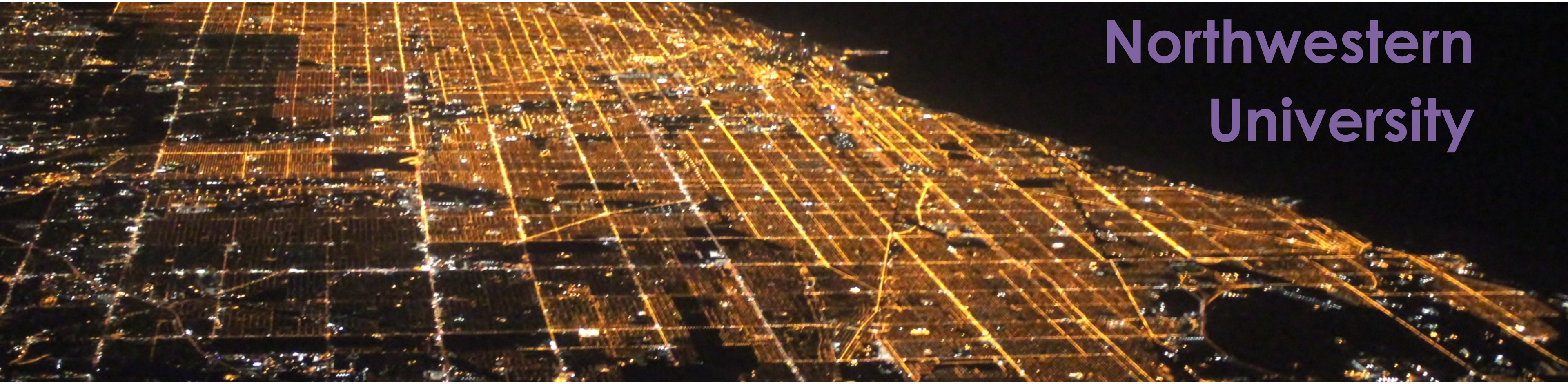
McCormick

Northwestern Engineering

Northwestern University Transportation Center

TRAJECTORIES IN 3D:
UNIFYING MODEL CALIBRATION AND NETWORK PERFORMANCE ANALYSIS

Hani S. Mahmassani



Symposium Celebrating 50 Years of Traffic Flow Theory
Portland, Oregon (USA), August 11-13, 2014

Contributors:

Jiwon Kim *(now at U. of Queensland, Brisbane, AU)*
Tian Hou
Meead Saberi *(now at Monash Univ., Melbourne, AU)*
Ali Zockaie

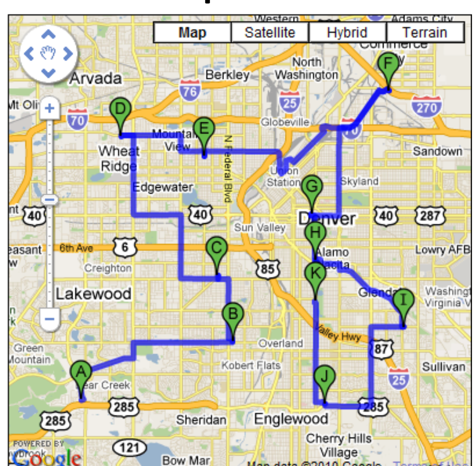
Doctoral researchers at NUTC

Outline

- **Motivation**
- **From 2D trajectories to 3D trajectories**
- **Application to Network Flow Modeling (NFD's)**
  - Vehicular networks
  - Pedestrians and crowds
- **Travel time reliability**
  - Signature relations and trajectory data
  - Within-day and day-to-day variability
- **Scenario-based approach to reliability modeling**
  - Trajectory Processor for particle-based simulators
- **Takeways, Limitations and Challenges**

Motivation

- **NETWORK TRAFFIC FLOW MODELING** needs high quality traffic data with broad network level coverage, for calibration, validation, and input to real-time predictive management strategies.
- **CHARACTERIZATION OF NETWORK PERFORMANCE,** and the quality of service experienced by users increasingly encompasses broader array of dimensions– e.g. reliability – that call for tracking vehicles as they travel through the network, and not only as they pass selected points
- Development of telecommunication and wireless technology are augmenting conventional point-based data collection methods with low-cost and widely available probe data.



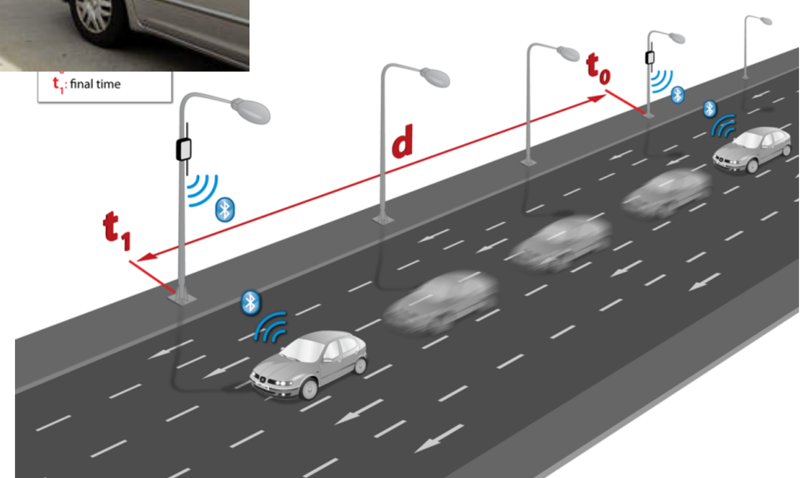
Motivation II

- LOOKING AHEAD–

  Autonomous Vehicles and *Connected Vehicles/Systems* will play a growing role part within the advanced traffic data environment, both as a major generator (data source) as well as end-user.

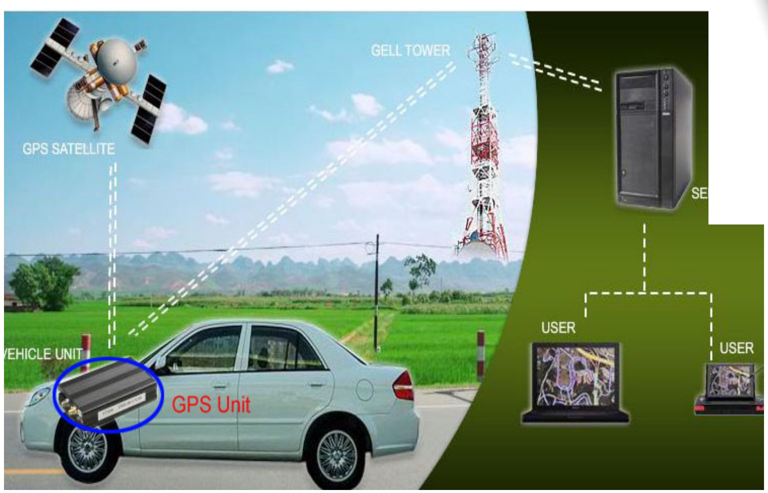
Segment Traffic Data

- Segment Data
  - collected by electronic transponders
  - Automatic Vehicle Identification (AVI), electronic toll data (I-PASS), blue tooth data, etc



Trajectory Data

- Collected by probe vehicles equipped with on board GPS devices
- A trajectory is the path followed by the moving object through the spatial area over which it moves



Trajectory Data

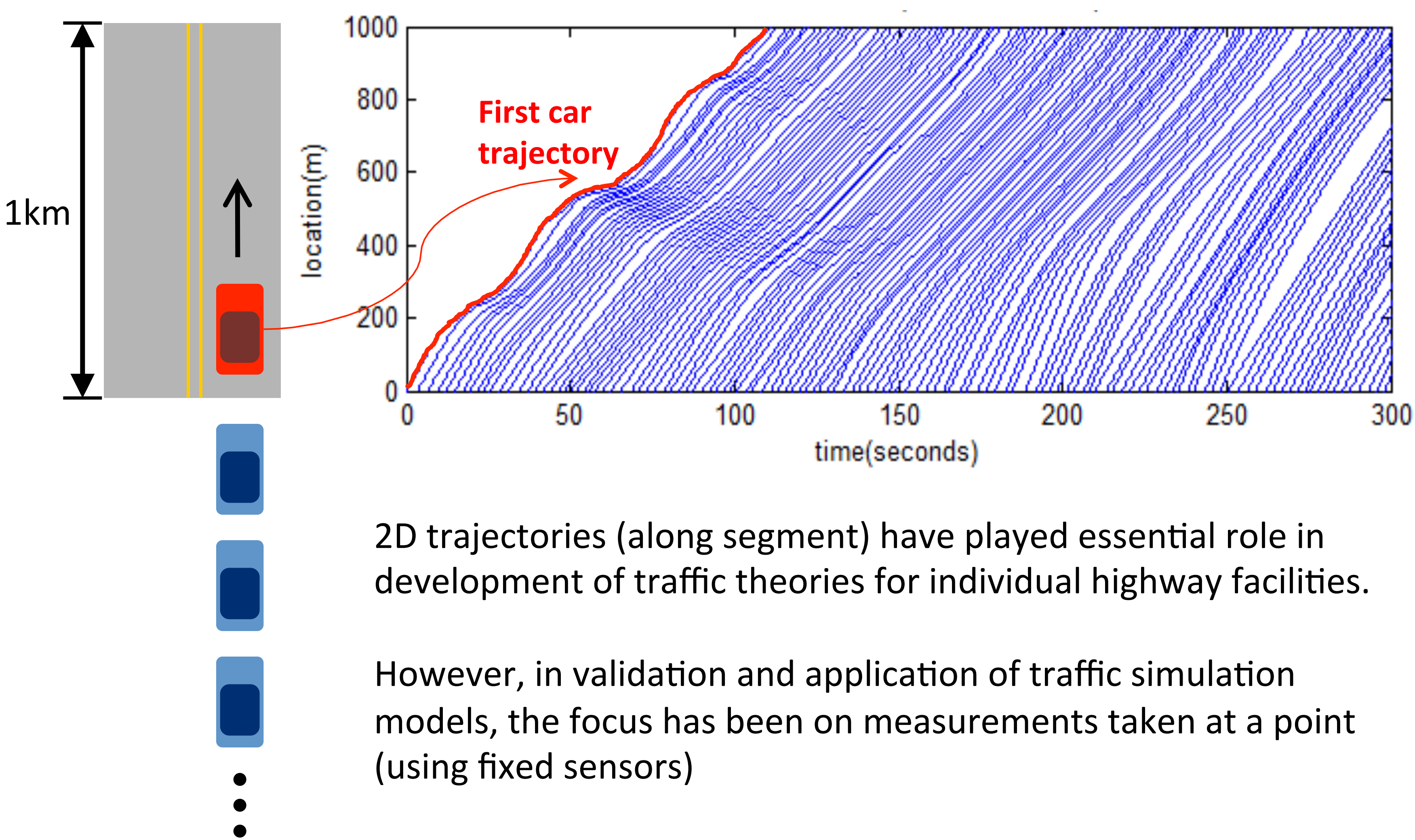
- Information that can be extracted from trajectory data
  - from individual trajectory:
    - Time, i.e. position of this moment on the timescale;
    - Position of the vehicle in space;
    - Trip origins and destinations ;
    - Direction of the vehicle's movement;
    - Speed of the movement;
    - Dynamics of the speed (acceleration/deceleration);
    - Accumulated travel time and distance.
    - Individual path and temporal characteristics
  - from groups of trajectories:
    - Distribution of speed/travel time;
    - Probe vehicle density;
    - Inferred traffic volume.

