USING THE UD CLUSTER TO DEVELOP A STATEWIDE SPEED SURVEY FROM VEHICLE GPS DATA

David P. Racca

Center for Applied Demography and Survey Research (CADSR)
University of Delaware
Graham Hall, Rm 284, Newark, Delaware 19716
Phone: (302) 831-1698
E-mail: dracca@udel.edu

Center of Applied Demography and Survey Research

- The Center’s primary mission is to ensure that the best possible data and information on important public policy issues are developed and made available
- Professional staff working in areas of transportation, education, finance, emergency management, health policy, demographics, land use planning, environmental protection.
- An advanced survey research center

A Focus on Measurement and Performance

- Team DelDOT – Transparent, Efficient, Accountable, Measured
- Increased capability with data systems. Detailed planning, large amounts of data being collected. Increasing automation of signal system and field measurements
- Besides safety measures, travel time and speed throughout the data is the most important performance measure

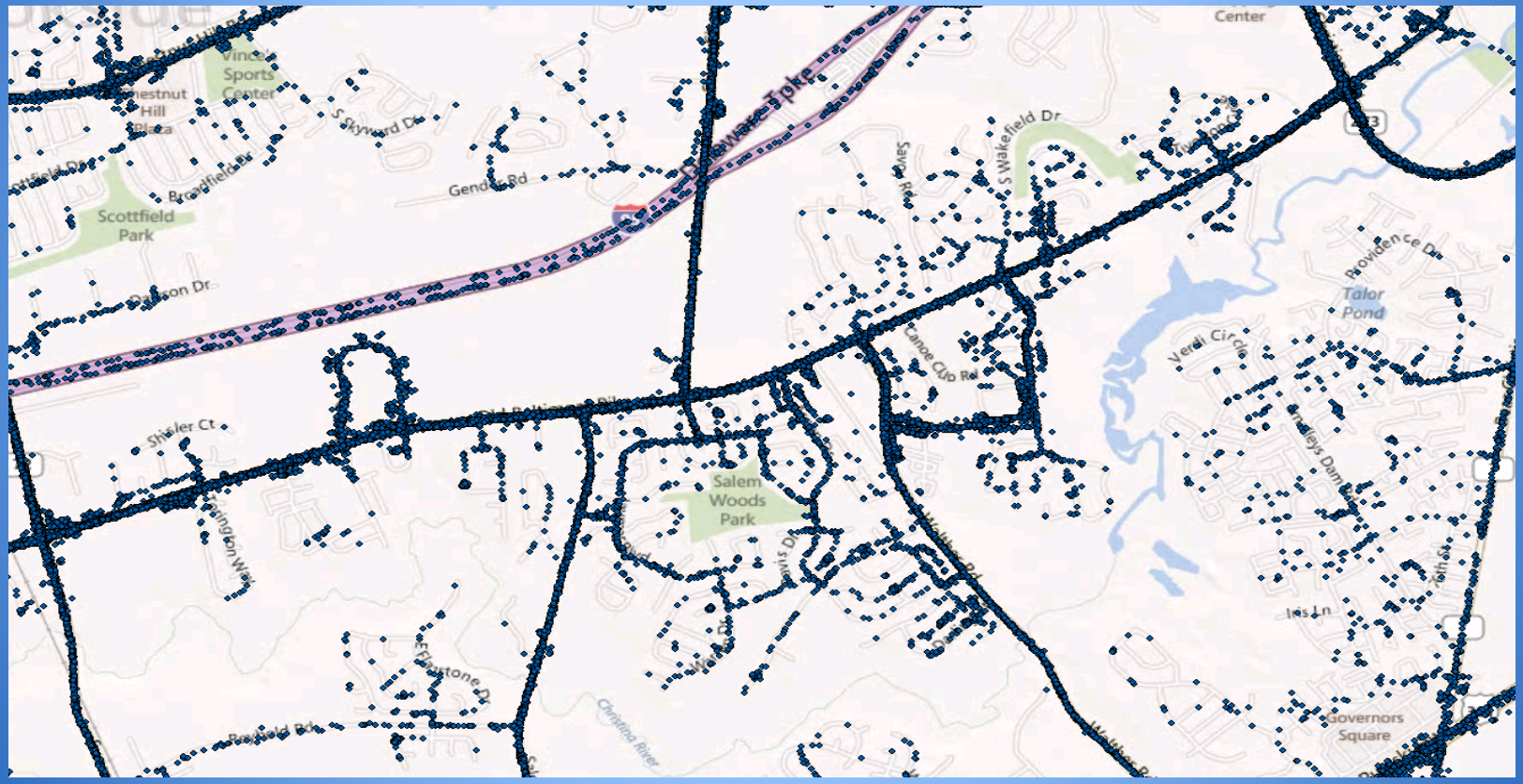
State of Delaware Public Vehicle GPS Data An Opportunity to Develop a Statewide Travel Time / Speed Survey

