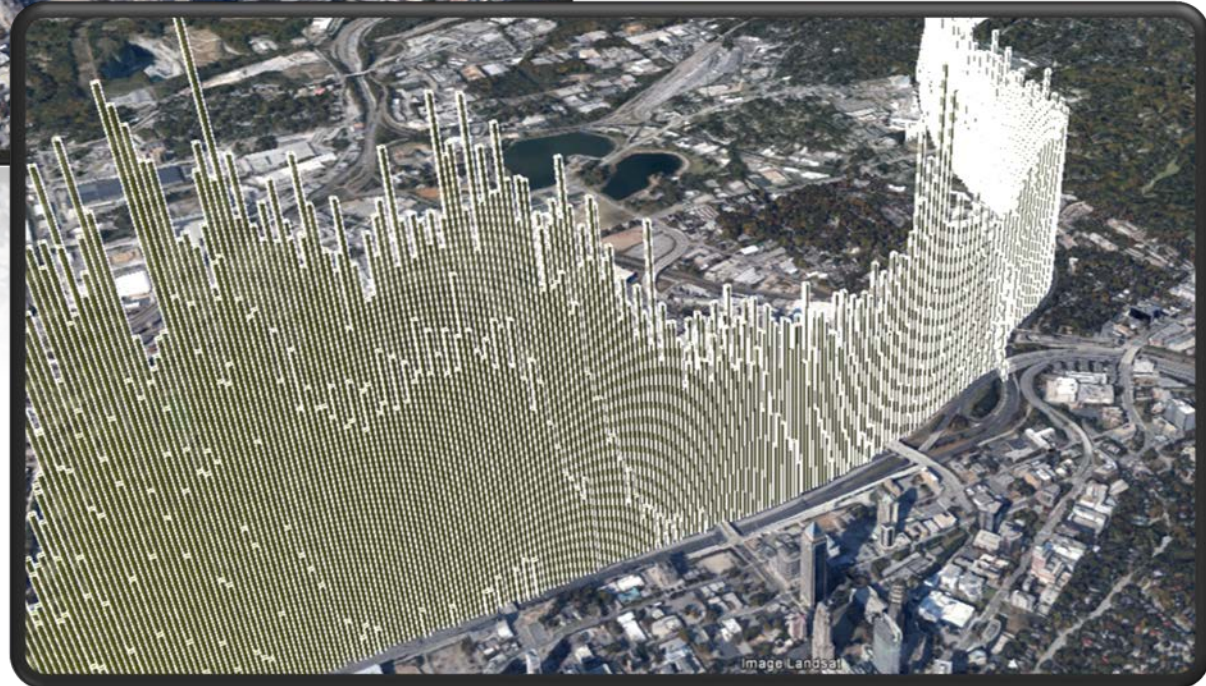
Road Use by Lane

Analysis Level: High resolution infrastructure utilization by time of day (5 a.m. to 10 a.m.)



Street Segments Driver Profiles

- *All point data collected for a grouping of roads ordered by trip visualized where speed is the z axis (left)*



Road Grid Segments Lane Summary

- *Road segment is divided into grid cells by lane*
- *Speed data are summarized by grid cell*
- *STDDEV of speed of travel by grid cell is visualized (right)*

Conclusion

- **By organizing sensitive, high-resolution data appropriately and linking the data to additional data sets using location and time, the quality and quantity of data available for researchers can be greatly improved while still maintaining restrictions placed upon the data**
- **Benefits**
 - Results in an infinite number of ways to group and sort the data for reporting
 - Increases the number of applications for the data, and the amount of data distributed without violating restrictions
- **Why never before?**
 - Location data are very restricted and have a high value, so very little data of this kind are available publicly
 - In-vehicle navigation systems utilize similar approach to analysis, but do not provide it publicly
 - Programming languages and databases now incorporate spatial analysis capability, providing free and open source tools with which to perform analysis