Trajectory Data

- Advantages and limitations of trajectory data as compared to traditional traffic data

Advantages	Limitations
• Low or no cost in installation and maintenance; • Wider geographic coverage (freeways and arterials); • Finer resolution (individual vehicle and shorter measurement time interval); • Contains additional traffic information (e.g. travel time); • Not affected by traffic interruptions or bad weather conditions. • Traffic simulation tools (microscopic, mesoscopic, or "particle-based" simulators) naturally produce trajectories	• Technology is not as mature as fixed sensors; • No direct occupancy or traffic density information; • Limited experience in analyzing data.

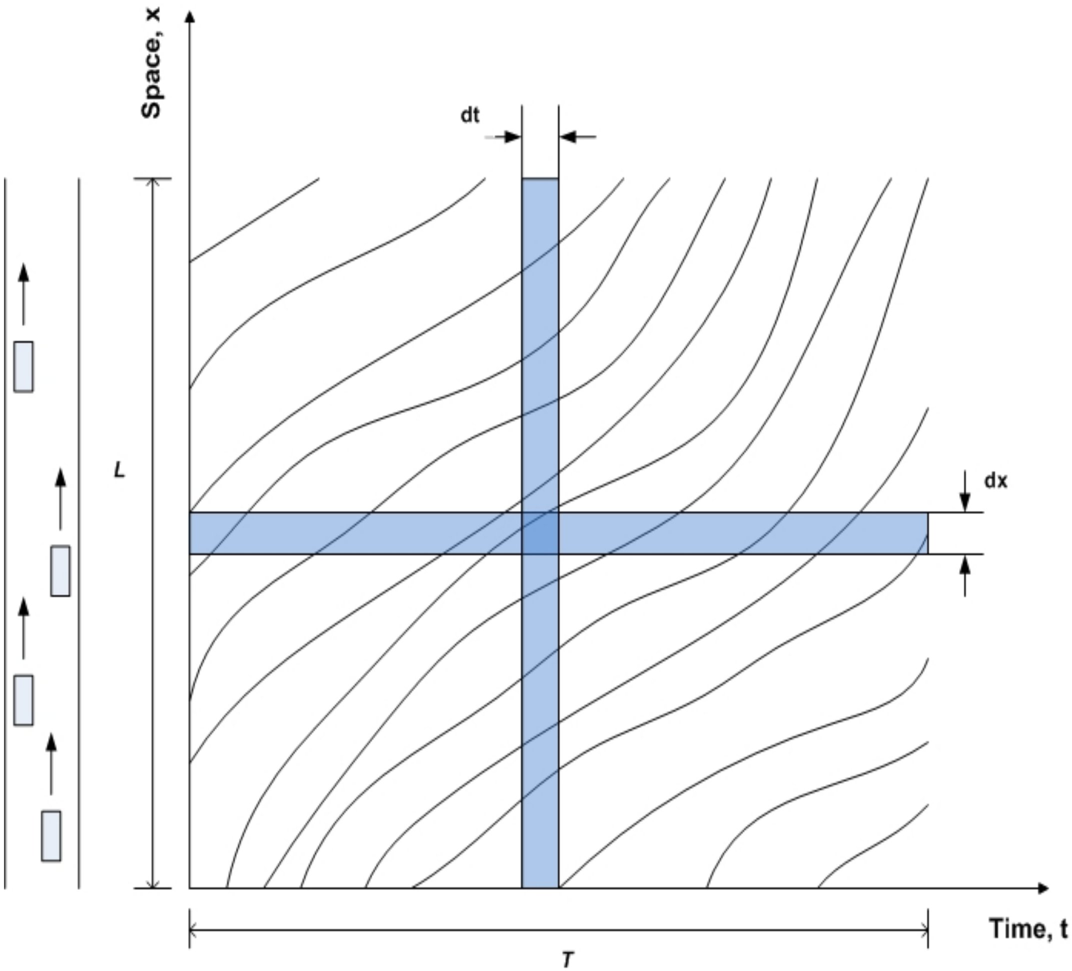
2D Trajectories



2D trajectories (along segment) have played essential role in development of traffic theories for individual highway facilities.

However, in validation and application of traffic simulation models, the focus has been on measurements taken at a point (using fixed sensors)

Measurements from Multiple Trajectories along a Single Road Segment

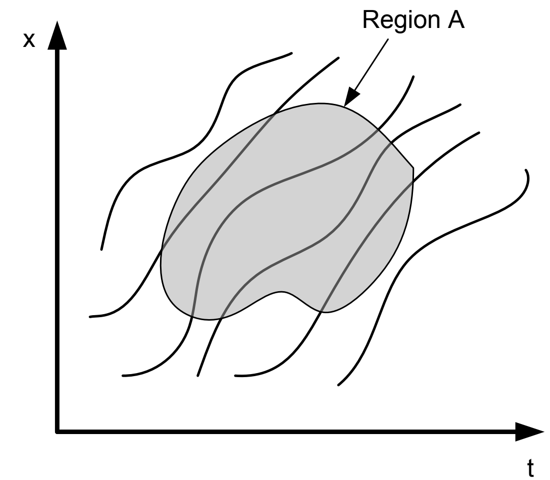


Edie's Definitions

Highway traffic (unidirectional)

Trajectory-based definitions of network flow variables

$$q(A) = \frac{\sum_{n \in N} d_n}{|A|} \qquad k(A) = \frac{\sum_{n \in N} \tau_n}{|A|}$$



where d_n is the total distance traveled by vehicle *n* in region *A*, τ_n is the total time spent by vehicle *n* in region *A*, and $|A|$ is the area covered by region *A.*

Edie (1965)

Network Fundamental Diagram

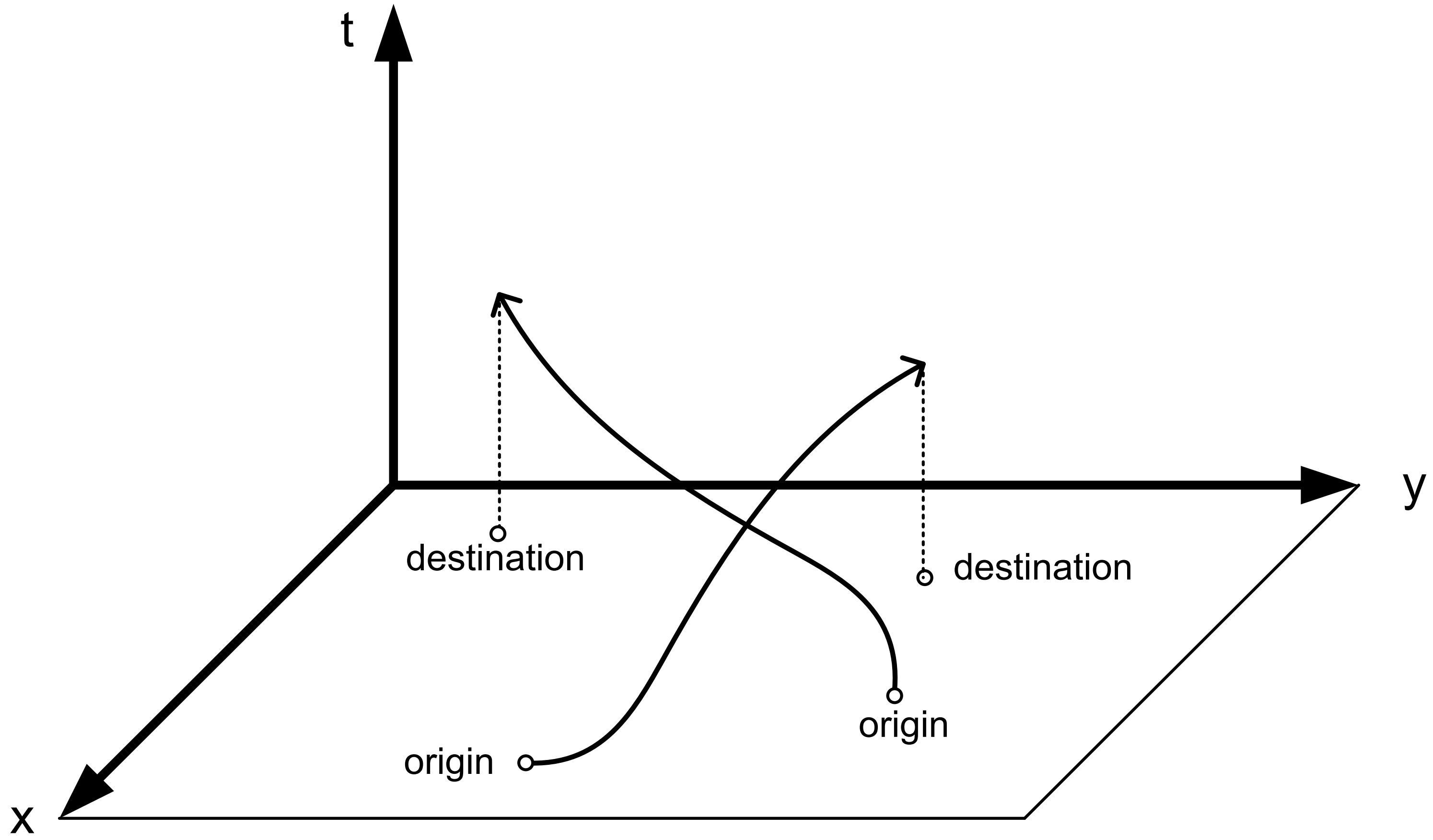
Trajectory-based definitions

It is possible to extend Edie's well-known generalized variable definitions of vehicle traffic flow along a highway to a network, as recently recognized by *Courbon and Leclercq (2011)*.

Recently, in *Saberi, Mahmassani and Zockaie (2014)*

- Operationalize and validate the extension of Edie's definitions to the network level.
- Formalize and test a method using three-dimensional (3D) vehicle trajectories in time and space to estimate network flow, density, and speed.