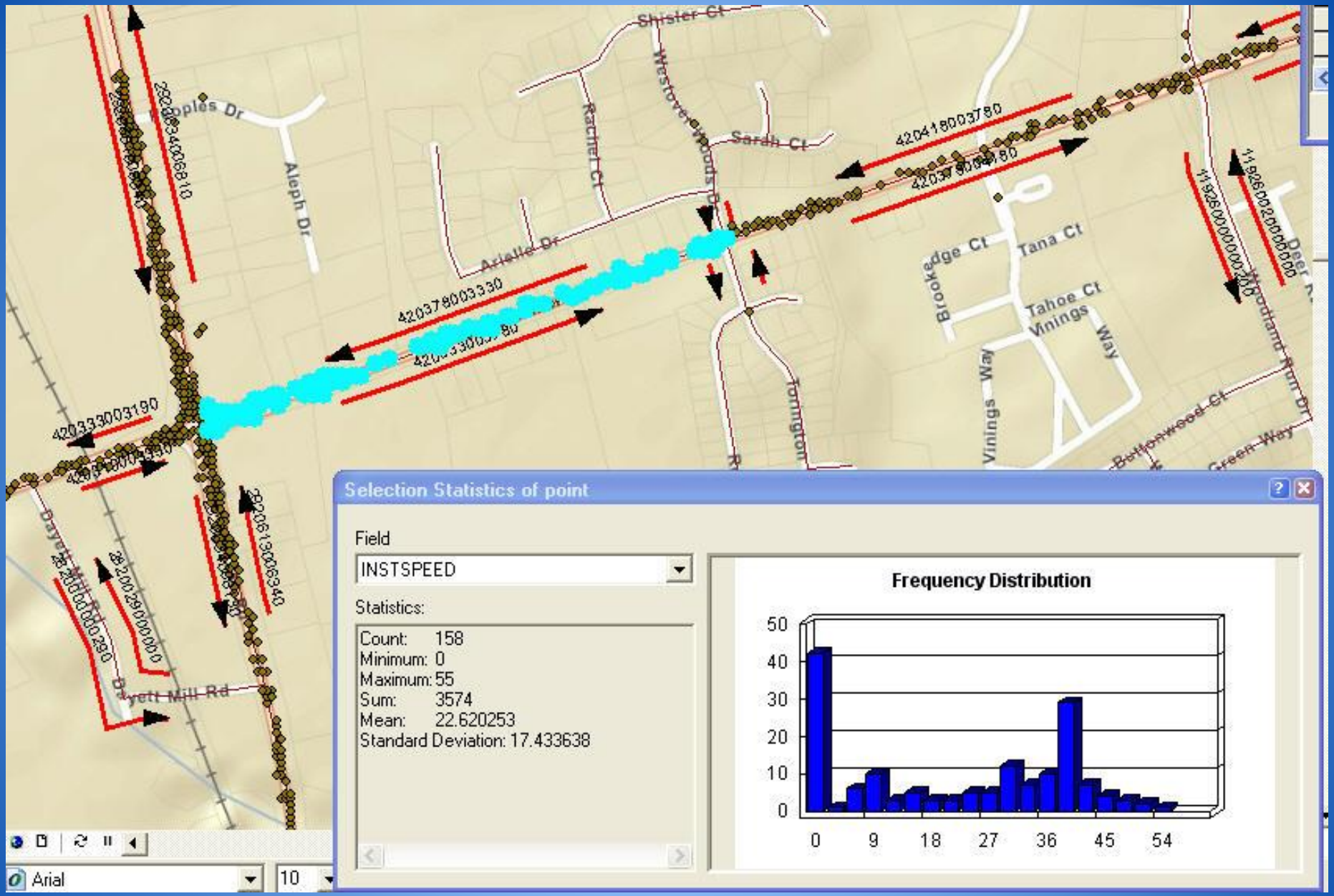
- About 2400 vehicles broadcasting location every two minutes.
- 2 million point measures per month
- Providing 5 to 10 million travel way measures per month
- Does not include public safety, transit buses, or road maintenance vehicles.
- 40% passenger cars, 34% passenger vans, 23% pickups and SUVs
- Data as far back as year 2007

An Opportunity to Develop a Statewide Travel Time / Speed Survey

- Addresses roads large and small
- Collection costs already covered
- Includes detailed trip data allowing for analysis of turning movement statistics

