Using AVL-equipped Buses to Detect Recurring Traffic Congestion on Surface Streets

Mark M. McCord,* Prem Goel,** Rabi G. Mishalani,*
Cheng Chen,* and Jinguo Gao**

*Dept. of Civil and Environ. Engineering and Geodetic Science
**Department of Statistics
The Ohio State University

Ohio Transportation Engineering Conference
Columbus, OH
October 28, 2009

Acknowledgements

- US DOT – Federal Transit Administration
- OSU Traffic and Parking Services

The views, opinions and statements contained in this presentation are solely those of the authors and do not represent the official policy or position of the Department of Transportation, the Federal Transit Administration, The Ohio State University, or any other entity or organization.

"Buses as probes"

- Many transit buses are presently equipped with AVL systems (low marginal cost)
- Buses regularly and frequently cover surface streets
- Let's use the "infrastructure" to help identify traffic conditions
- Any AVL equipped public fleet would do?

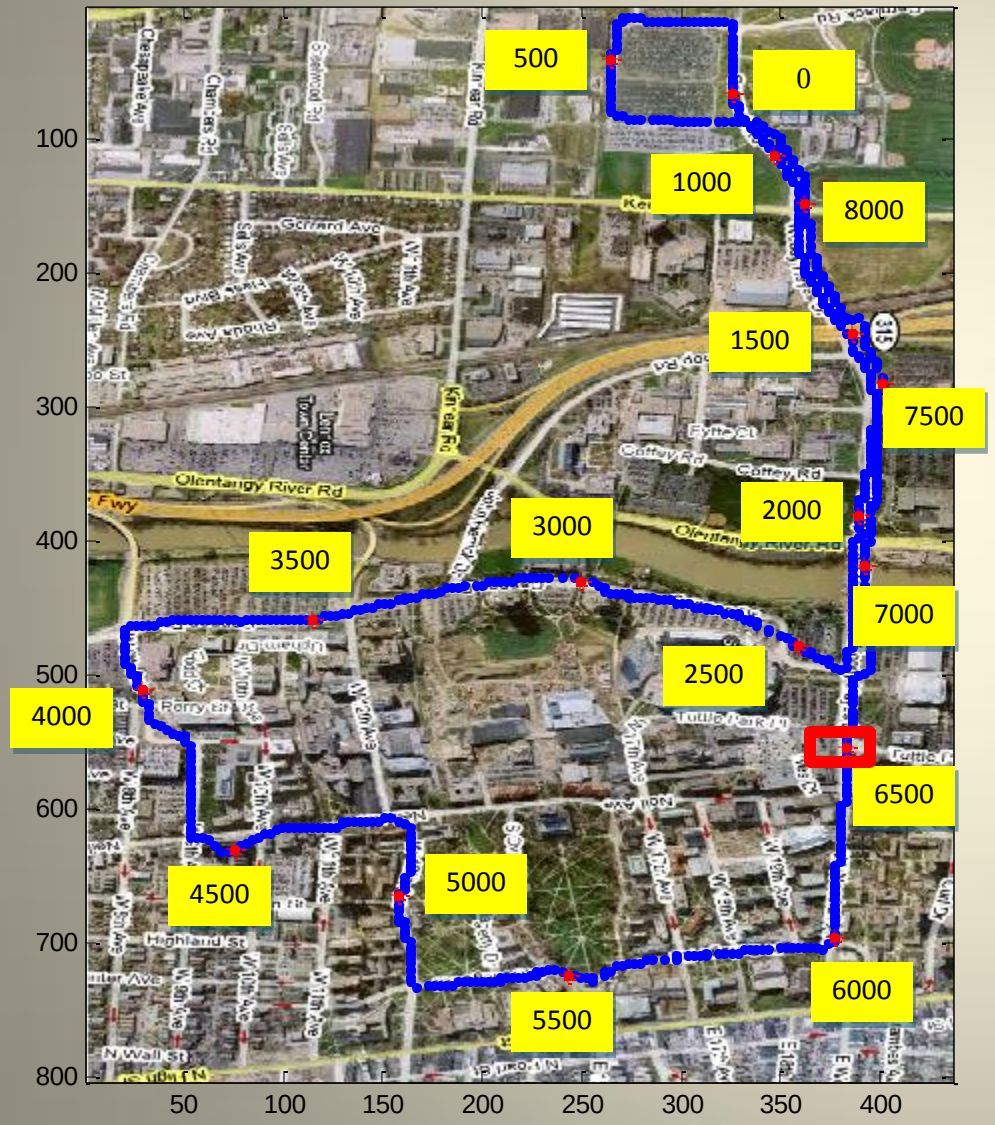
Buses as probes on surface streets

- Traditional difficulty: Buses behave differently than prevailing traffic
- Approach used in this effort
  - "Compare buses to buses"
  - Find systematic time-of-day differences in bus speeds to infer time-of-day differences in traffic conditions
  - Recurring congestion: More appropriately, time-of-day reductions in bus speeds to *indicate* possible recurring congestion

Overview of remaining presentation

- Location specific concept
- Route-based concept
- Empirical study and results
- Comments with respect to classical approaches
- Ongoing and future work

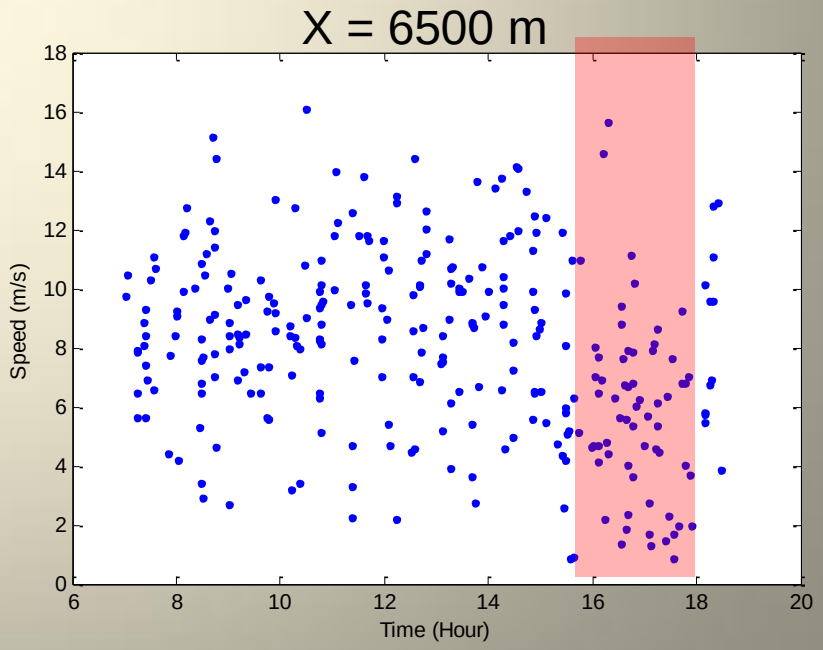
Location-specific concept



At a location X on a route

Group bus speeds over several days assumed homogeneous in bus operations (e.g., weekdays)

Identify time periods where speeds are markedly lower than the speeds at other time periods



Congestion Identification at Location X: "Low-speed" Method

- Identify low speed range (e.g., 0.35 x 95%-ile of speeds)
- Determine:
  - *N*: # overall speed observations in data set
  - *L*: # overall low speed observations data set
- Divide day into time intervals
- For a time interval, determine:
  - *n*: # speed observations in selected time interval
  - *l*: # low speed observations in selected time interval