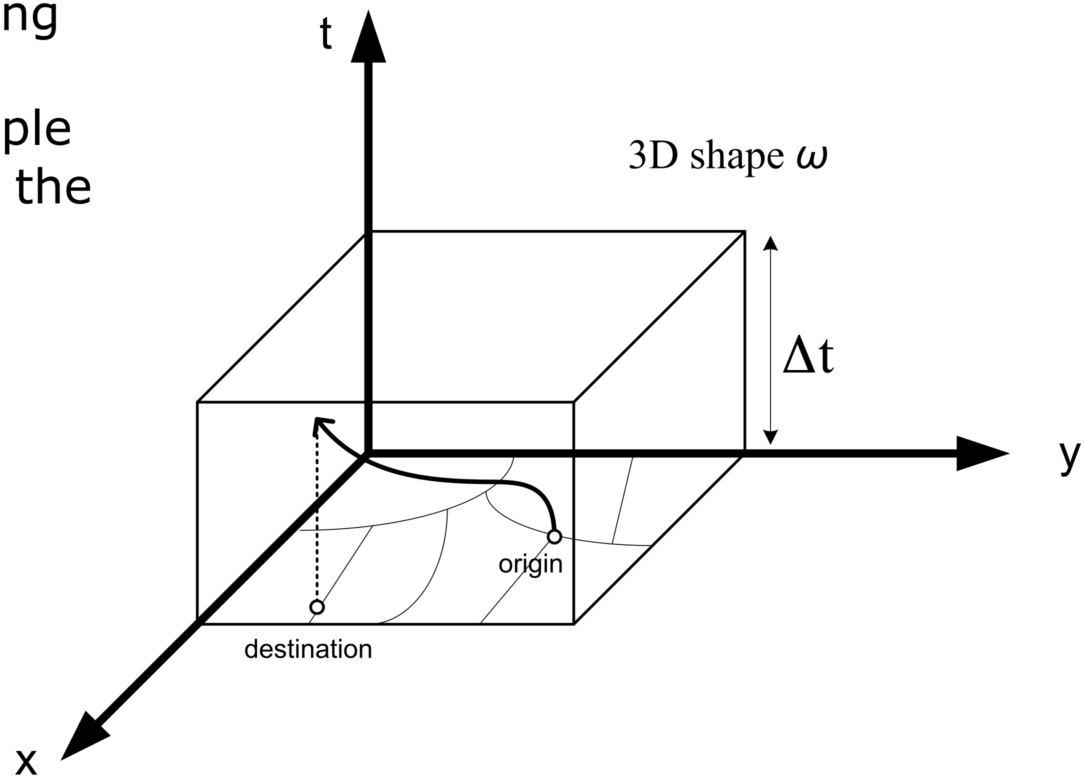
3D Trajectories in a Network



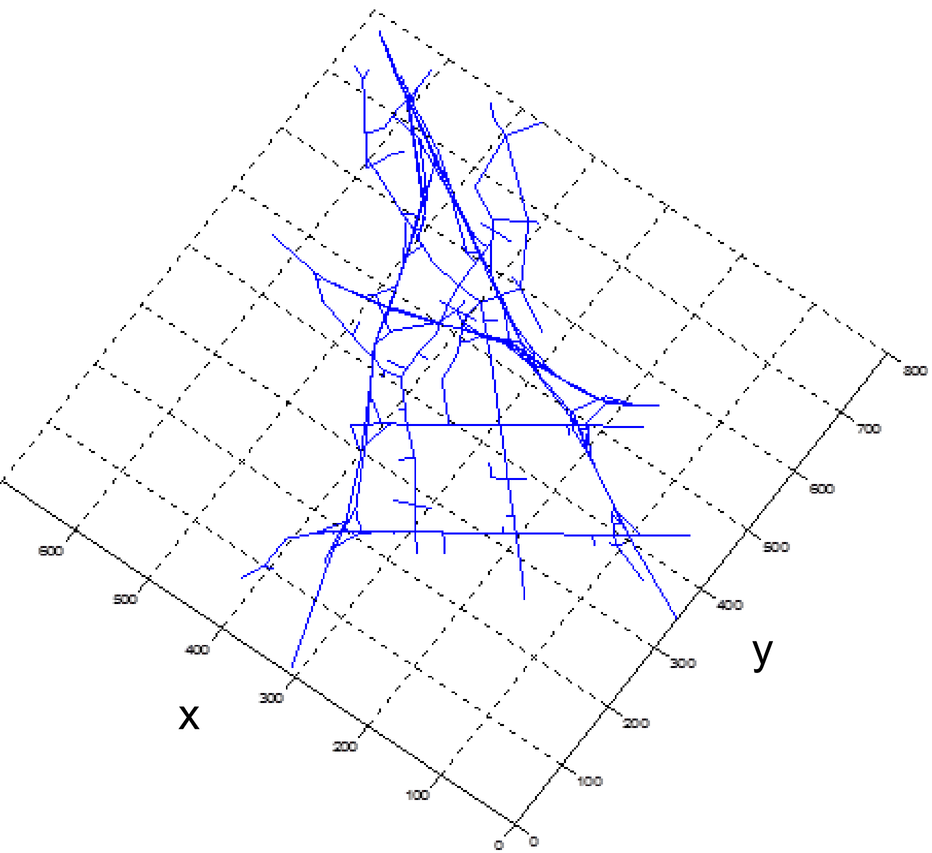
Network 3D Time-Space Diagram

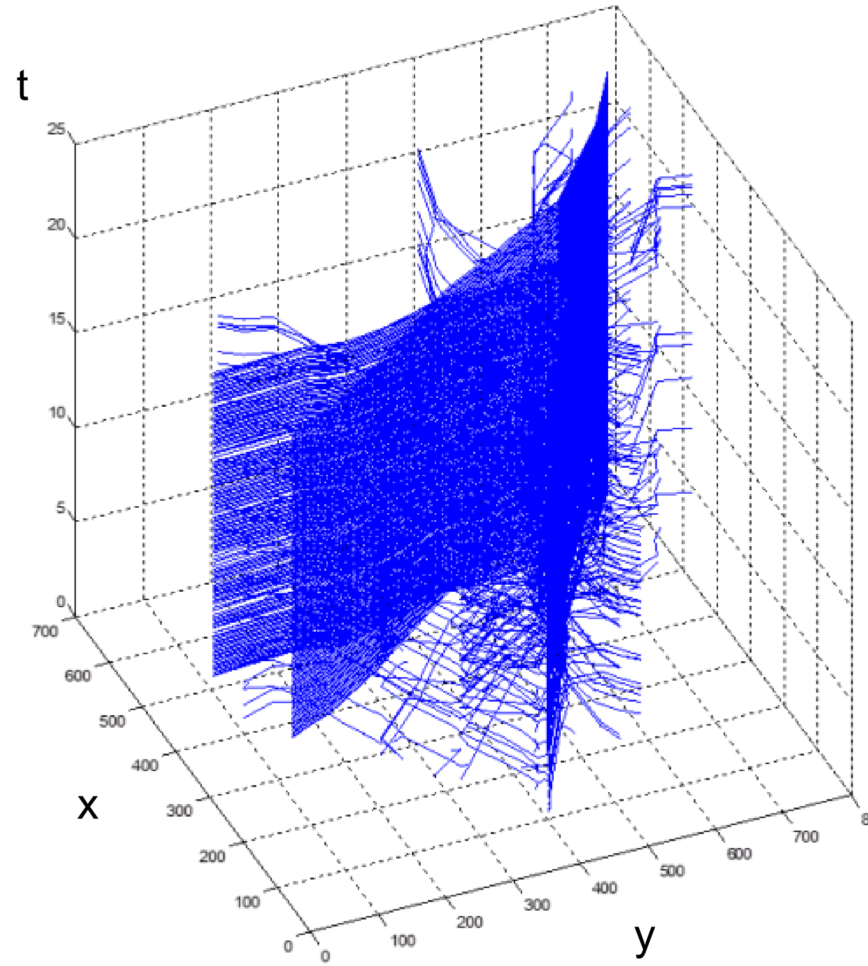
In order to estimate network-wide traffic flow variables using trajectories, we introduce a closed 3D shape ω, for example a cube, similar to region A in the 2D time-space diagram.

The network structure is laid down on the *x*-*y* plane.



Network 3D Time-Space Diagram





3D trajectories of 1,000 simulated vehicles in Irvine, California

Saberi, Mahmassani, Zockaie (2014)

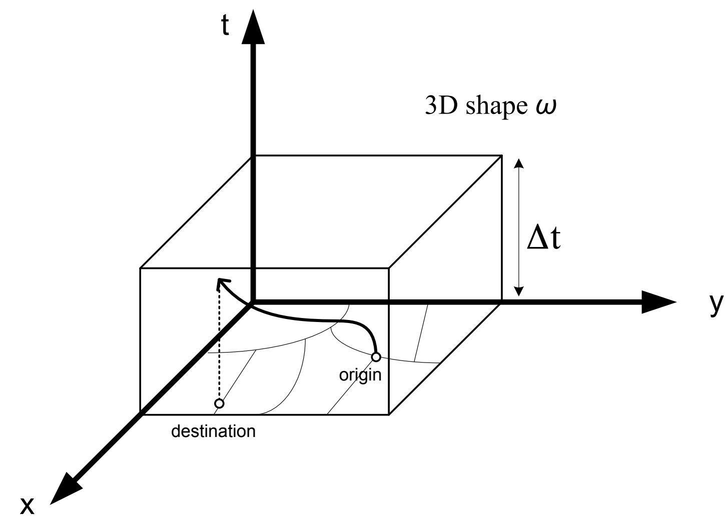
Edie's Definitions

Extension to Networks

Courbon and Leclercq (2011)
Saberi, Mahmassani, Zockaie (2014)

$$Q(\omega) = \frac{d(\omega)}{L_{xy}(\omega) \times \Delta t}$$

$$K(\omega) = \frac{t(\omega)}{L_{xy}(\omega) \times \Delta t}$$



where $Q(\omega)$ and $K(\omega)$ are the network-wide average flow and density for the specified shape ω; $d(\omega)$ is the total distance traveled by all the vehicles in the shape ω, $t(\omega)$ is the total time spent by all vehicles in the shape ω, $L_{xy}(\omega)$ is the total length (in lane-miles or lane-kms) of the network on the x-y plane associated with the shape ω, and Δt is the time height of the shape ω.

Background

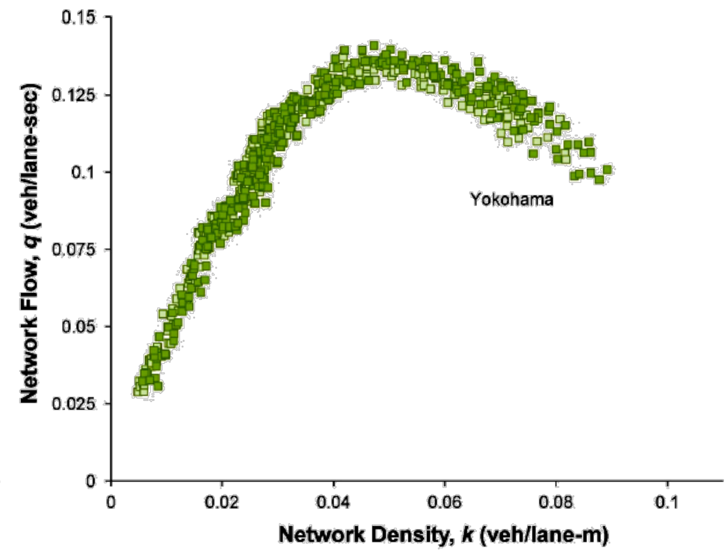
Network Fundamental Diagram

Link-based definitions

Most of the studies to date have used the classical link-based measurement method to estimate the NFD by taking the distance-weighted averages of flow and density over all the links in the network.

$$Q = \frac{\left(\sum_{i=1}^{M} l_i q_i \right)}{\left(\sum_{i=1}^{M} l_i \right)} \qquad K = \frac{\left(\sum_{i=1}^{M} l_i k_i \right)}{\left(\sum_{i=1}^{M} l_i \right)}$$

Source: *Mahmassani et al. (1984)*



Source: *Geroliminis and Daganzo (2008)*

Network Fundamental Diagram

Trajectory-based definitions

It is possible to extend Edie's well-known generalized variable definitions of vehicle traffic flow along a highway to a network, as recently recognized by *Courbon and Leclercq (2011)*:

$$\tilde{q}(t \rightarrow t + \Delta t, x \rightarrow x + \Delta x) = \frac{\sum_i l_i}{\Delta t \Delta x}$$

$$\tilde{k}(t \rightarrow t + \Delta t, x \rightarrow x + \Delta x) = \frac{\sum_i t_i}{\Delta t \Delta x}$$

where l_i and t_i are respectively the distance traveled and the time spent by vehicle *i* in a time-space area of $\Delta x.\Delta t$.