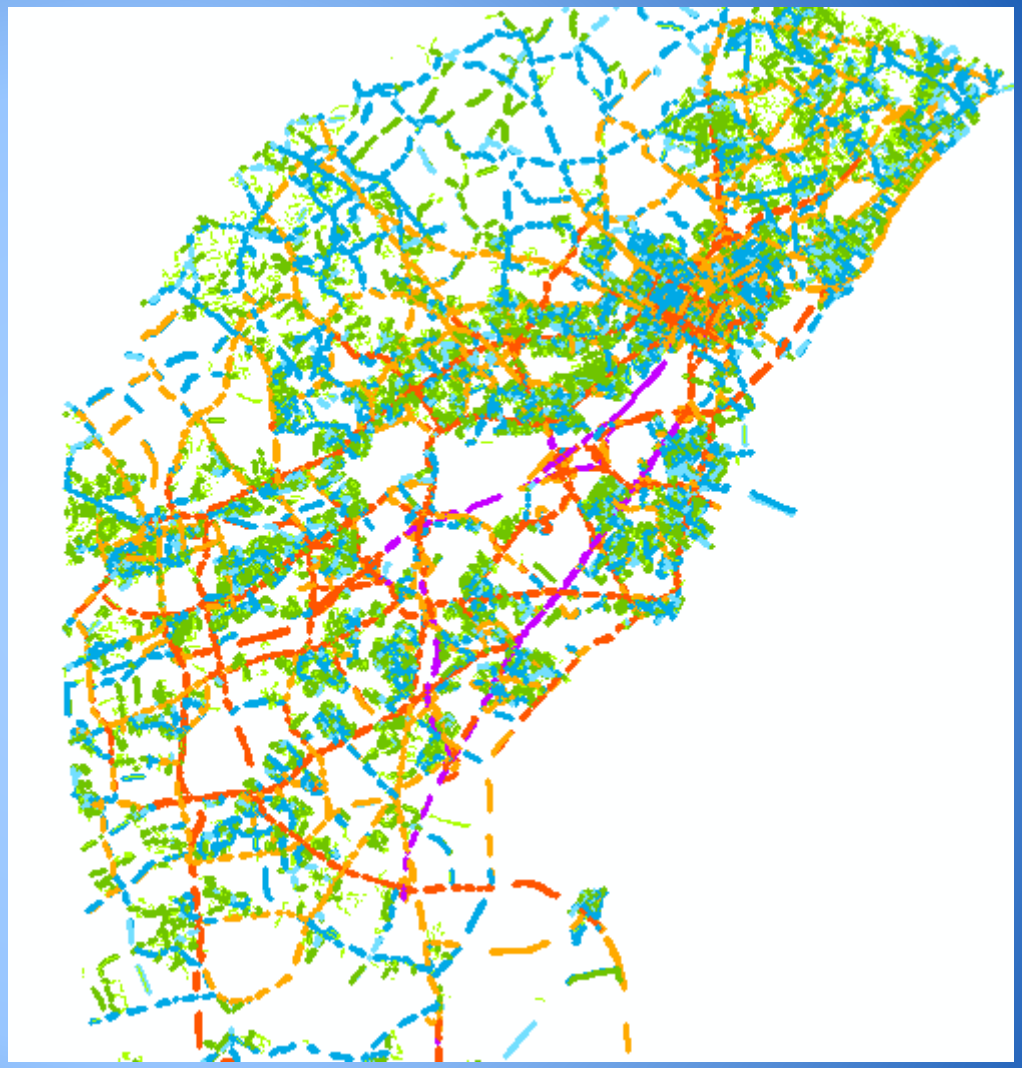
State Vehicle GPS Measurements



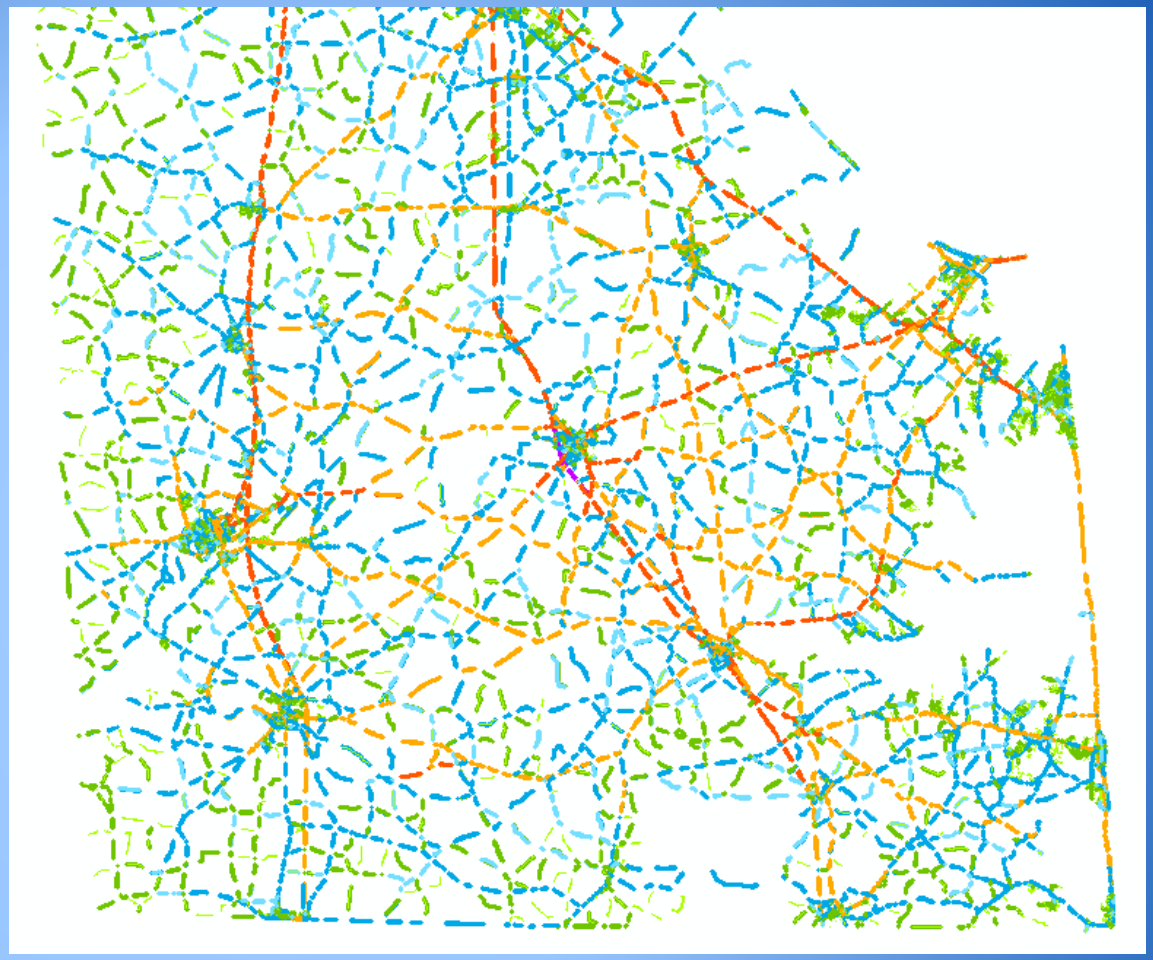

Shrister Ct
Westover Woods Dr
Rachel Ct
Sarah Ct
Aleph Dr
Arielle Dr
Brookedge Ct
Tana Ct
Tahoe Ct
Vinings Way
Torrington
Bottomwood Ct
Green Way
Deer
Bayett Mill Rd
420418003780
420378003330
420333003190
Selection Statistics of point
Field
INSTSPEED
Statistics:
Count: 158
Minimum: 0
Maximum: 55
Sum: 3574
Mean: 22.620253
Standard Deviation: 17.433638
Frequency Distribution
50
40
30
20
10
0
0
9
18
27
36
45
54
Arial
10

State Vehicle GPS Measurements Sample Trip

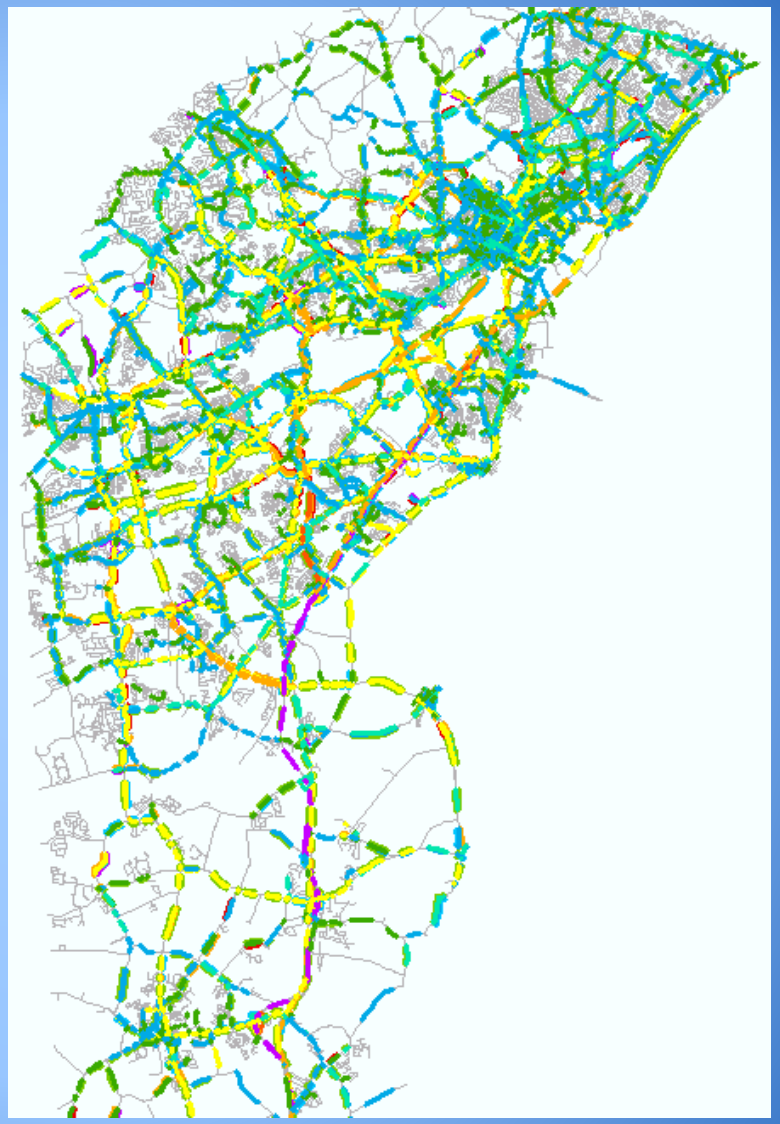
Northern New Castle County Weekday Observations 2012



Sussex County Weekday Observations 2012



Hourly 8am Observations



- 5 - 10
- 11 - 50
- 51 - 100
- 101 - 500
- 501 - 2000
- 2001 - 3000
- 3001 - 4016

Processing

- Capture historical GPS data by querying Networkfleet internet services for Delaware vehicles. Process GPS XML response data
- Create GPS point databases and GIS files
- Extract and associate GPS points with particular trips taken through time
- Build a trip and link based version of the GPS data
  - Estimate the path taken between GPS readings
  - Associate particular point measures with a particular road link, direction, and tuning movement
- Where portions of a particular trip include road links with no actual measurement, interpolate speeds estimates and travel times between GPS measures
- Screen the data for errors and anomalies.