Location X

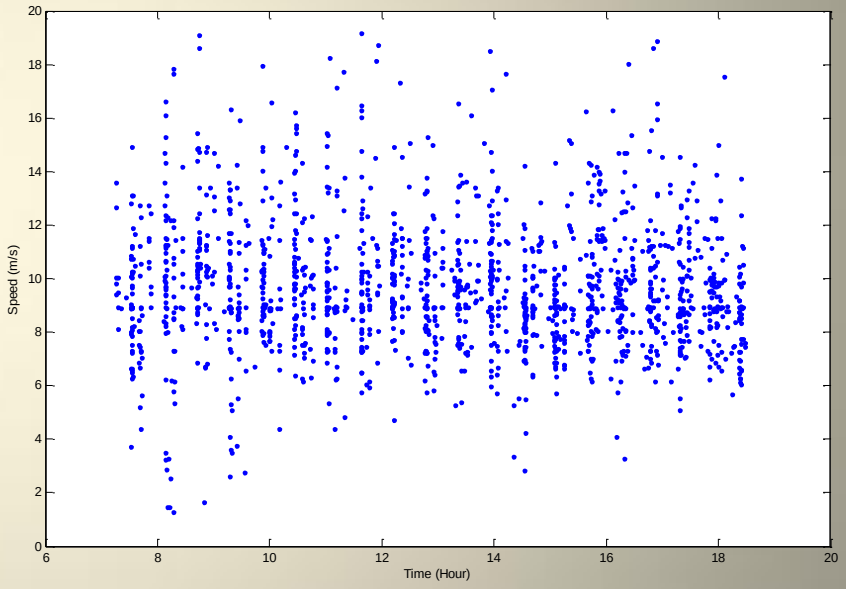
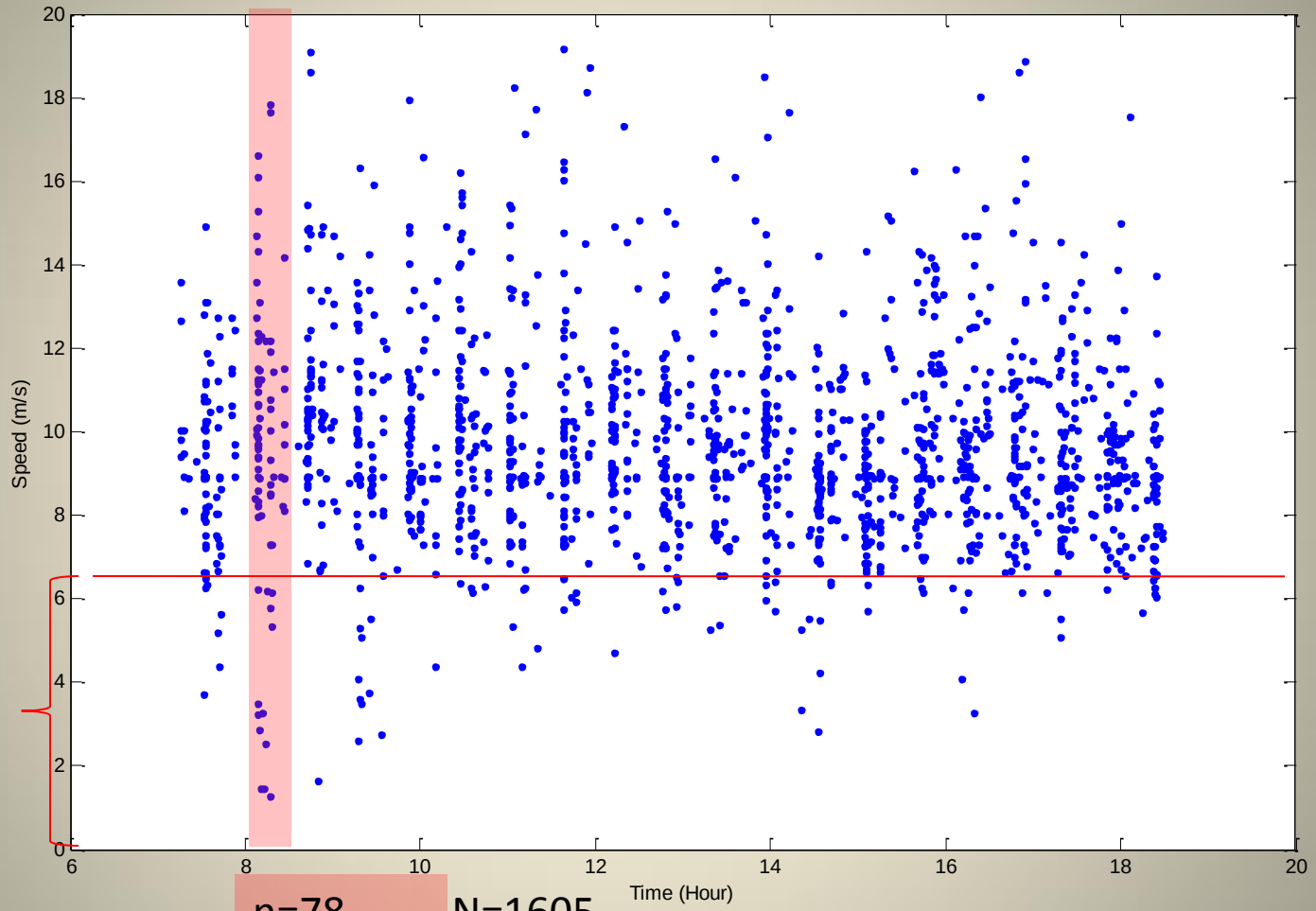


Illustration of Low-speed Method at Location X = 2000



Low Speed Range

n=78	N=1605
l=12	L=73
p=0.154	P=0.045

Congestion Identification at Location X: "Low-speed" Method

Given (*N*,*L*; *n*,*l*)

- Test: Is it unlikely to observe ≥ *l* low speed observations in *n* trials when considering a distribution (N, L)?
  - If *yes*, time interval *indicates* recurring congestion
  - If *no*, time interval does not indicate recurring congestion (**is *not* same as indicating no congestion**)
- Repeat for all time intervals

n=78

N=1605

l=12

L=73

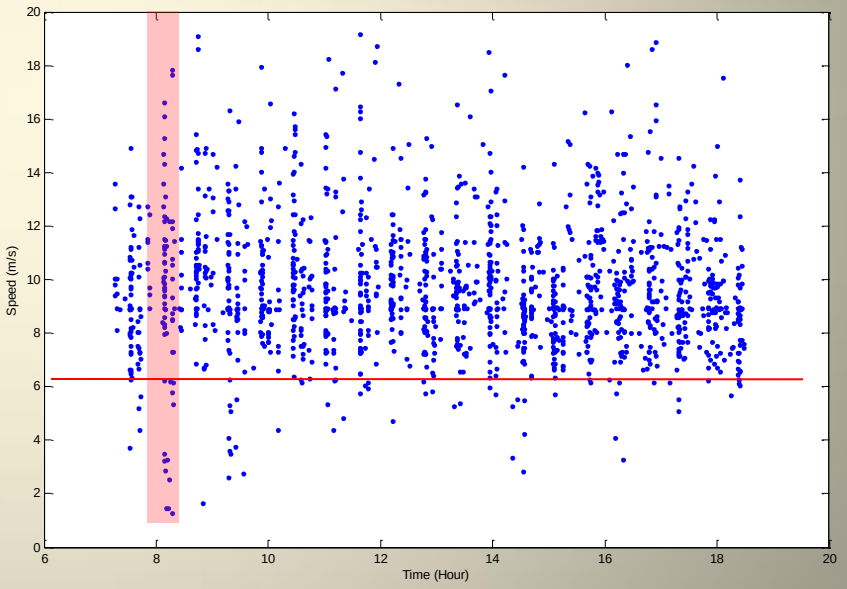
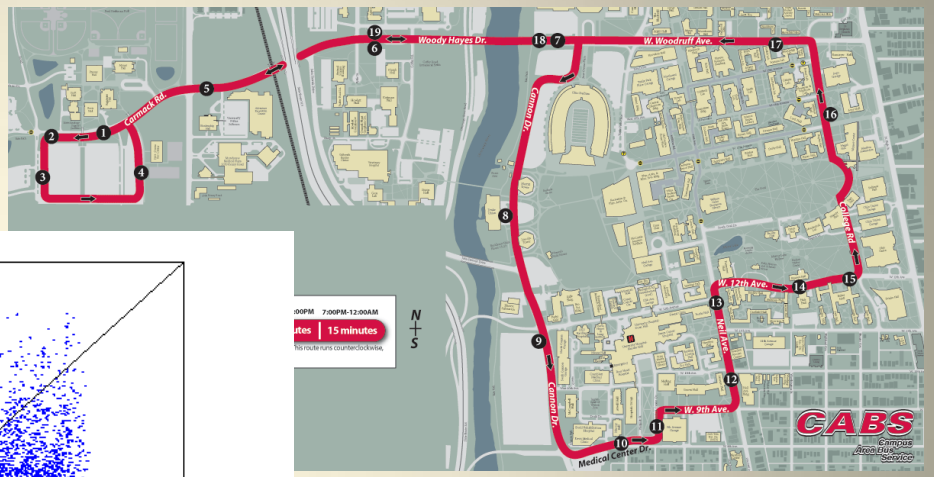