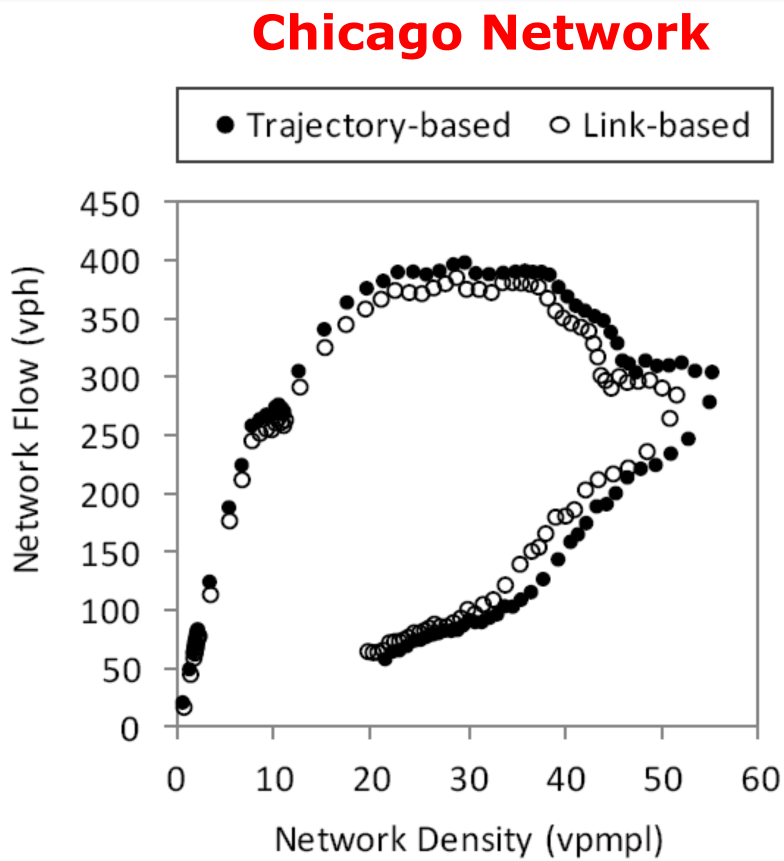
Network Traffic Simulation

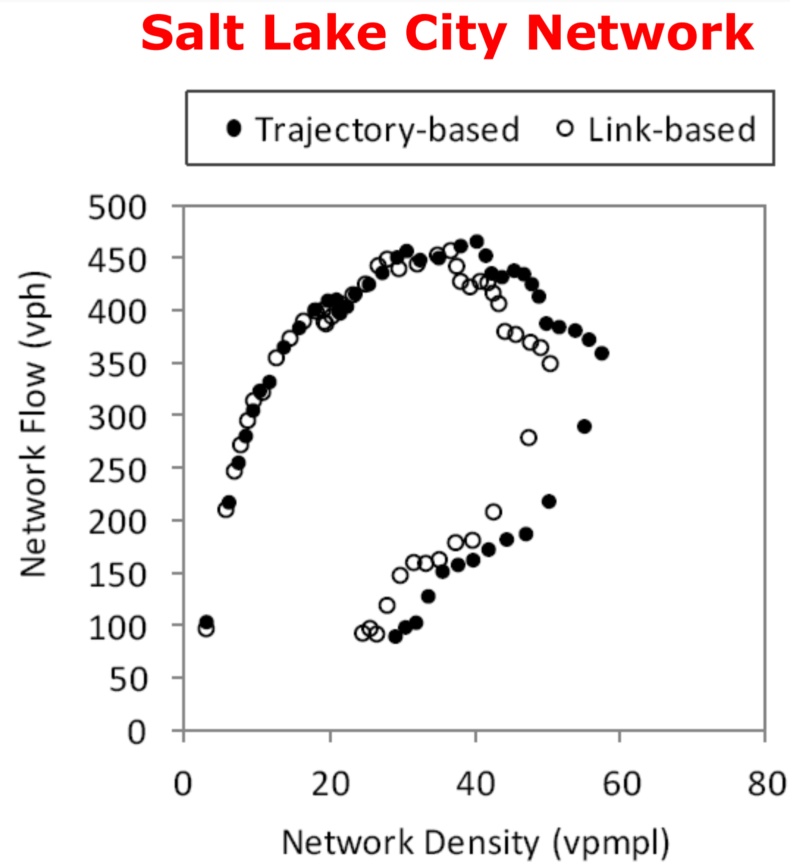
Networks of Chicago and Salt Lake City are simulated in a simulation-based dynamic traffic assignment platform (DYNASMART-P) with normal daily demand and 20% adaptive drivers to prevent formation of large gridlock.

Smaller Chicago sub-network

Salt Lake City sub-network

Trajectory vs. Link based NFD





NFDs with Different Measurement Methods

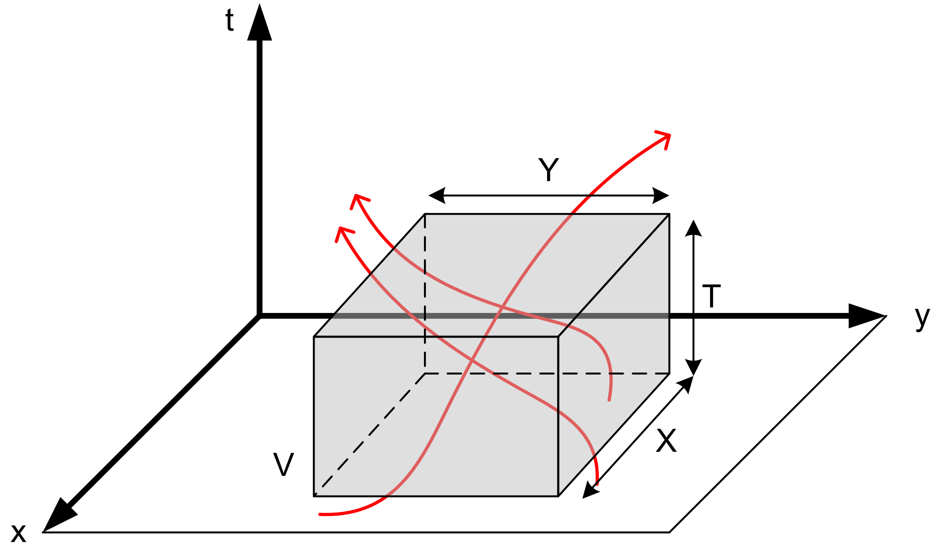
Trajectory vs. Link based NFD

- In both networks, for network densities greater than 20 vpmpl, the link-based method underestimates average network densities.
- Both estimation methods yield near-identical network flows.
- When densities are high, the link-based method does not fully capture the variability of the congestion effects in the network.
- Averaging the number of simulated vehicles on individual links in each time interval creates a bias in the link-based method when estimating network densities.

Outline

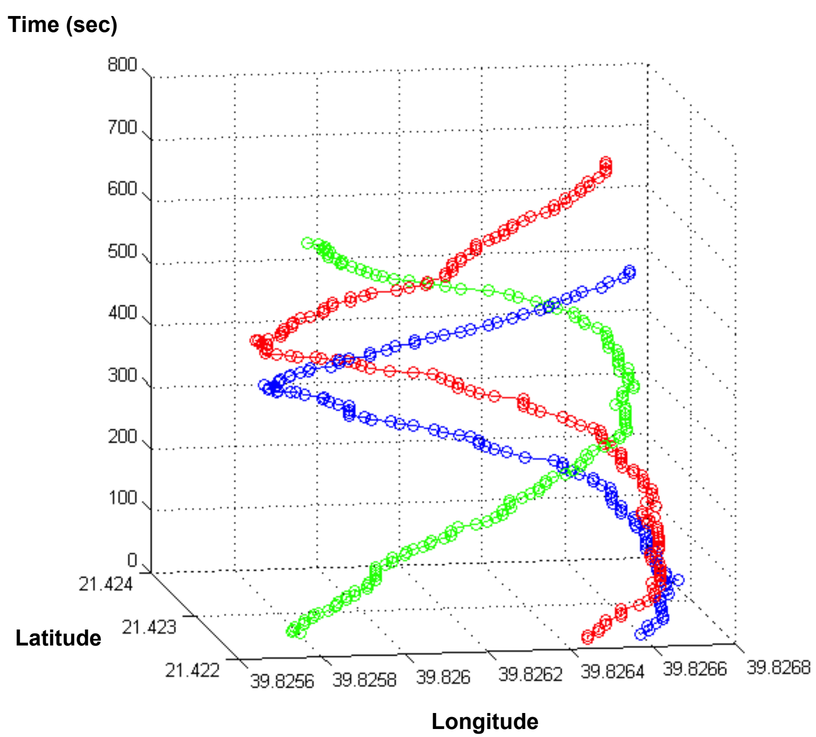
- Motivation
- From 2D trajectories to 3D trajectories
- Application to Network Flow Modeling (NFD's)
  - Vehicular networks
  - Pedestrians and crowds
- Travel time reliability
  - Signature relations and trajectory data
  - Within-day and day-to-day variability
- Scenario-based approach to reliability modeling
  - Trajectory Processor for particle-based simulators
- Takeways, Limitations and Challenges

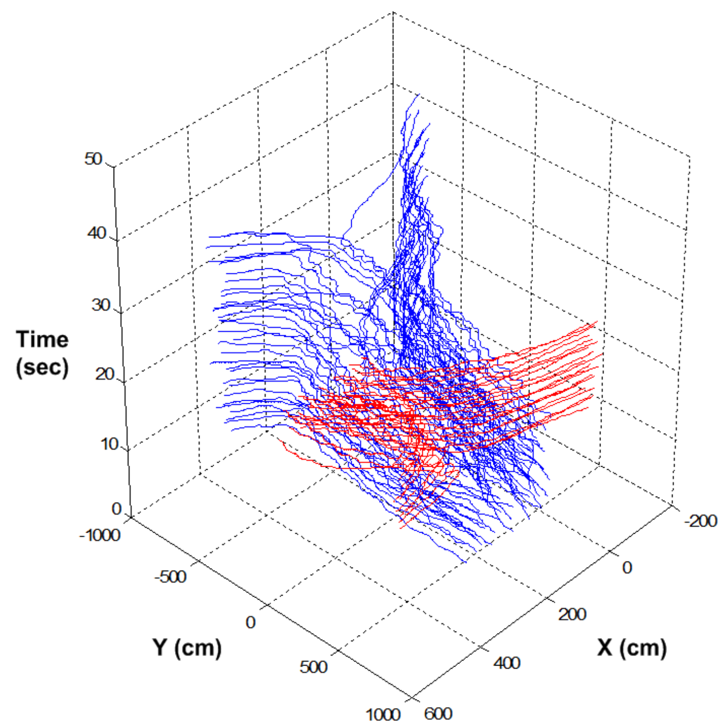
3D Time-Space Diagram



The three-dimensional time-space diagram of walking areas can be defined as a space in which the *x* and *y* axes represent the walking surface and the *z* axis represents time.

3D Time-Space Diagram





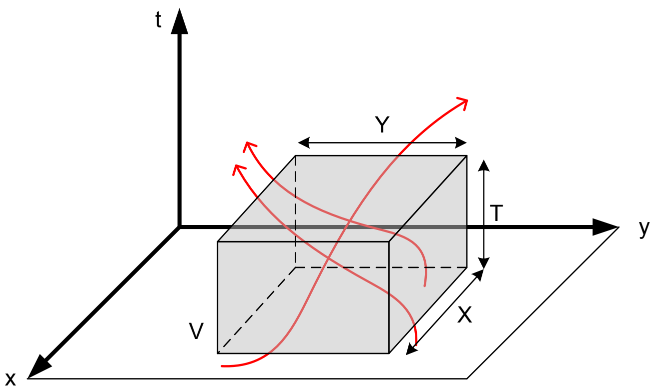
Source: Prof. Seyfried's group

Three-Dimensional Illustration of *(left)* Pilgrims Trajectories in a Circular Environment (n=3) and *(right)* Pedestrian Trajectories in a Bidirectional Environment (n=100)

Edie's Definitions

Extension to multi-directional pedestrian areas

$$k = \frac{\sum_{n \in N} \tau_n}{|V|} = \frac{\sum_{n \in N} \tau_n}{T.|A|} = \frac{\sum_{n \in N} \tau_n}{T.(X.Y)}$$

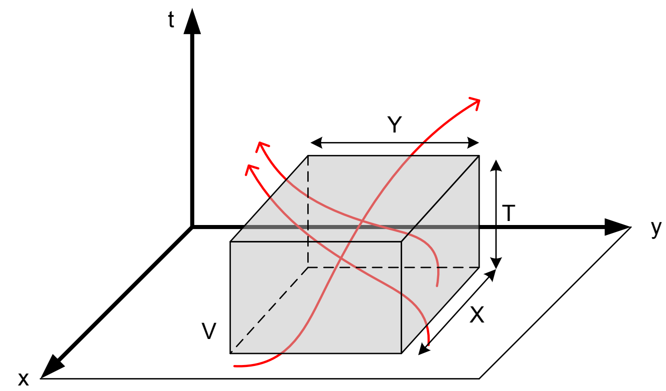


where τ_n is the total time spent by pedestrian *n* in shape *V* and |*V*| is the spatial volume covered by shape *V*. Also, |*V*| can be expressed as the geometric area of the walking area ($|A|=X.Y$) multiplied by the time interval $T=(t_1-t_0)$.