The Processing Challenge

- Processing takes a great deal time, on a fast personal computer it can take 3 days to process 5 days worth of data.
- The development of the processing routines took many revisions as the data was processed and examined. The initial stages of development required that datasets be reprocessed a number times.
- As the process develops and errors are found, particularly in the underlying network model it is necessary to reprocess the data. Sometimes when corrections are identified in portions of the network, it is necessary to reprocess previous months' data.
- This quickly got beyond personal computer resources.

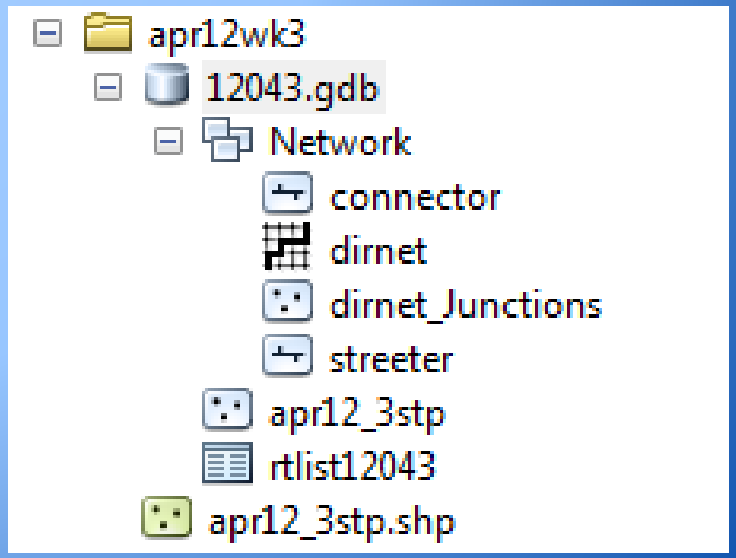
The UD Community Cluster To the Rescue

- The Mills Cluster allows for the operation of geographical information system software (ARCGIS) that includes access and processing of geodata sets.
- This capability includes the running of complex Python code to make the calculations.
- The ability to easily run and time batch jobs was crucial in being able to keep the processing going. The use of notifications/emails when jobs were completed was very helpful
- Processing speed was effectively increased by at least 10 times.
- Having this type of power allowed for additional quality control steps and more testing of routines. Development time was significantly improved (made possible)

Discussion of the process

- Data for each month was divided into 6 portions. Each portion of vehicle GPS data was usually about 400 to 600 MB. Through experimentation this appeared to be close to an optimum size. Experience so far with ARCGIS has shown that when data sets get close to 1 million features things slow down. Each portion contained around 400,000 measurements.

- Each portion was developed as a file geodatabase that included a routing network for a time corresponding to the time the measures were taken.



Details of the process (continued)

- Data transfer all done with FileZilla
- Mills was accessed within SSH Secure Shell
- Processes were run in exclusive mode

```
qsub -l exclusive=1 arc-submit-RTXMLRun.qs
```

- Use of Python processing for ARCGIS is provided in mills by referencing the ARCPY library

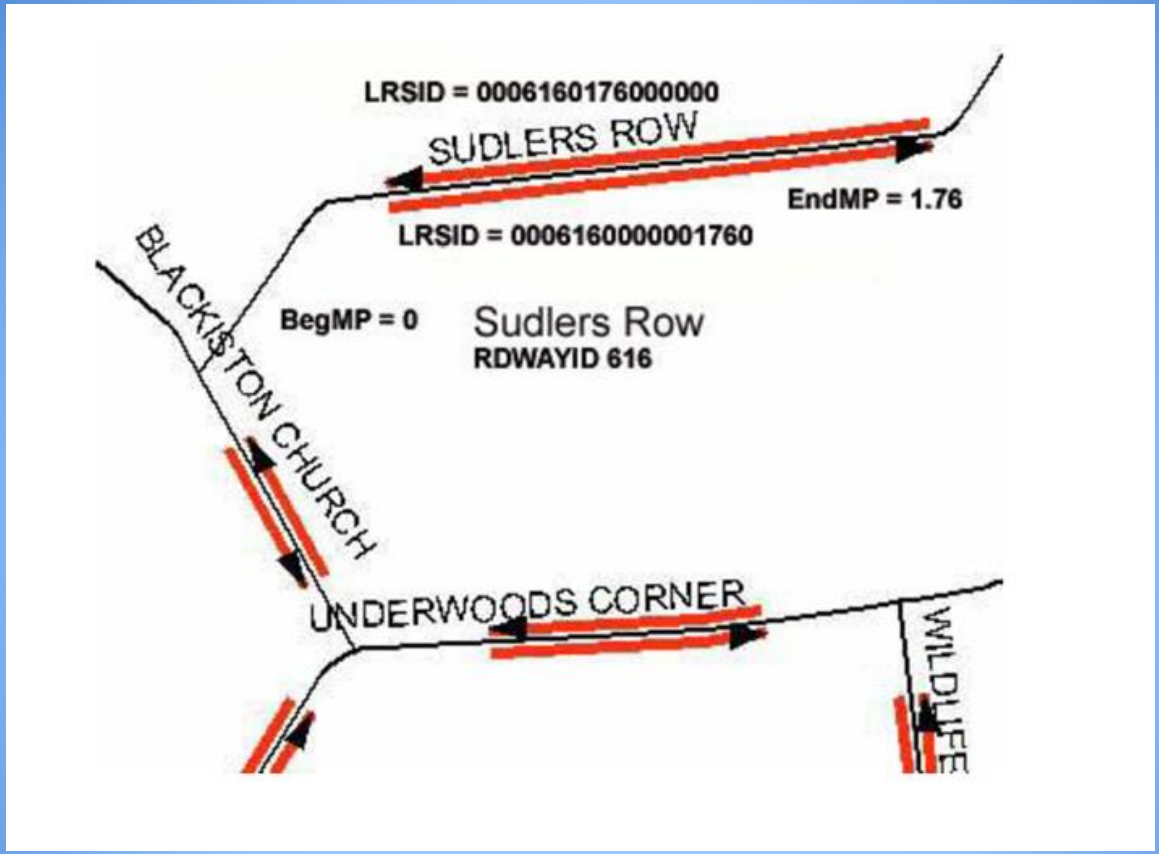
```
import arcpy
from arcpy import env
```

- ARCPY include a very large set of GIS functions that can be incorporated into python programs running on Mills. The primary functions used were those that involved determination of optimum path and allocation of speed measures to links in the transporation network

Details of the process (continued)

- This effort did not employ parallel processing as this is currently unsupported with ARCGIS.
- It does seem feasible to replace higher level ARC functions with programmed functions that would allow for parallel processing. Algorithms for optimum path on a link/node network have been in use for decades.