Illustration of Route-based Approach

Speed data by (x,t) collected over homogeneous days



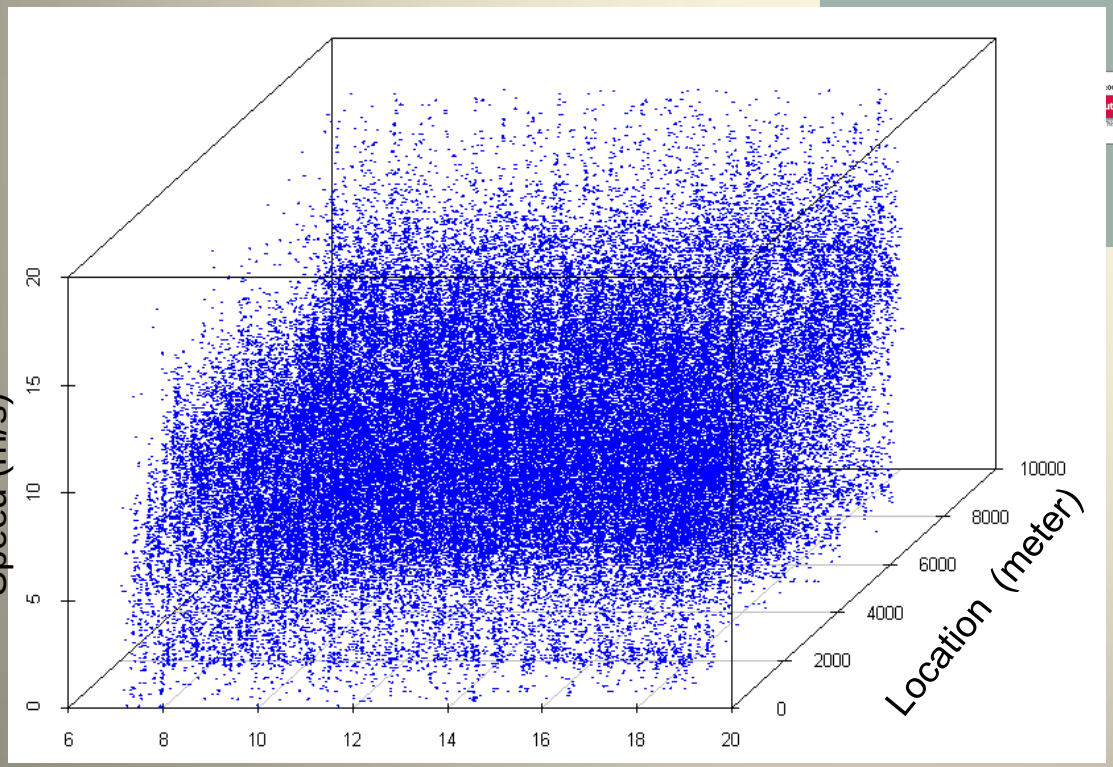
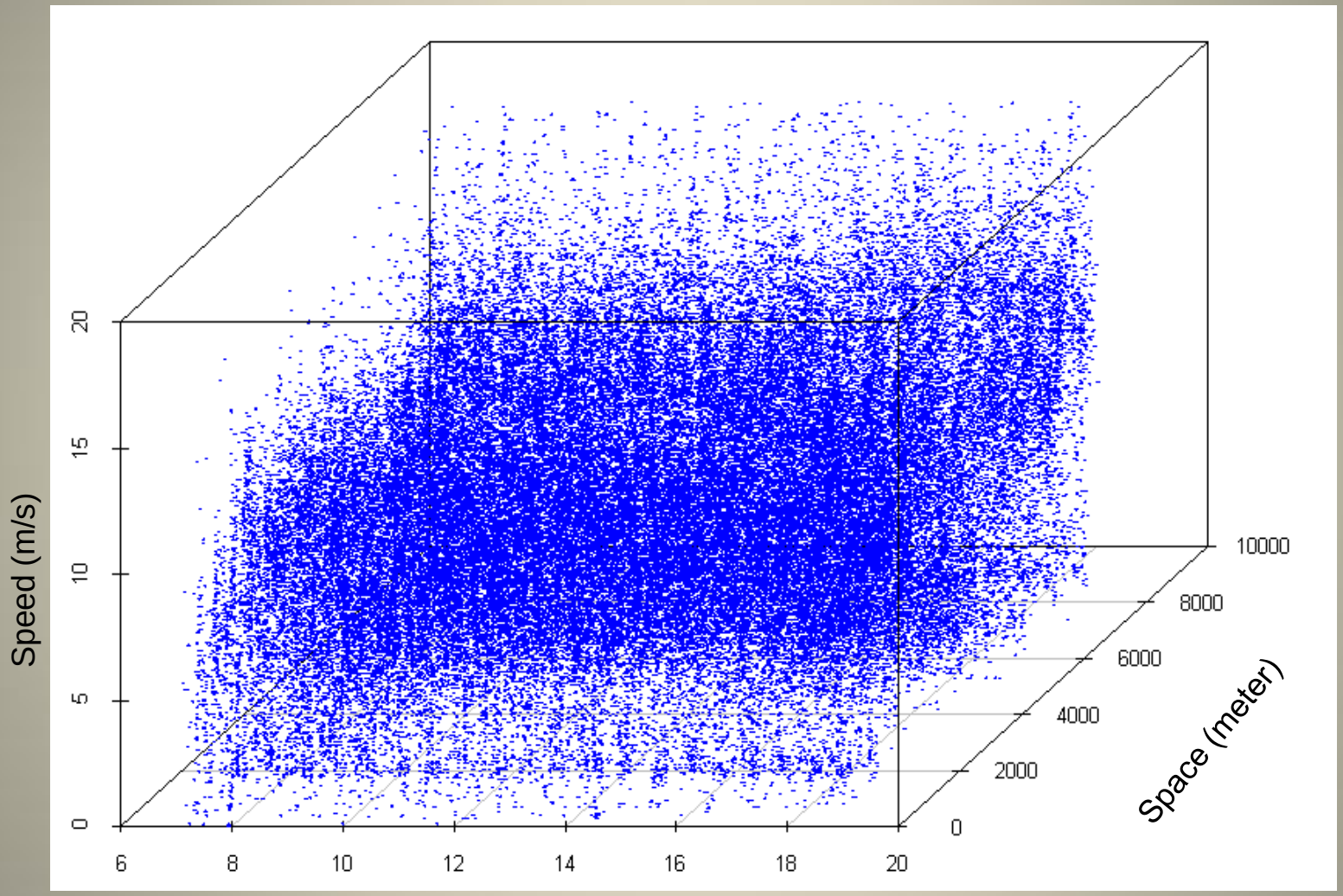
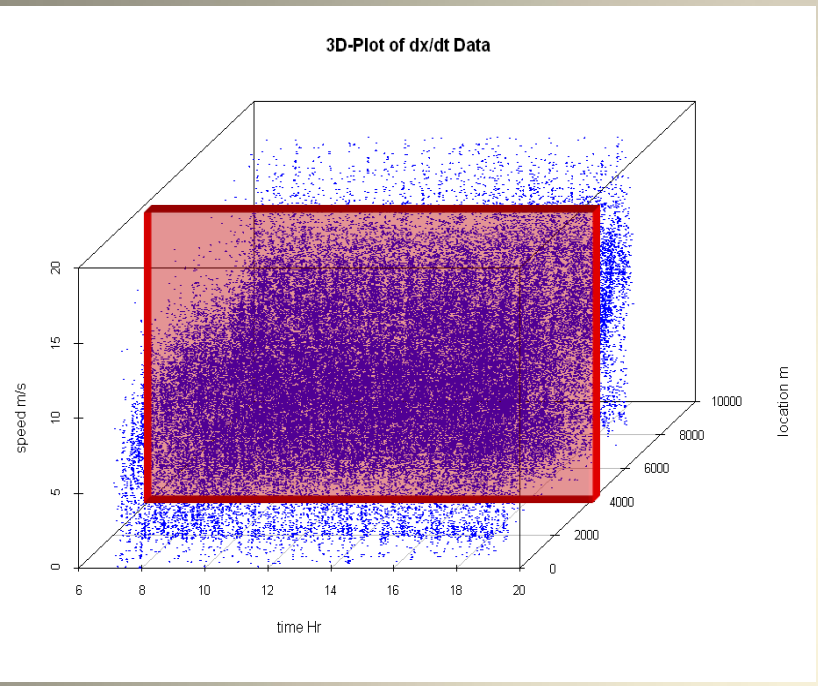
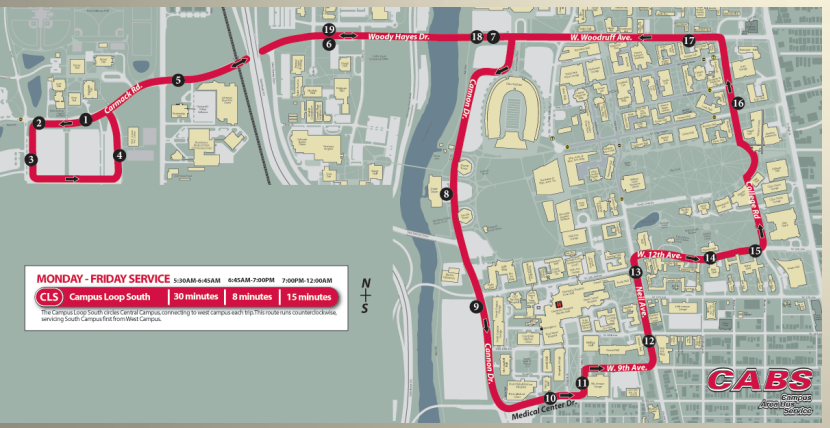