Edie's Definitions

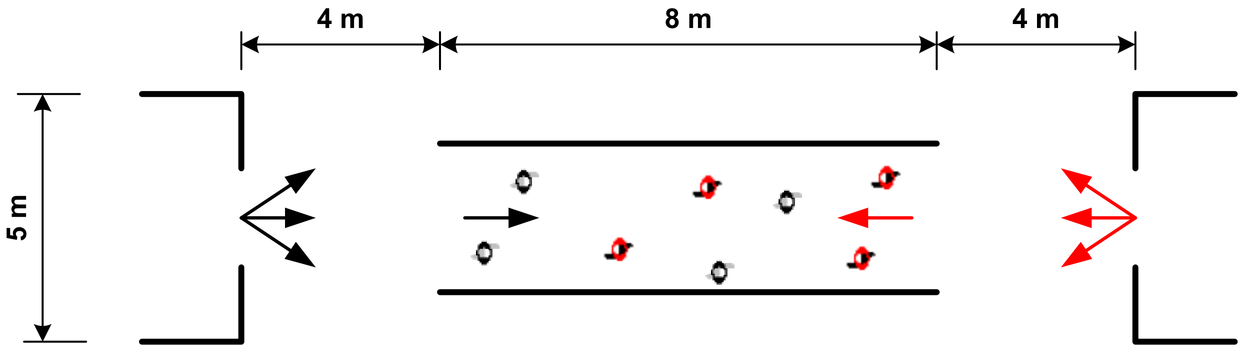
Extension to multi-directional pedestrian areas

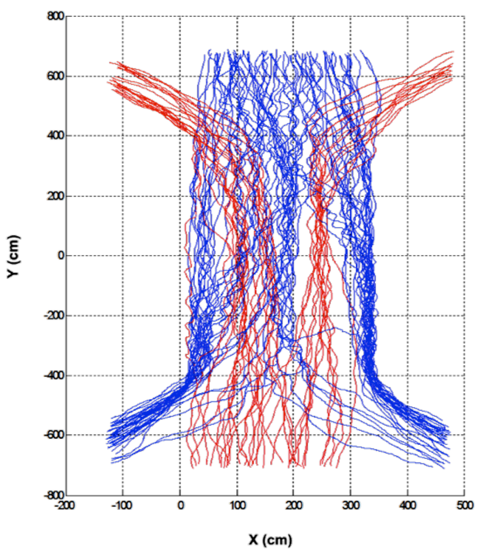
$$q = \frac{\sum_{n \in N} d_n}{|V|} = \frac{\sum_{n \in N} d_n}{Y.X.T}$$

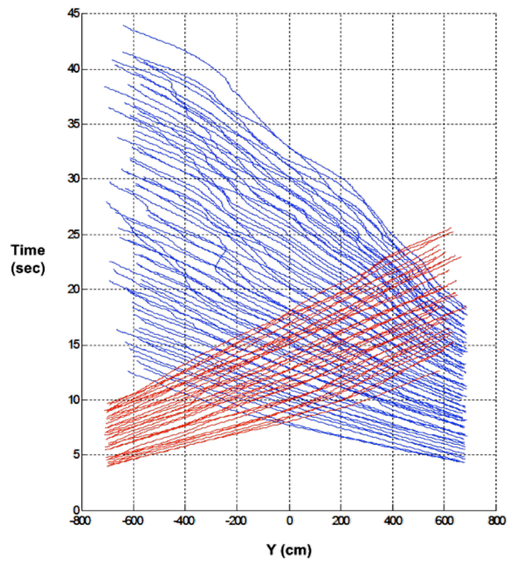


where d_n is the total distance traveled by pedestrian n in shape V and $|V|$ is the spatial volume covered by shape V.

Experimental Data

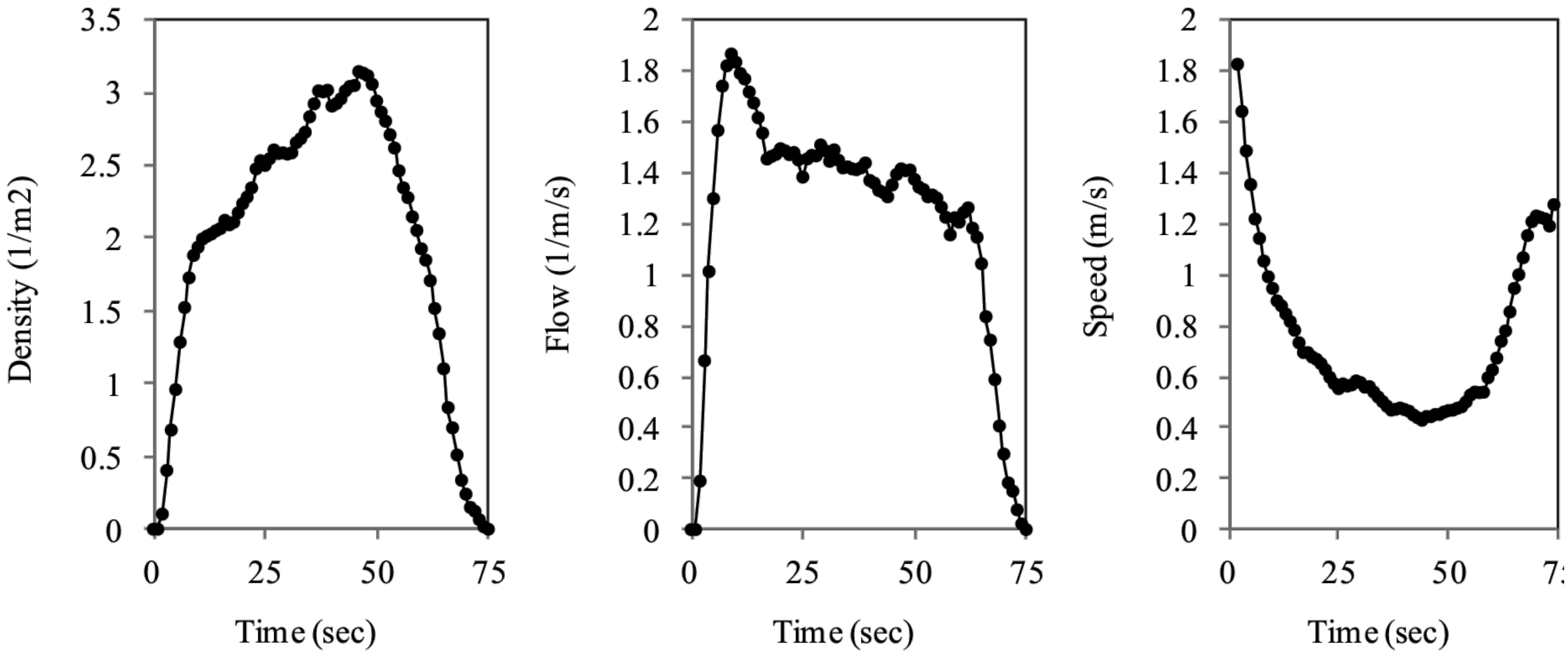






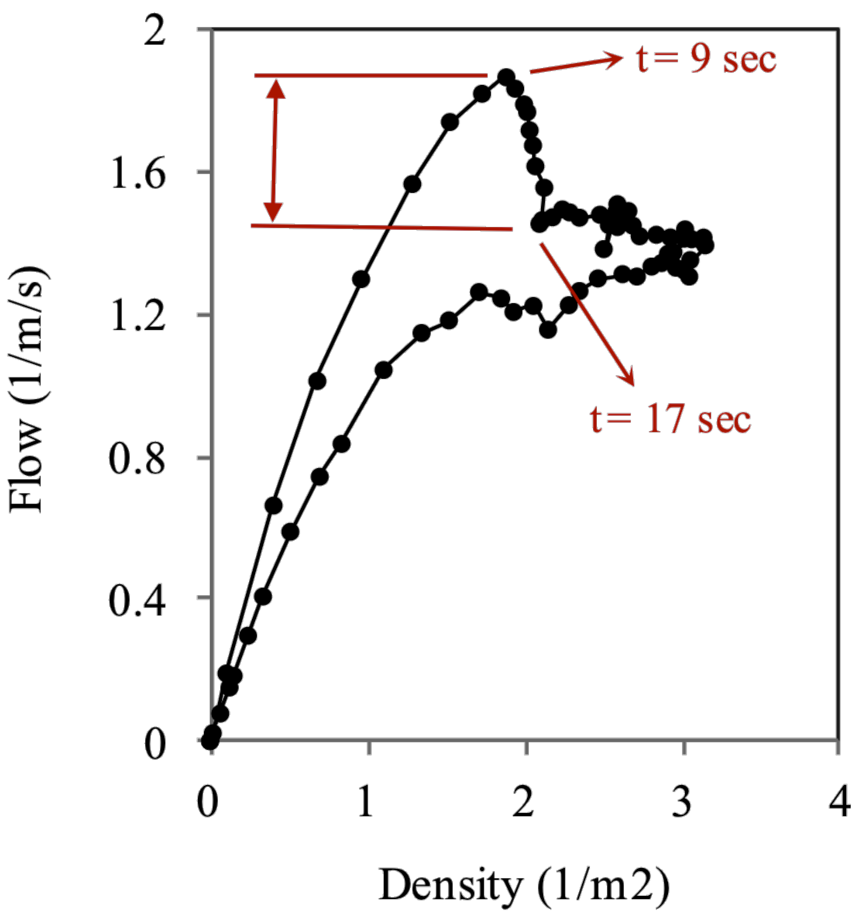
Zhang et al. (2011)

Pedestrian Traffic Measures



Time series of area-wide density, flow, and speed

Area-wide Fundamental Diagram



Similar to vehicular traffic flow on both individual facilities and networks, pedestrian traffic exhibits hysteretic behavior too.

The capacity drop phenomenon seems to exist in pedestrian crowds too.

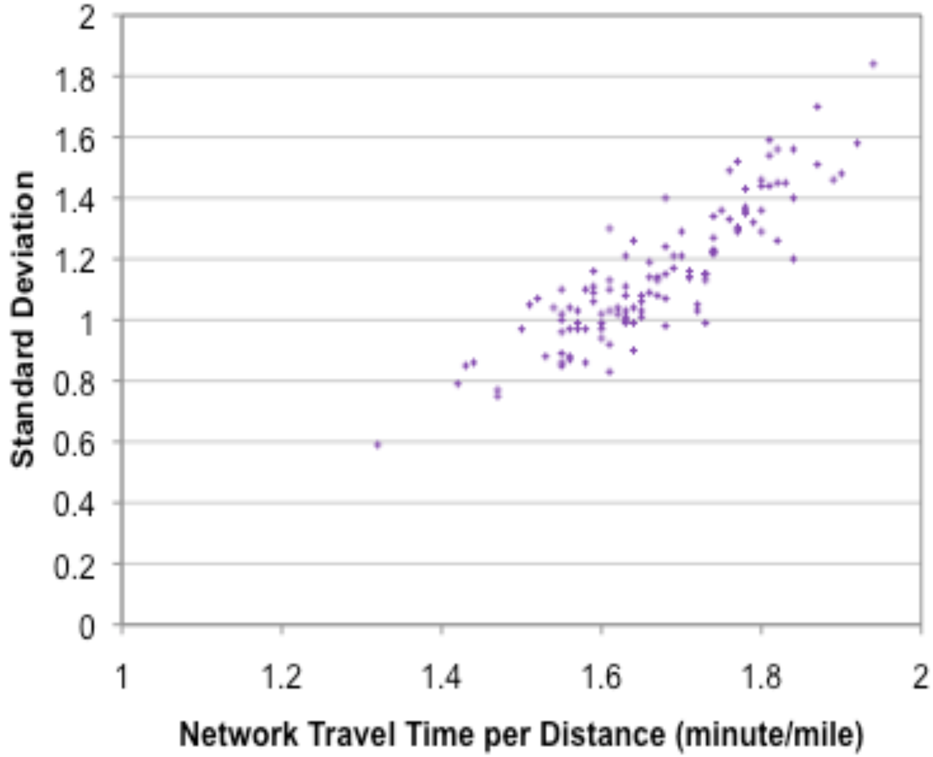
The observed capacity drop (t = 9-17 sec) is followed by a relatively stable period in which the area-wide flow remains roughly constant while density continues to increase due to formation of stable self-organized lanes.