Identification illustration



To relate a measure to a particular turning movement a “S”, “L”, “R”, or “U” is appended to the LRSID, for example. Left turn from Sudlers Row → LRSID = 0006160176000000L

lrsid	weekday	hourofday	speed_mean	hourobs	freeflow75
0000010098001020S	1	1	41.1	213	44.5
0000010098001020S	1	2	47.3	58	44.5
0000010098001020S	1	3	43.2	5	44.5
0000010098001020S	1	4	40.6	13	44.5
0000010098001020S	1	5	47.2	16	44.5
0000010098001020S	1	6	42.2	287	44.5
0000010098001020S	1	7	38.9	509	44.5
0000010098001020S	1	8	39.4	2058	44.5
0000010098001020S	1	9	40.7	2697	44.5
0000010098001020S	1	10	40.2	1895	44.5
0000010098001020S	1	11	39.8	1393	44.5
0000010098001020S	1	12	39.0	1111	44.5
0000010098001020S	1	13	39.3	1198	44.5
0000010098001020S	1	14	39.8	1066	44.5
0000010098001020S	1	15	40.7	975	44.5
0000010098001020S	1	16	41.4	1007	44.5
0000010098001020S	1	17	41.8	1280	44.5
0000010098001020S	1	18	40.5	622	44.5
0000010098001020S	1	19	42.5	284	44.5
0000010098001020S	1	20	44.2	168	44.5
0000010098001020S	1	21	44.6	137	44.5
0000010098001020S	1	22	49.6	180	44.5
0000010098001020S	1	23	50.7	73	44.5
0000010098001020S	1	24	46.2	55	44.5