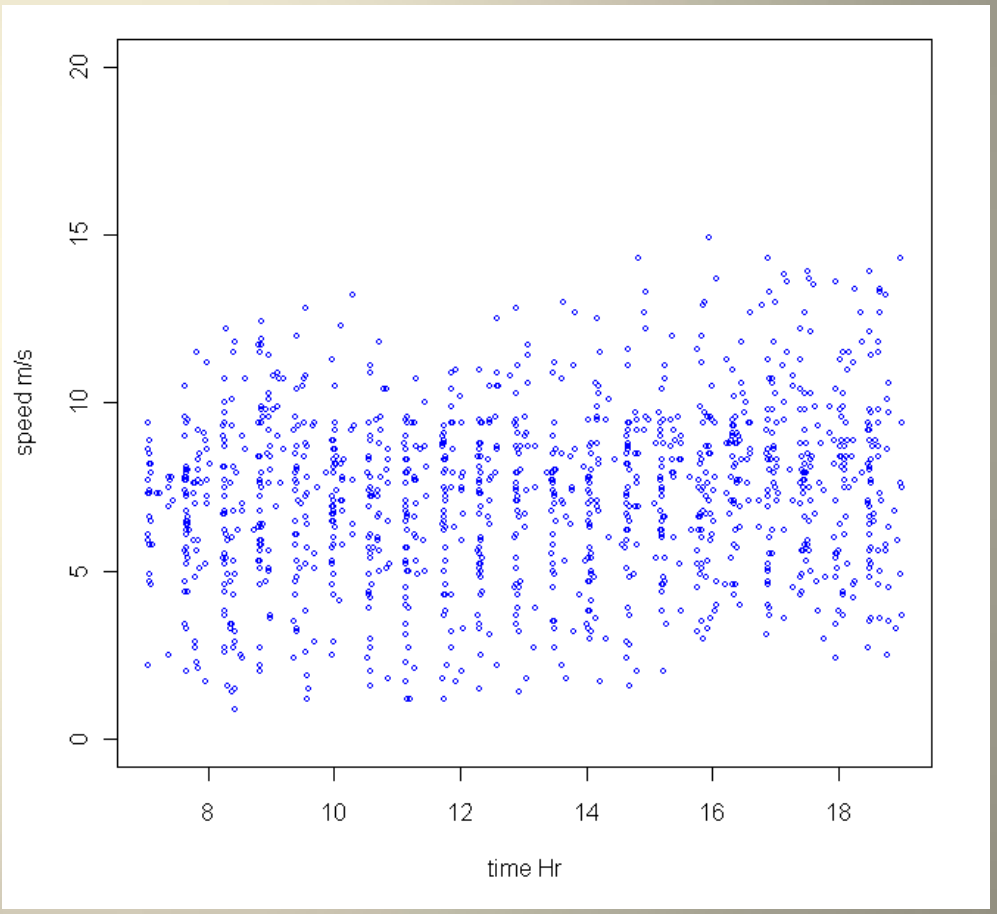
Illustration of Route-based Approach

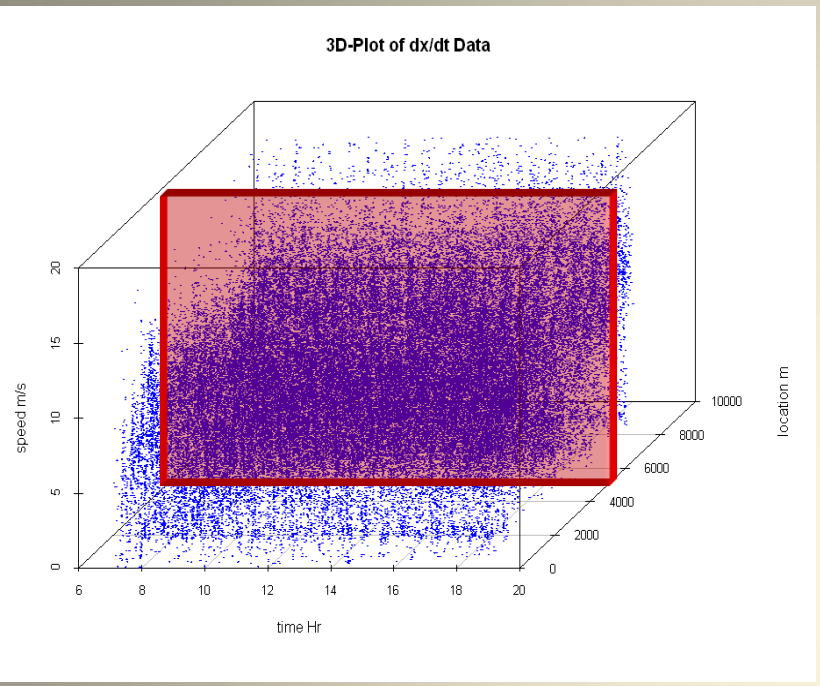




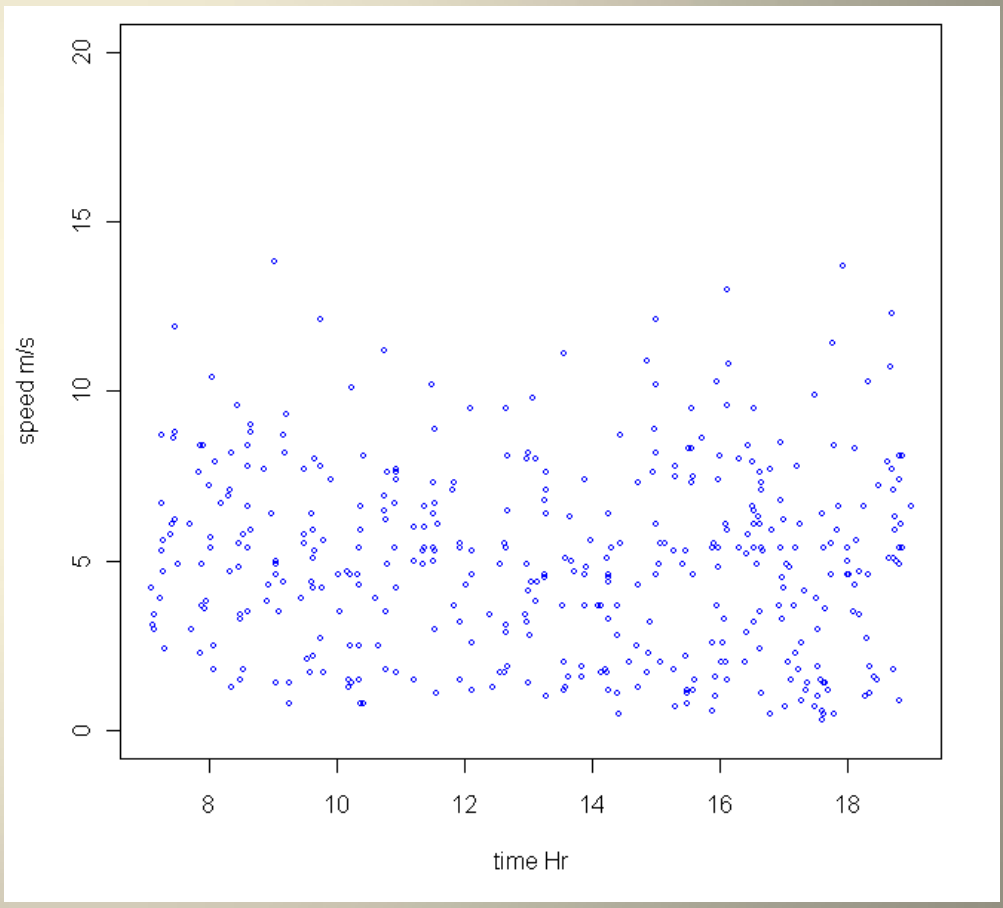


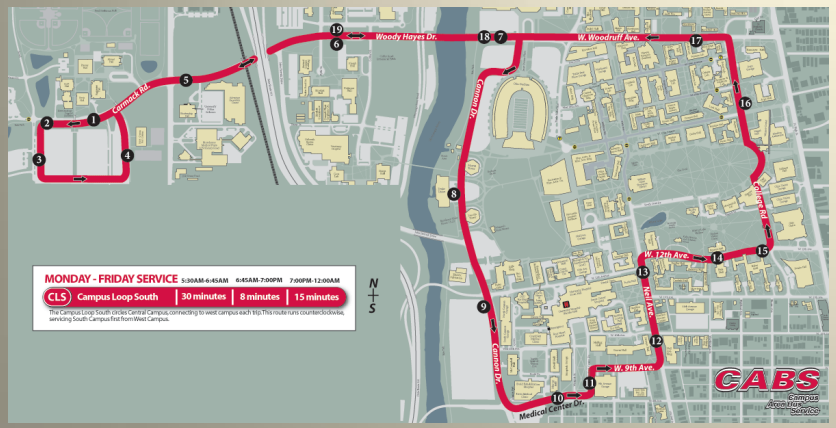
Location 4050 m

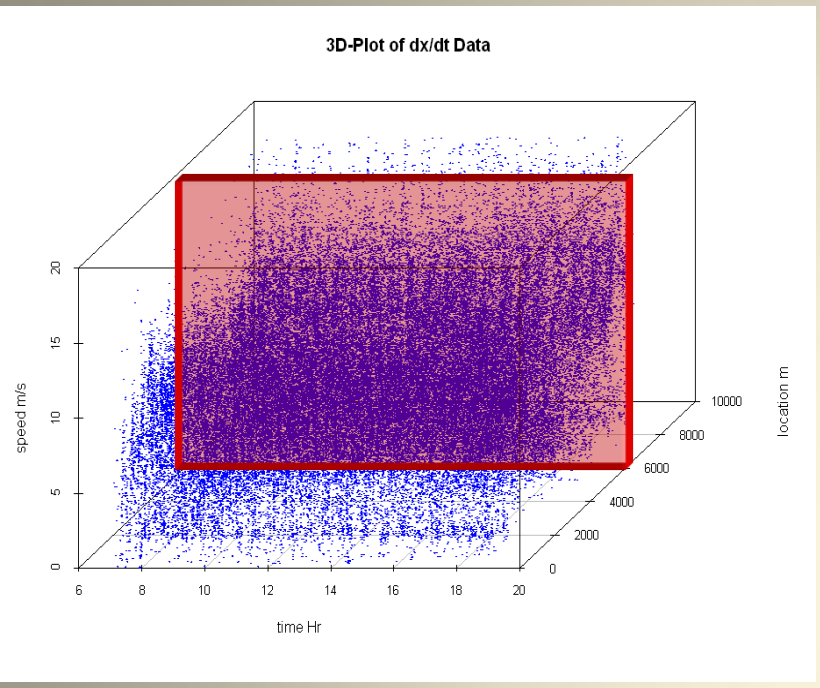


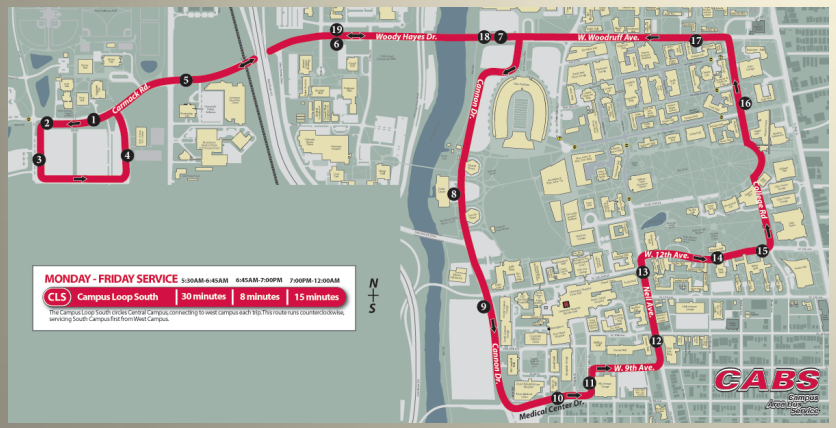


Location 5050 m

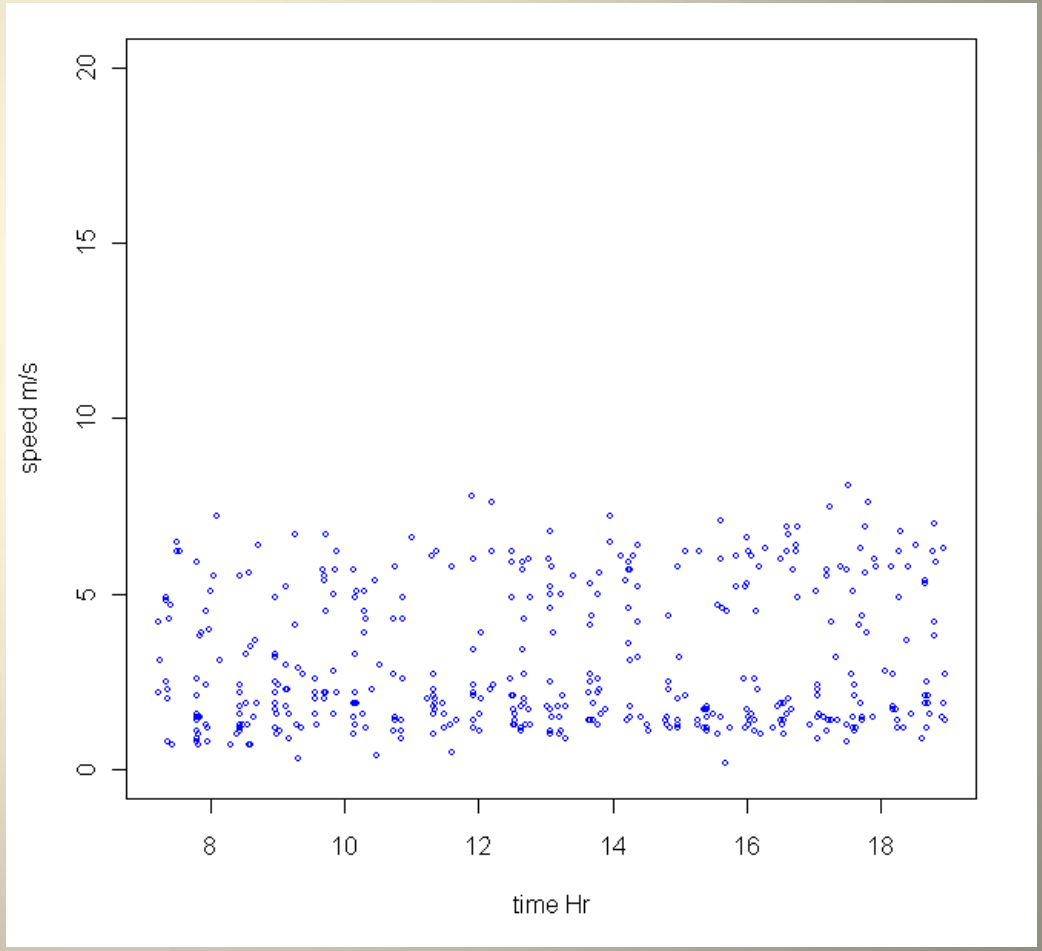






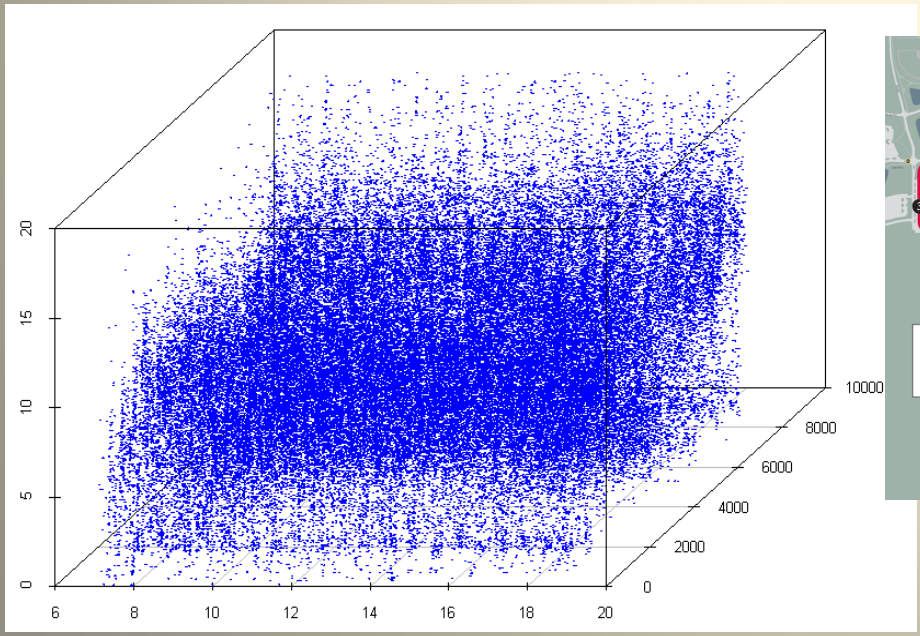


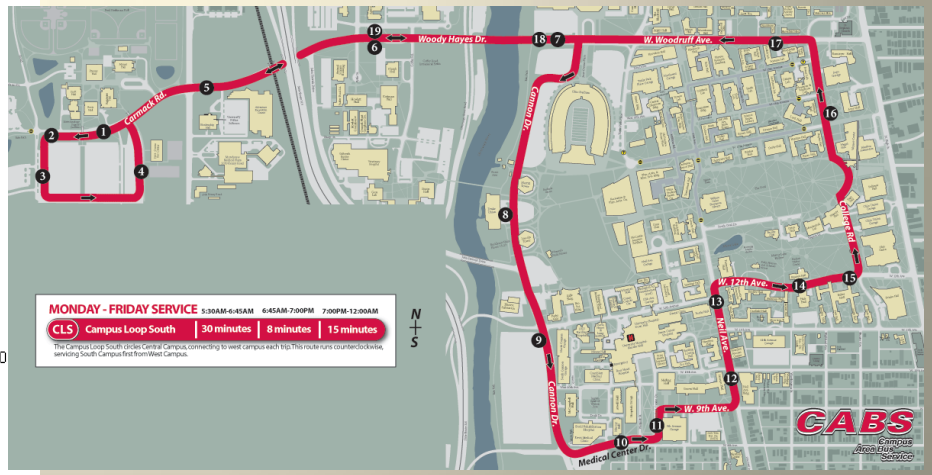