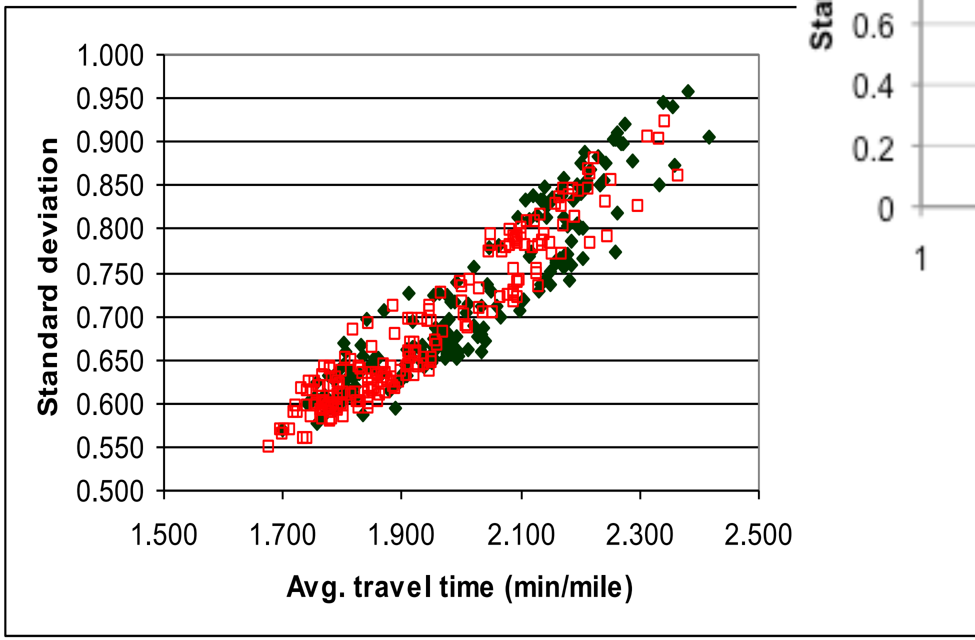
Outline

- Motivation
- From 2D trajectories to 3D trajectories
- Application to Network Flow Modeling (NFD's)
  - Vehicular networks
  - Pedestrians and crowds
- Travel time reliability
  - Signature relations and trajectory data
  - Within-day and day-to-day variability
- Scenario-based approach to reliability modeling
  - Trajectory Processor for particle-based simulators
- Takeways, Limitations and Challenges

Complex interactions, Collective Effects

Relation between standard deviation of trip time per mile and mean trip time per unit distance





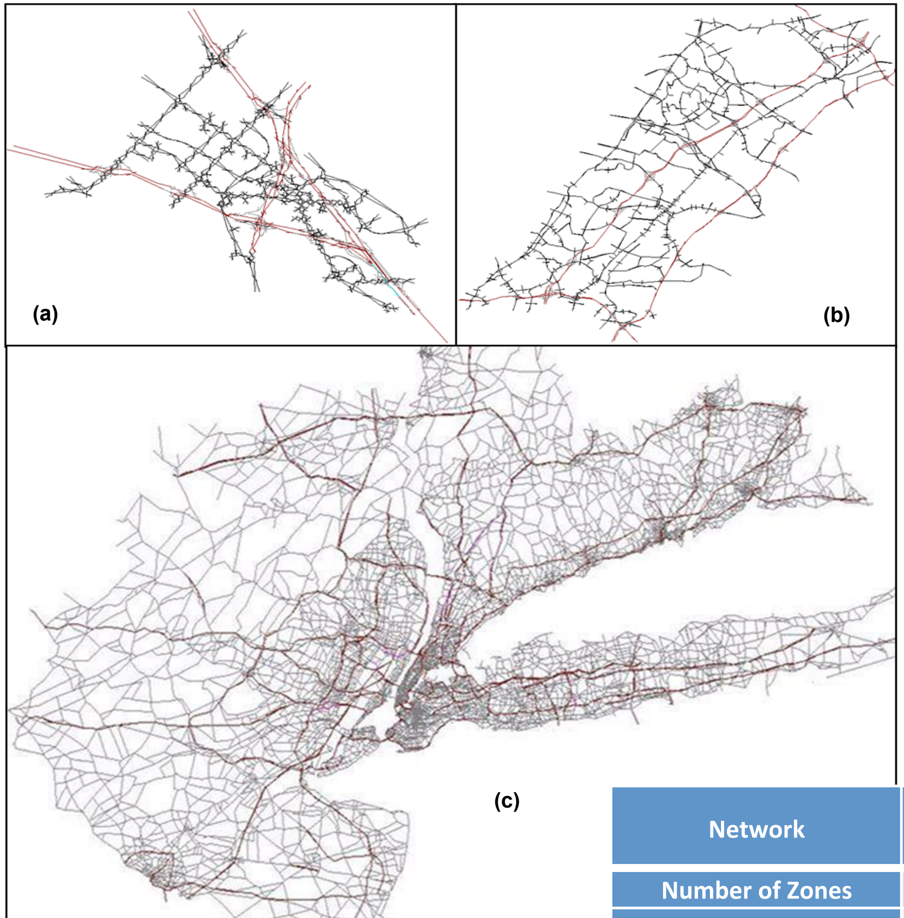
Theoretical Background – Travel Time Reliability

- Model has been validated and tested at different aggregation levels using different data sources (Mahmassani, H., Hou, T., Dong, J., TRB 2012)

- Data sources
  - Vehicle trajectories from simulation output
  - GPS probe data (location and time)

- Model works for different aggregation levels
  - Network level
  - O-D level
  - Path level
  - Link level

Simulated Trajectory Data

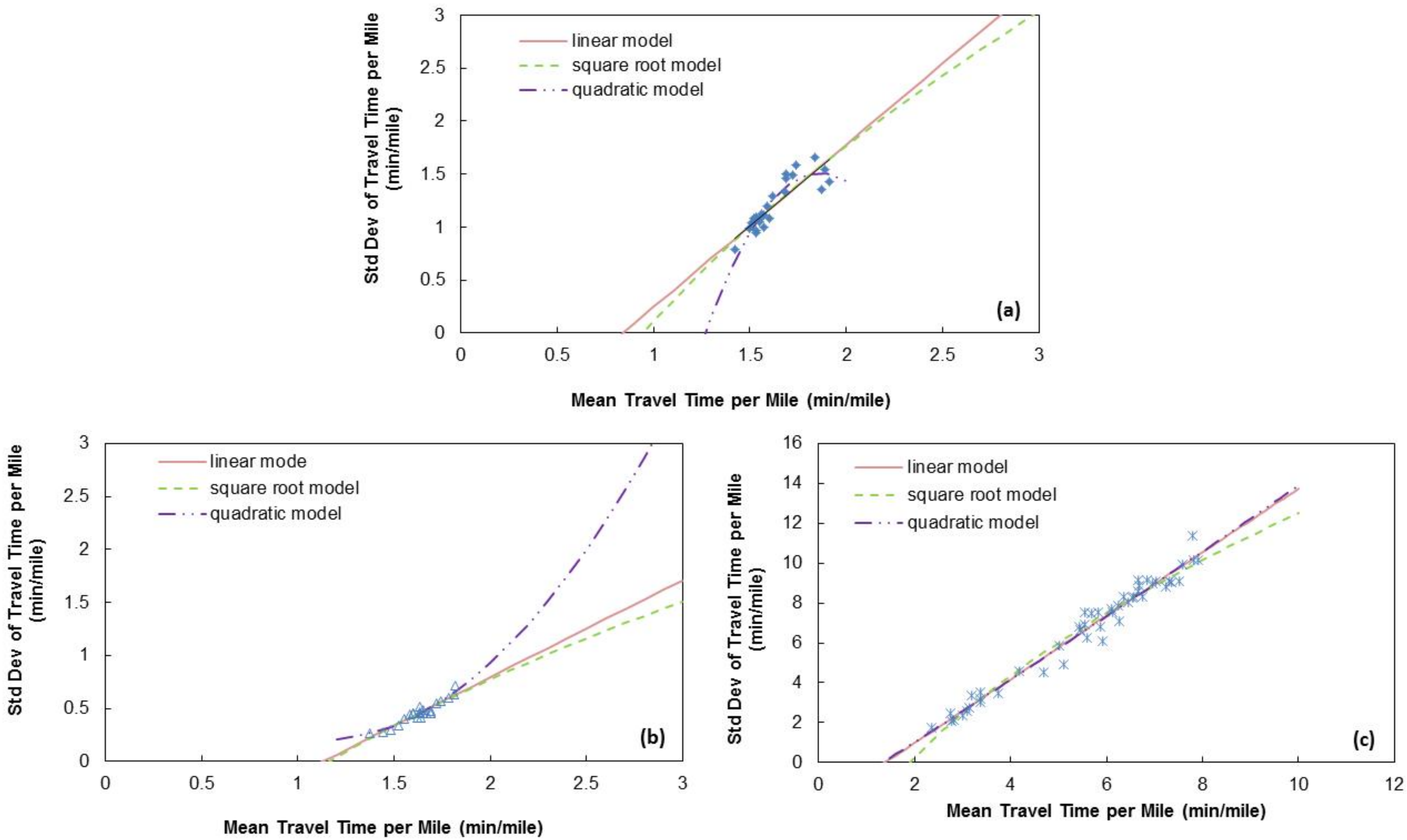


- Models are calibrated for different sizes of networks at different aggregation levels
- Three model forms are tested
  - Linear model
  - Square root model
  - Quadratic model
- Linear model gives best results
- Model parameters are estimated by Weighted Least Square (WLS) to accommodate heteroscedasticity

Network	Irvine	CHART	New York City
Number of Zones	61	111	3697
Number of Nodes	326	2182	28406
Number of Links	626	3387	68490
Number of Vehicles	58385	151973	6766805
Demand Duration (hr)	2	2	4