Sample Output

Weekday Hourly

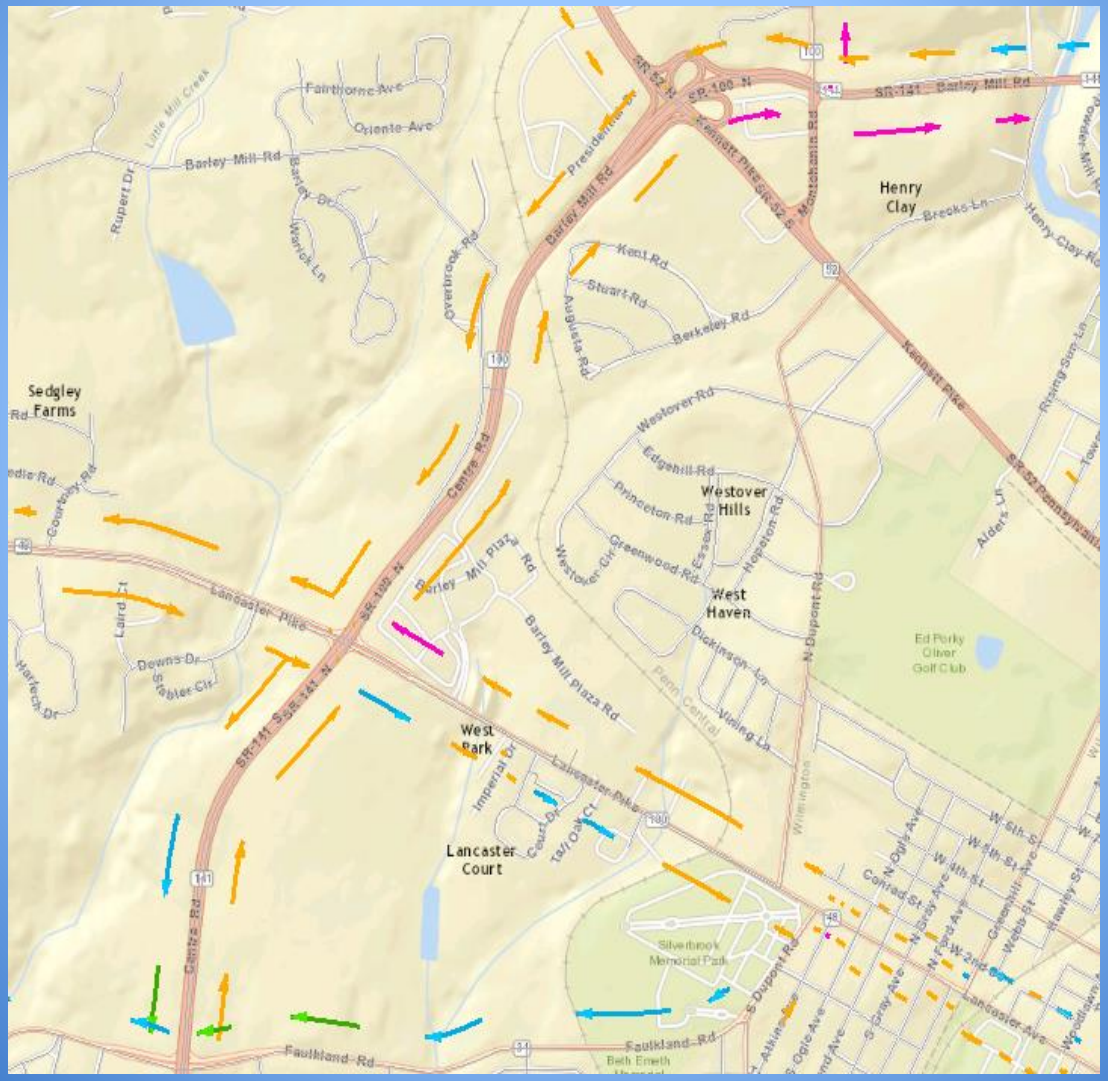
"S" Straight or Thru Shown

Also available are Right & Left

Segments statewide included

Example Detail Captured for Road Links

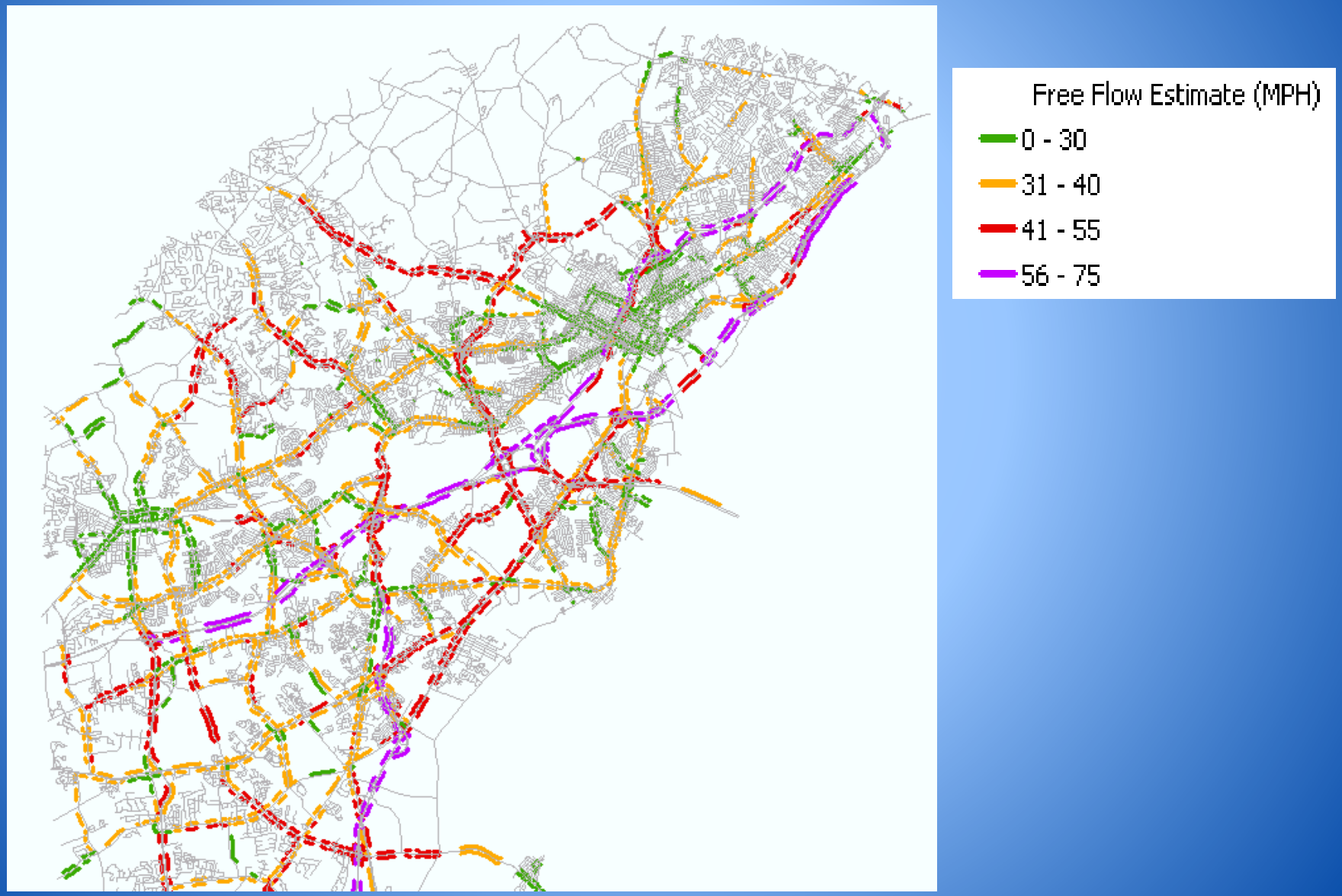
Routing network is segmented at every major or minor intersection



Summary of Features of the State Vehicle GPS

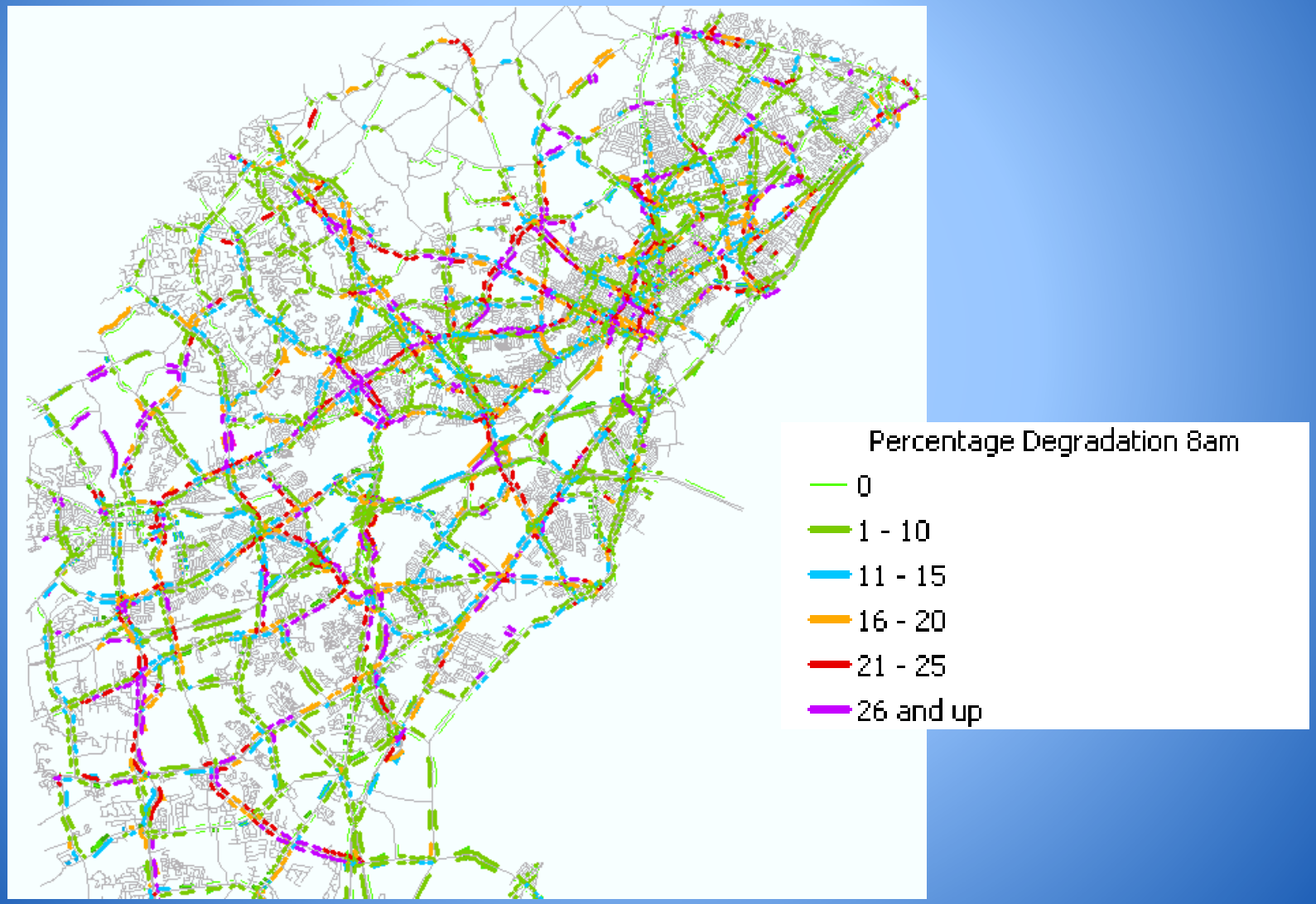
- Data available for up to 6 past years
- Wide coverage, data for small and large roads
- Captures speeds and travel times relative to turning movement
- Measures available at great detail, road link breaks at all intersections, large and small
- Delay at intersections by turning movement , incorporated into road link speed / travel times. Ideal for generation of time sensitive routing network impedance.
- Cost of collection covered in existing program