Location 6050 m



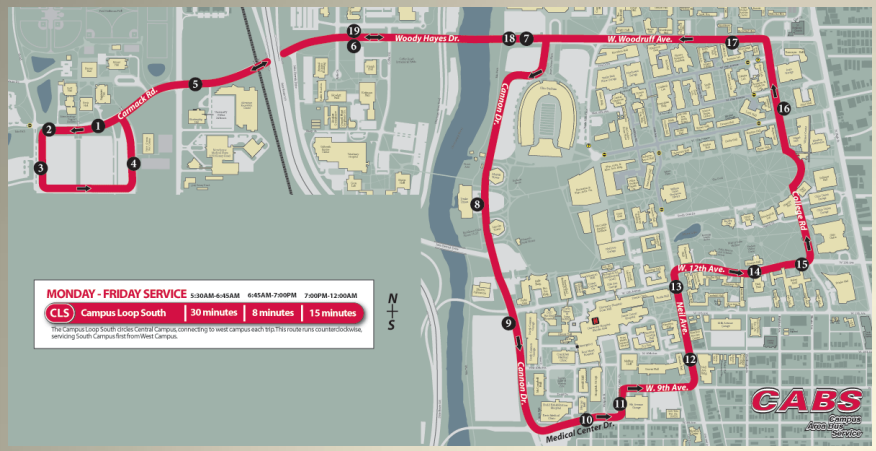
Route-based Approach

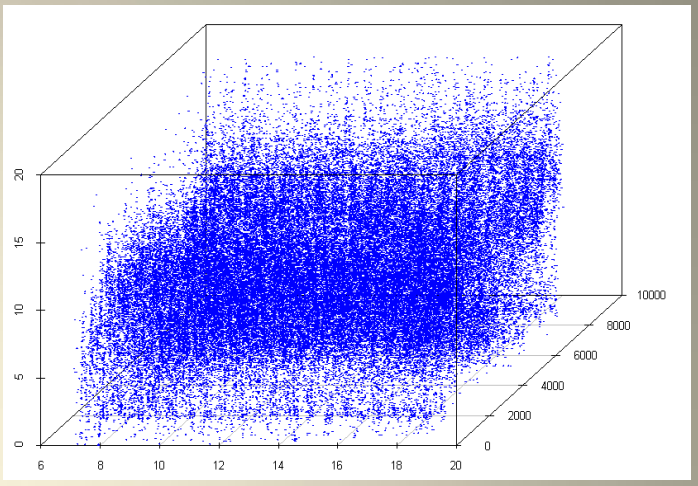
- Use speed data from multiple homogeneous days
- Start at location X = Xo near beginning of route
- Gather speed data in [X – INC, X + INC]
- Apply “low speed” method to identify time intervals indicating congestion
- Move to next location X+Δ; repeat until end of route

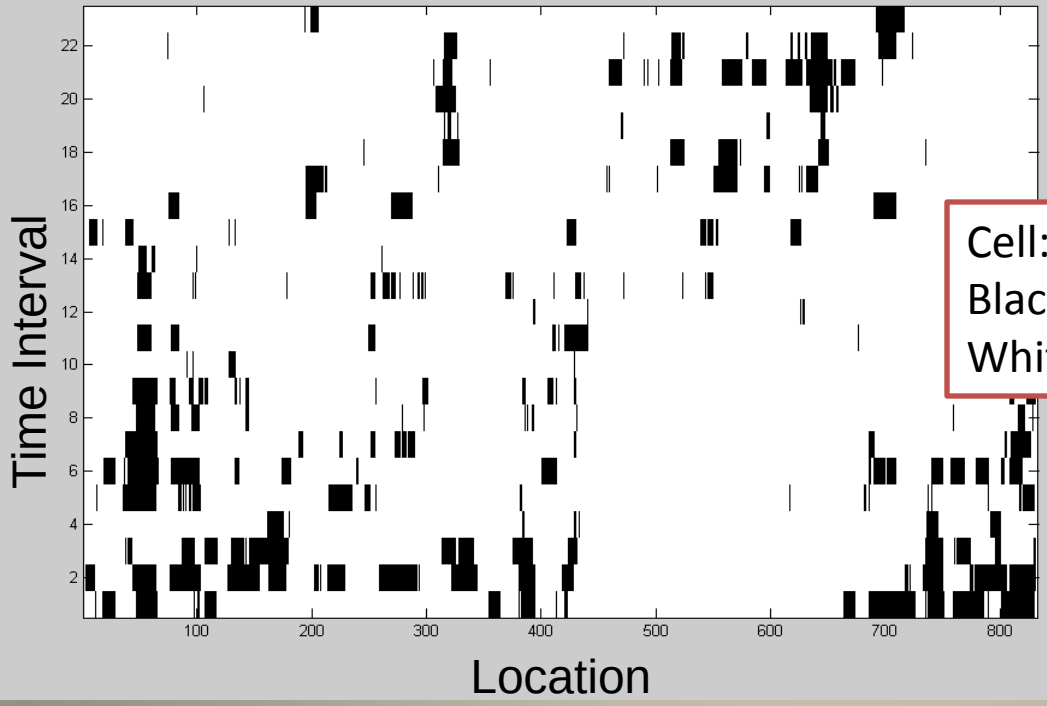