Simulated Trajectory Data

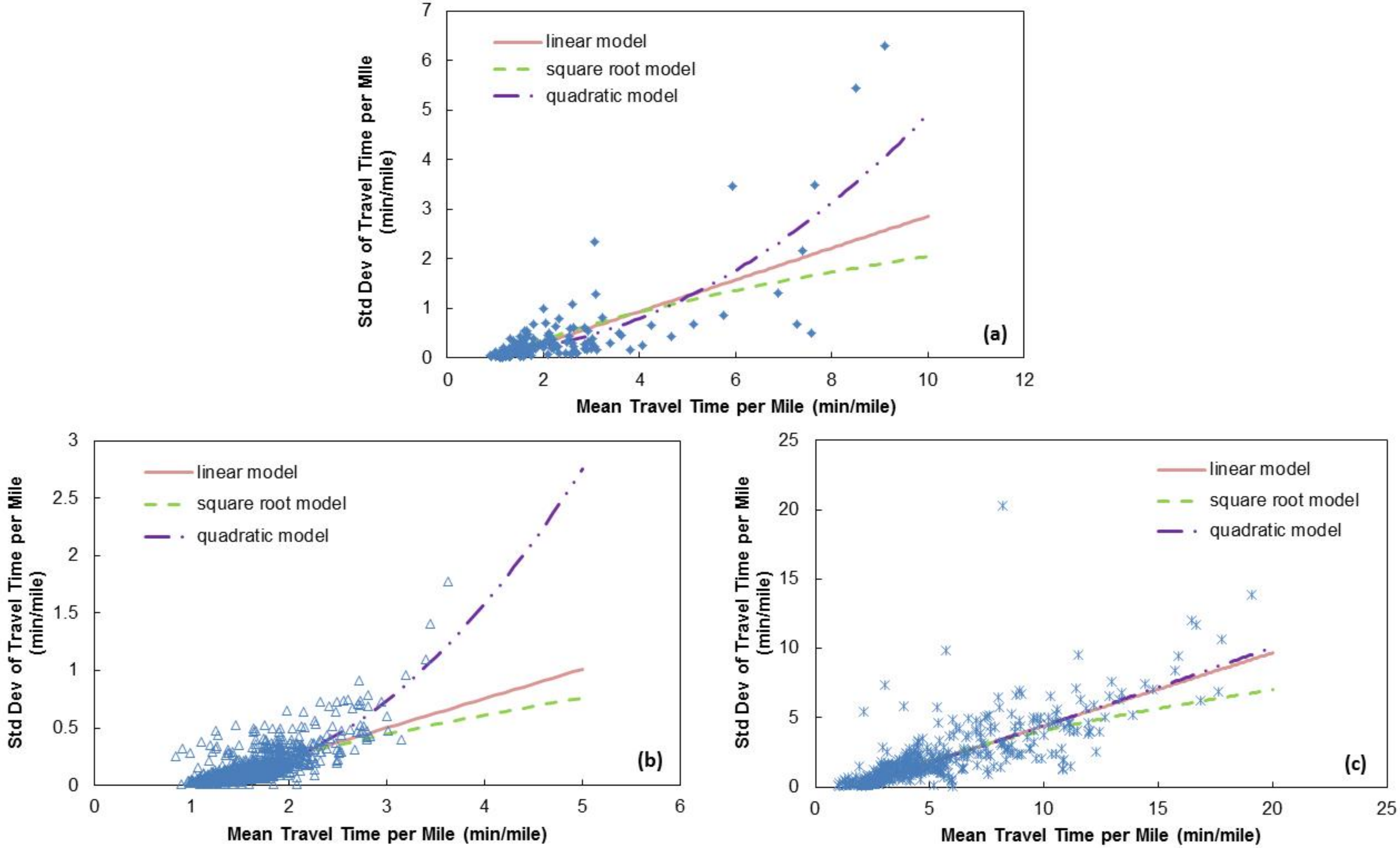
- Model comparison – network level analysis



(a) Irvine; (b) CHART; (c) New York City

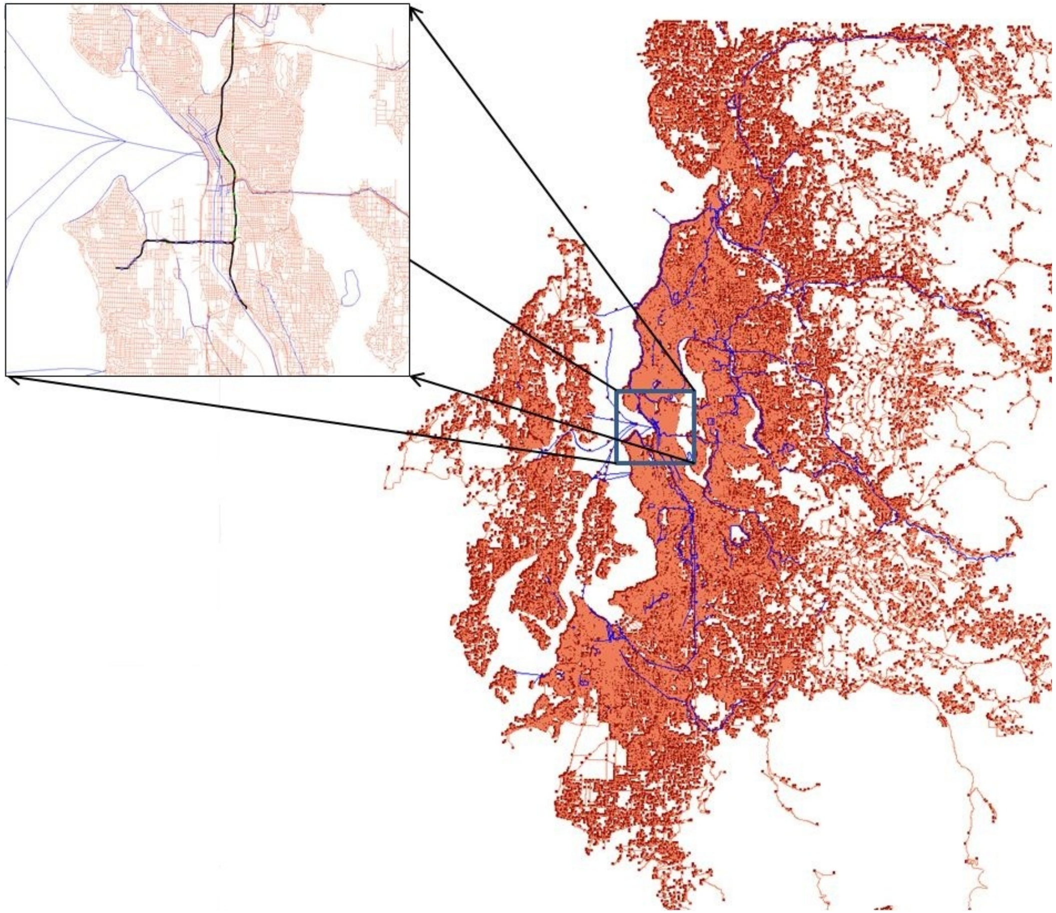
Simulated Trajectory Data

- Model comparison – path level analysis



(a) Irvine; (b) CHART; (c) New York City

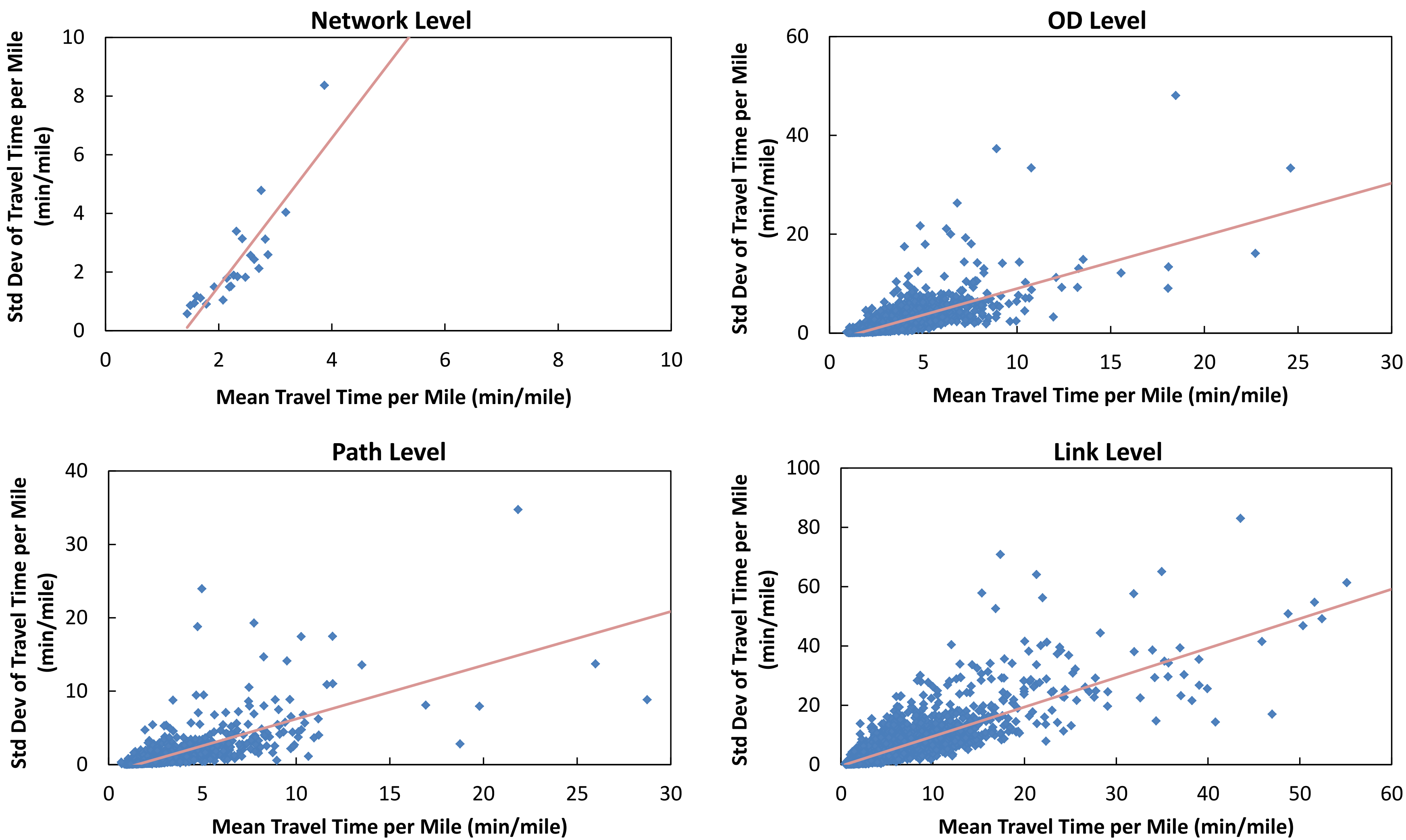
GPS Probe Data



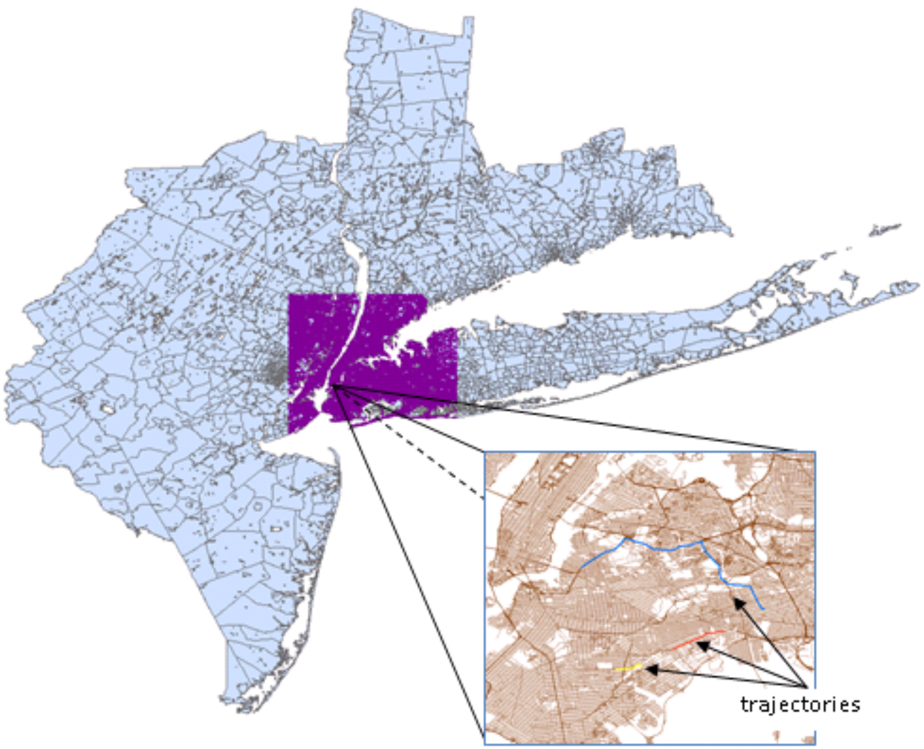
- Seattle network
  - ~600 zones
  - ~6000 nodes
  - 549,624 trips
  - ~400 participating vehicles