Calculation of Free Flow Speed as the 75 percentile Of Hourly Averages (just major roads shown)



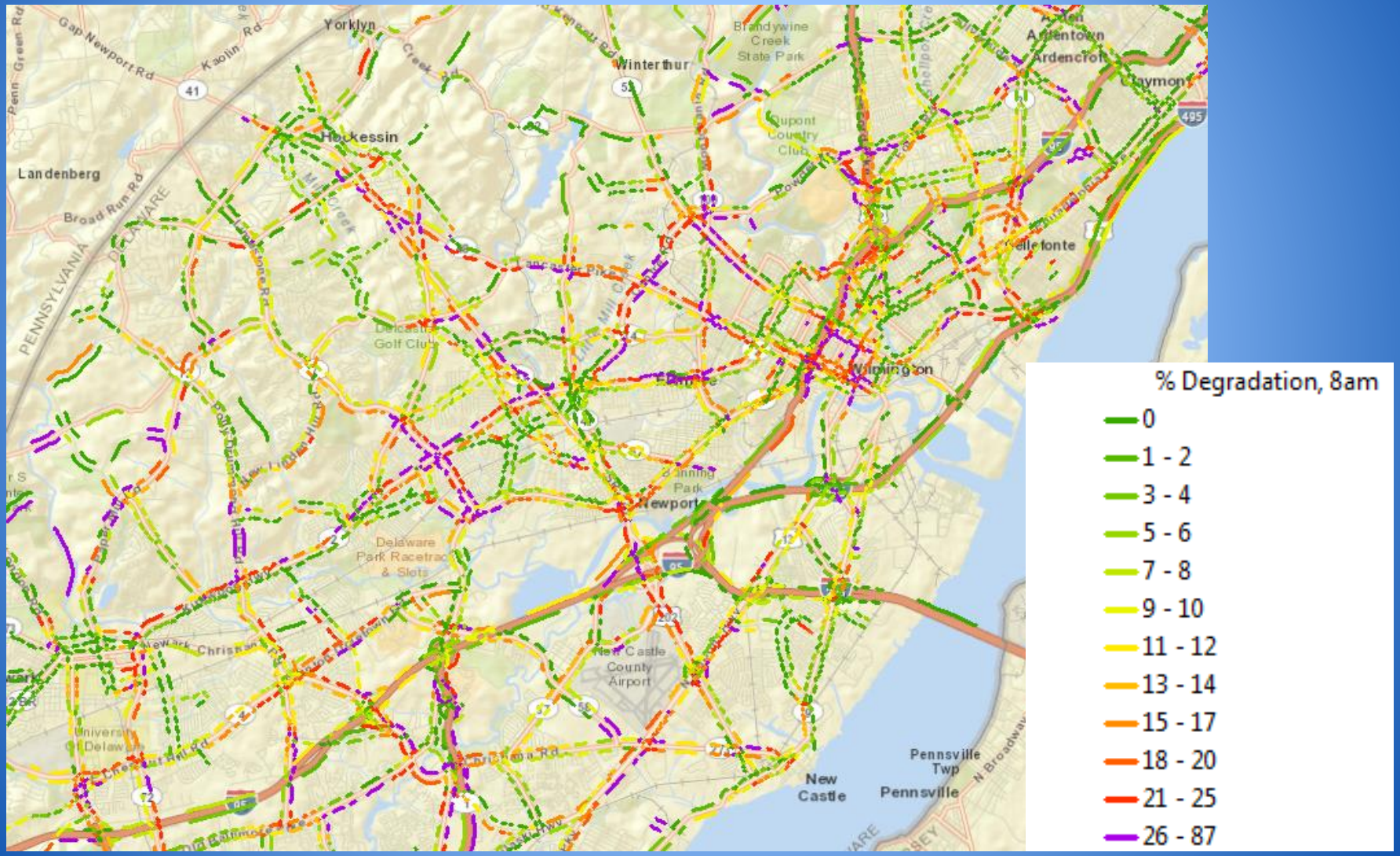
Calculation of Percent Degradation at 8am, weekdays

Percent degradation = 100 * (freeflow75 – speed) / freeflow75
Calculated from weekday hourly averages



Calculation of Percent Degradation at 8am, weekdays

Percent degradation = 100 * (freeflow75 – speed) / freeflow75
Calculated from weekday hourly averages