Location x Time Binary Matrix



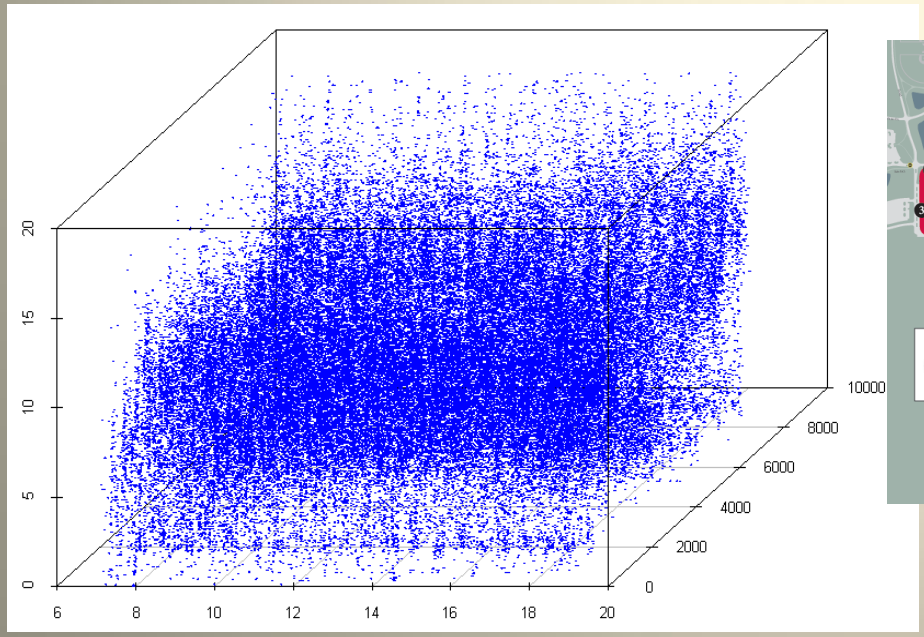


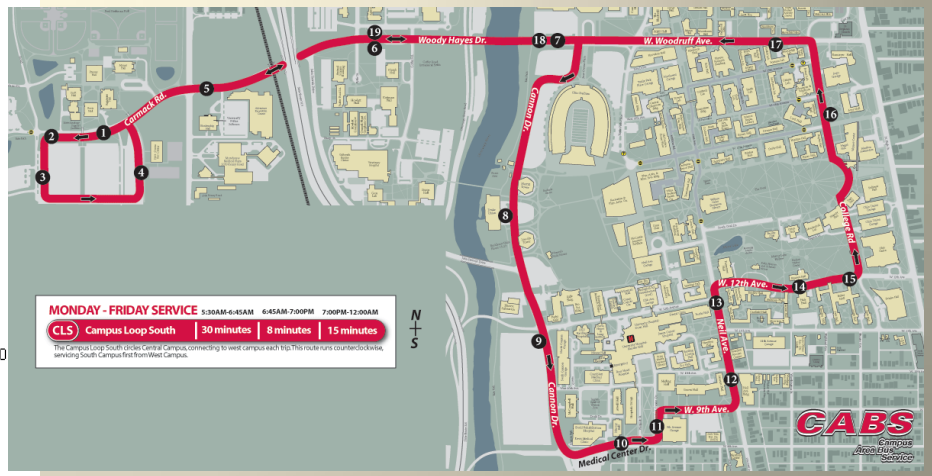


Cell: Location and time interval
Black: Test indicated congestion
White: Test does not indicate congestion

Empirical Study

- OSU CABS Campus Loop South route (8.3 km, 19 stops)
- 99,072 weekday speed observations April – June, 2004
- Arc speed calculated based on two consecutive space-time AVL points



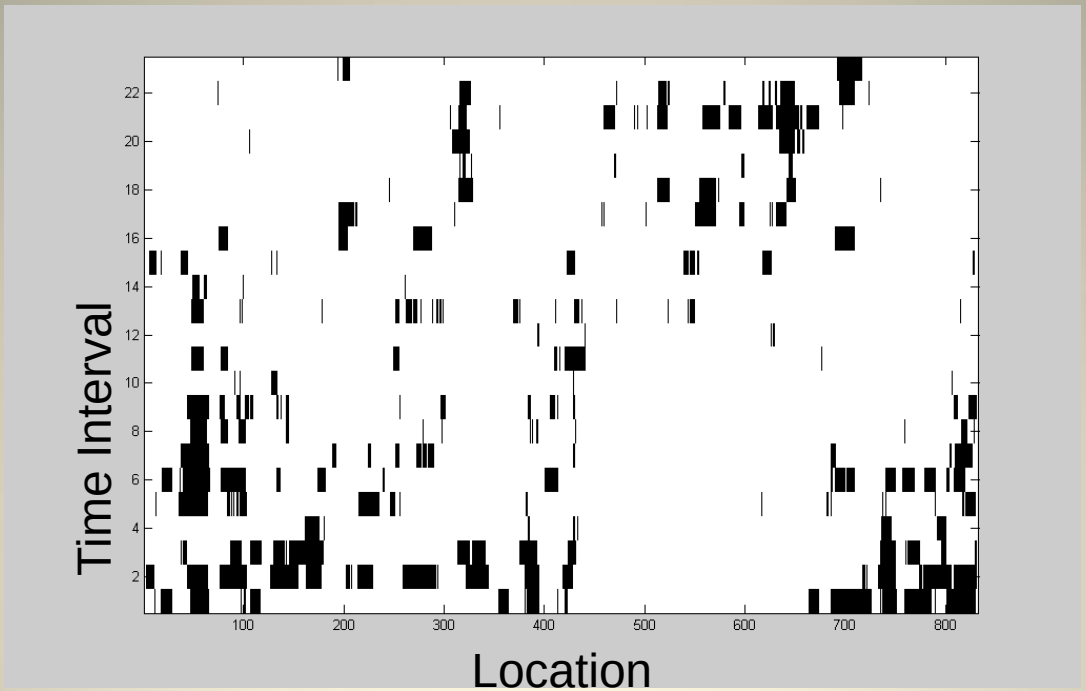


Empirical Study Parameters

- *Time intervals*: 23 non-overlapping, 30-min intervals from 7:00 am to 6:30 pm
- *Low speed threshold at X*: 0.35 x 95-percentile of speed distribution at X
- *Gather speed data in [X–INC, X+INC]*: INC = 50 m (100m space window to allow sufficient data)
- *Apply "low speed" method*: $\alpha = 0.05$
- *Move to X + Δ and repeat*: $\Delta = 10$ m (cover entire route; 833 spatial windows)

Binary location-time matrix: 833 x 23 cells

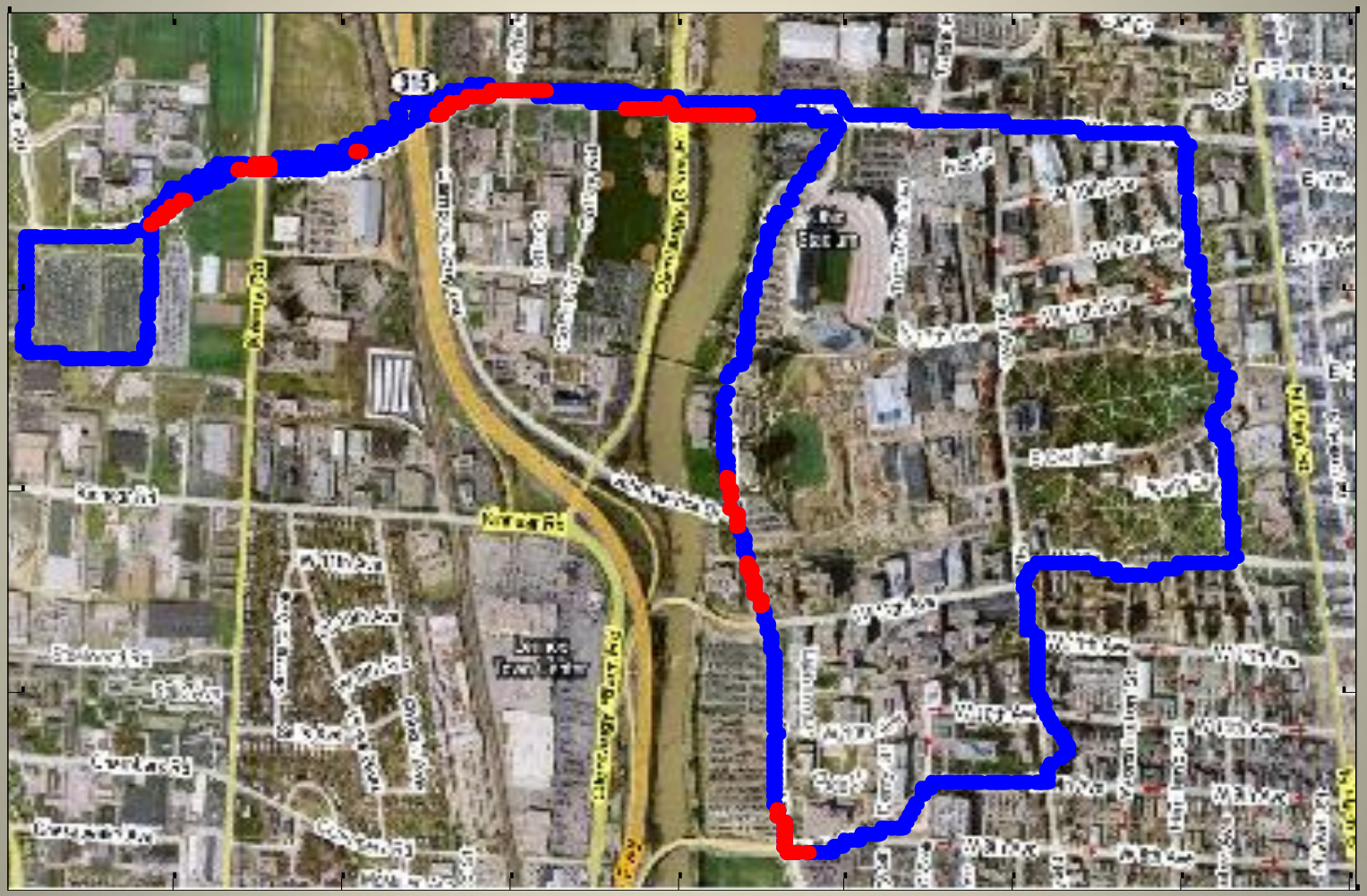
Spatial “Rule of 3 Filter”



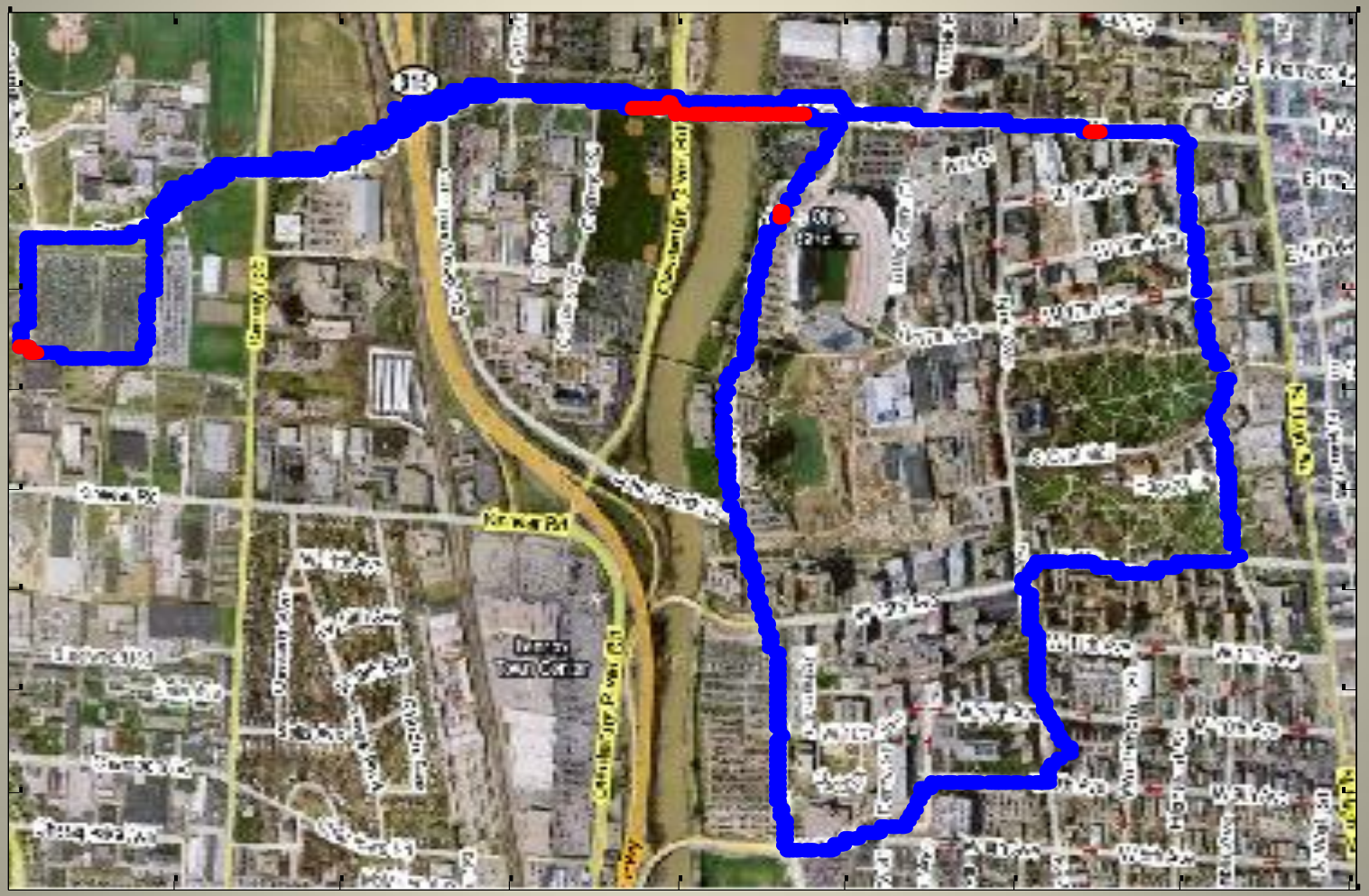
- Motivation:
  - Filter for “non-causal positives”
  - Acknowledges overlapping data and spatial extent of congestion
- 3 consecutive location cells must be “positive” to be considered as indicating congestion
- 3 cells “cover” 30 meters (includes 120m of data, with 80m common to all cells, 90 m common to 2 cells)