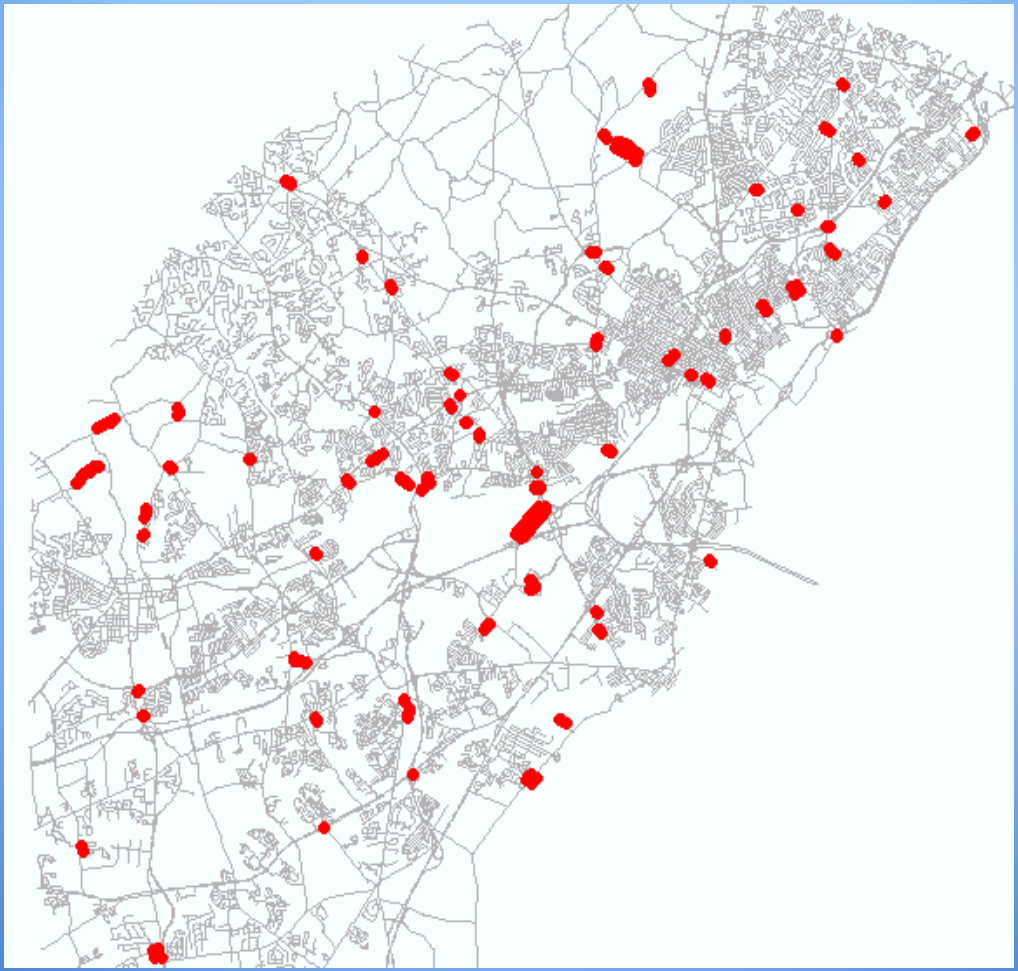
Travel Time Reliability

Trip Category	Mean Time	95%	Buffer Index	Planning Time Index	Avg Hourly Volume	#Trips
All Trips	6.4	8.1	.27	1.3	334	712
Peak (AM/PM)	6.7	8.5	.27	1.4	664	149
AM Peak (7-9)	**6.4**	7.8	.22	1.3	632	76
PM Peak (16-18)	7.0	9.8	.40	1.6	697	73
Midday (9-16)	6.4	7.9	.23	1.3	506	366
Weekend (all day)	6.5	8.3	.28	1.4	278	143
Off Time (<7 & >17)	6.0	7.8	.30	1.3	152	54
All Trips 2010+Jan11	6.6	8.4	.27	1.4	328	316
All Trips 2011 -Jan11	6.3	7.8	.24	1.3	336	396
Peak 2010+	7.1	9.8	.38	1.6	647	71
Peak 2011-	6.4	8.0	.25	1.3	674	78
Midday 2010+	6.5	8.3	.28	1.4	507	138
Midday 2011-	6.3	7.8	.24	1.3	508	228

Buffer Index is extra time that travelers must add to their average time to ensure on=time arrival. Planning Time Index represents how much total time a traveler shou allow to insure on-arrival.

Intersection Study

Left turns that are most effected (> 40% degradation) by morning (8am) congestions



Other potential applications

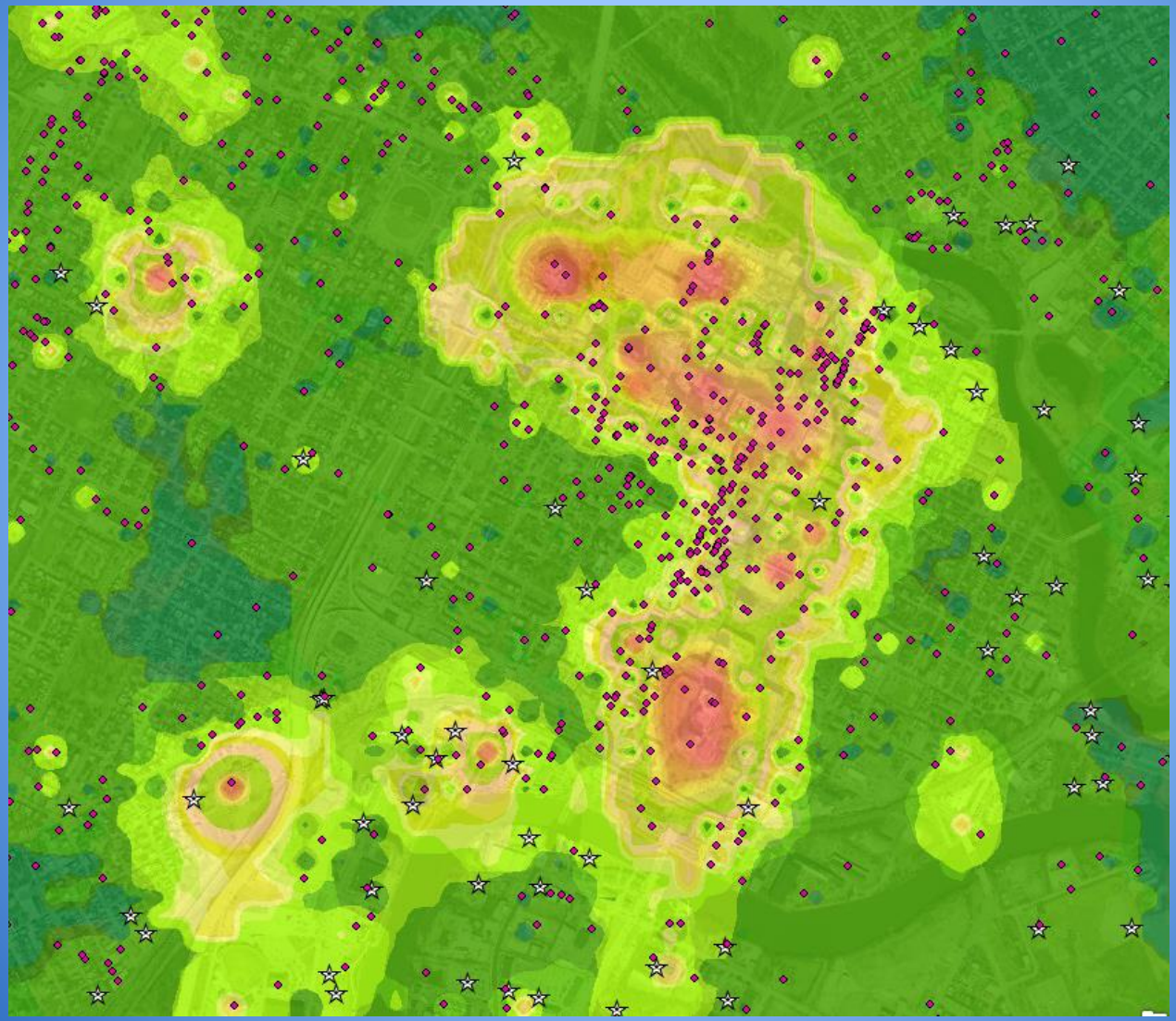
- Before and after studies, land use and facility changes
- Examining delay at intersections
- Estimations of capacity and studies of volume speed relationship
- Relating traffic flow to land use and travel demand
- Multimodal studies
- Applications of a detailed time sensitive routing network, such as accessibility studies

Other Possibilities For Using The Power of Mills For Complex Geoprocessing

- Origin and destination tables getting completely out of hand
- High resolution spatial distributions such as employment, property value, and accessibility.
- Rastor processing
- Emergence of Big Data in transportation and land use analysis.
- GIS in its higher forms involves the study and display of spatial relationships

Accessibility

Distribution of Employment



Summary

- Mills provided the capability to develop a valuable resource for examining the performance of roads in Delaware. Processing and development time improved by a factor of 10.
- GIS software in Mills could serve numerous applications some yet to be imagined.
- Mills is a powerful tool that includes features that make staging and processing easy.
- Project report is at the CADSR website,

www.cadsr.udel.edu , search for "speed survey"