Empirical Route-wide Results

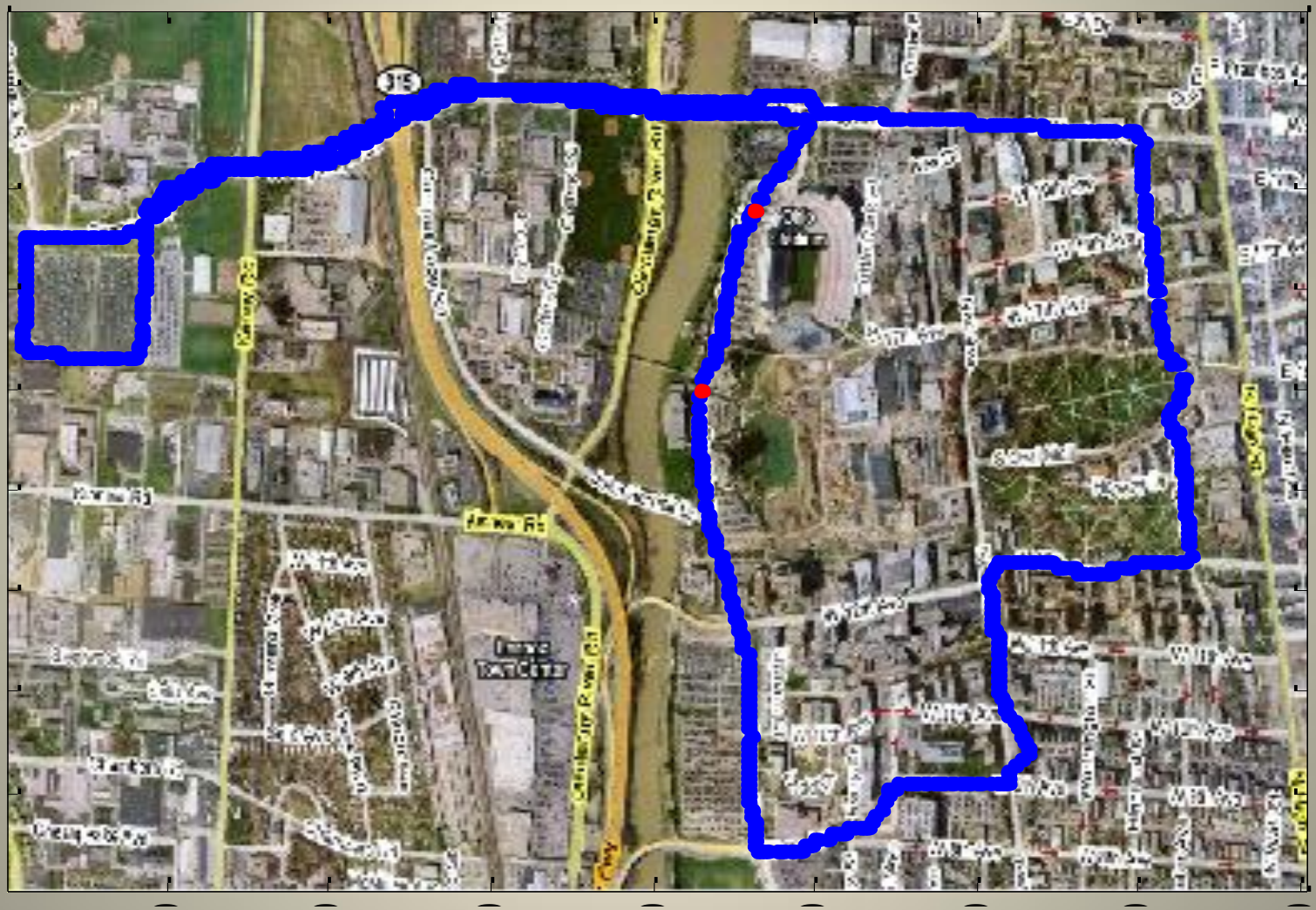
8:00-8:30 am



9:00-9:30 am



10:30-11:00 am

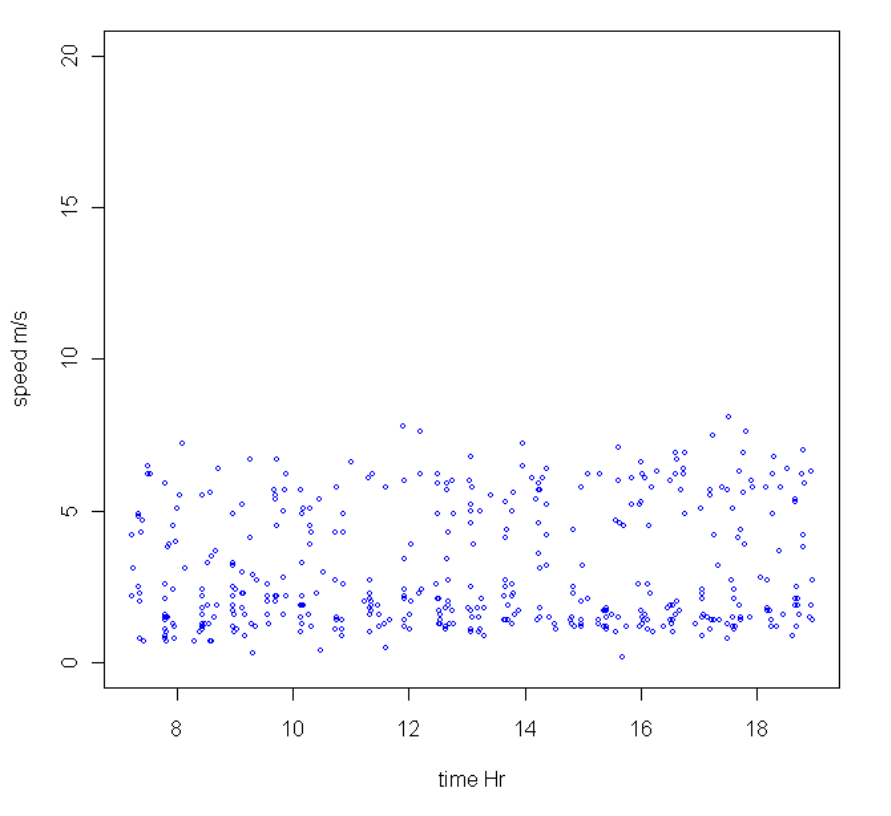


A Few Comments and Observations

Congestion Identification: Classical Statistical Methods

- Mean based approach (ANOVA-based test)
  - test if the mean speed in a certain time interval is lower than the others
  - assume normality
- Median based approach (Rank sum test)
  - test if the median speed in a certain time interval is lower than the others
  - assume same shape distribution

Deficiency of Classical Statistical Methods



Necessary assumptions violated (normality violated in over 70% of cells)

High variance because of bi-modality

Ongoing/Future Work

- Systematic investigation of reasonableness of indicated results
- Contribution of/Integration with other (classical) methods
- Impacts of
  - space window and time interval size
  - overlapping space window
  - overlapping time intervals
  - “rule of 3”

Ongoing/Future Work

- Use of arc speeds vs. instantaneous GPS speed
- Impact of data frequency
- Investigations with OSU CABS smart bus system
  - Replications
  - Investigations of assumptions
- Validate with other transit systems