GPS Probe Data

- Seattle network



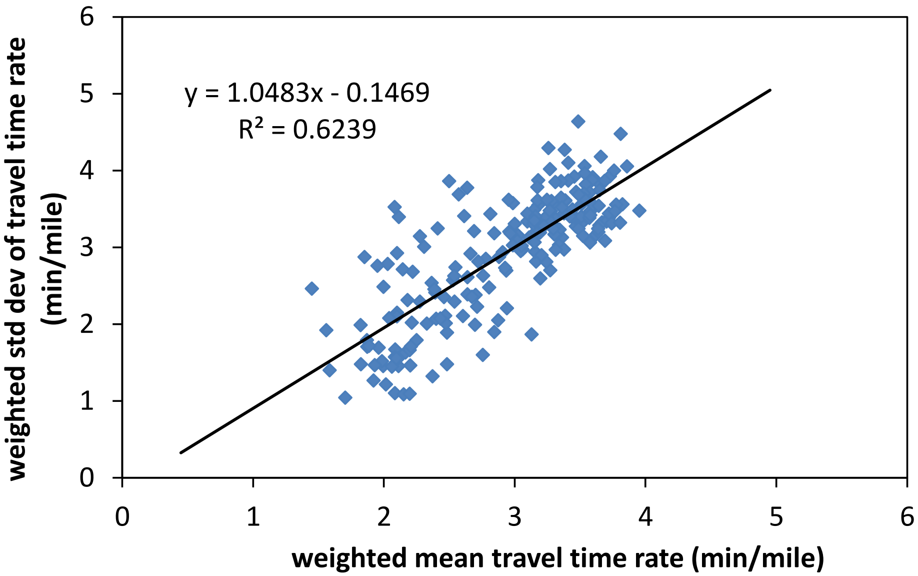
Validation by GPS Trajectory Data



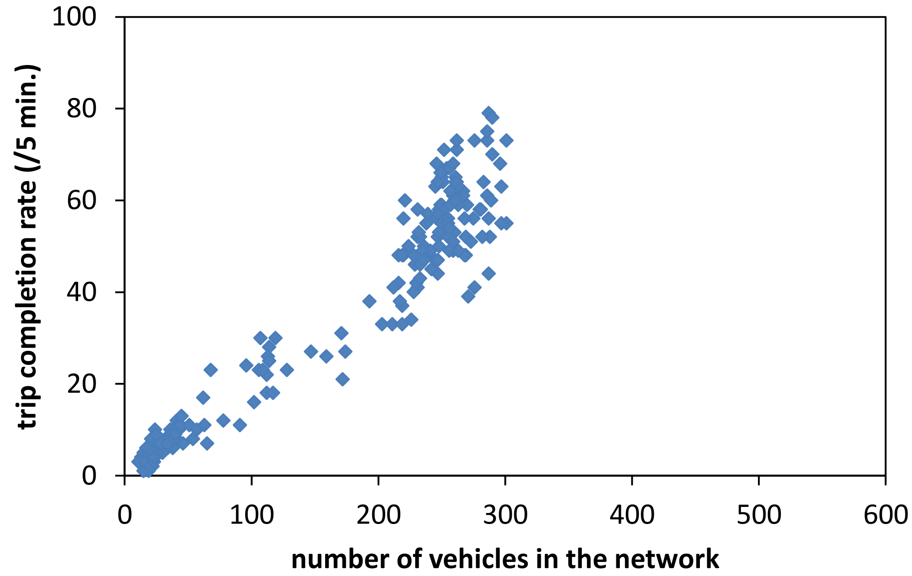
- NYC network
  - Vehicle trajectories collected by GPS devices
  - Two-week period from 2010/05/02 to 2010/05/17
  - ~10,000 trips are recorded on each day

Validation by GPS Trajectory Data

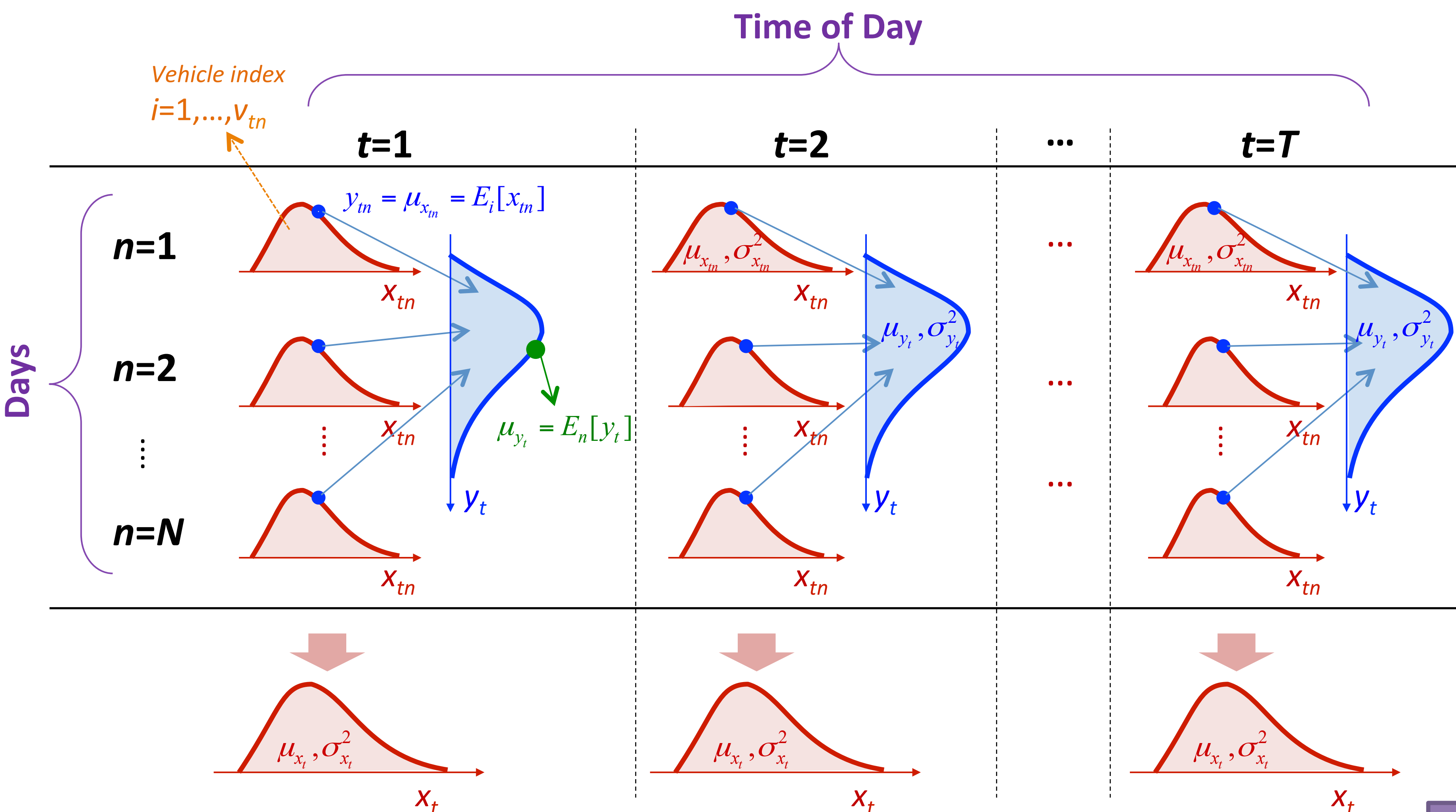
validation of linear travel time reliability model



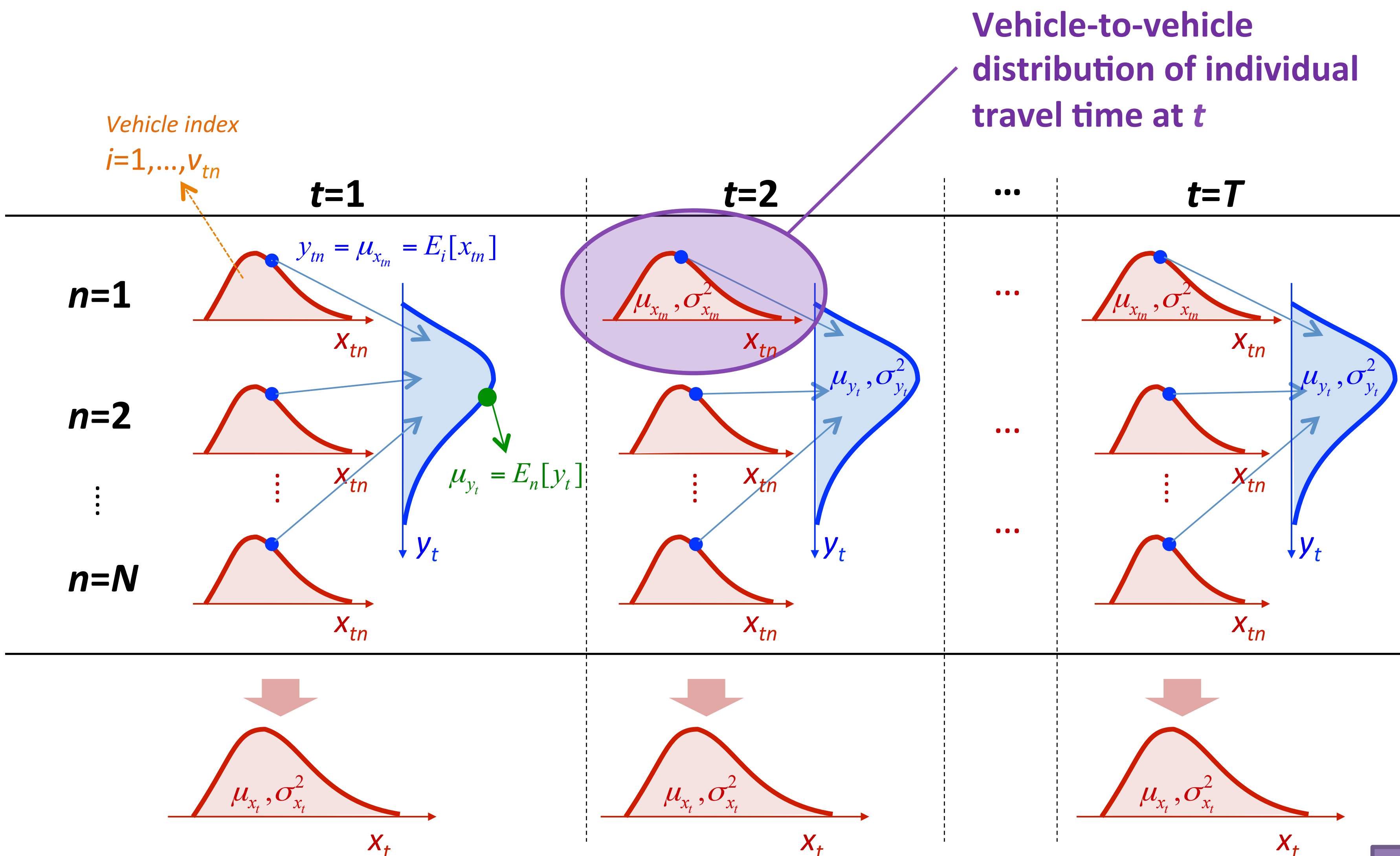
validation of NFD (network flow vs. network density)



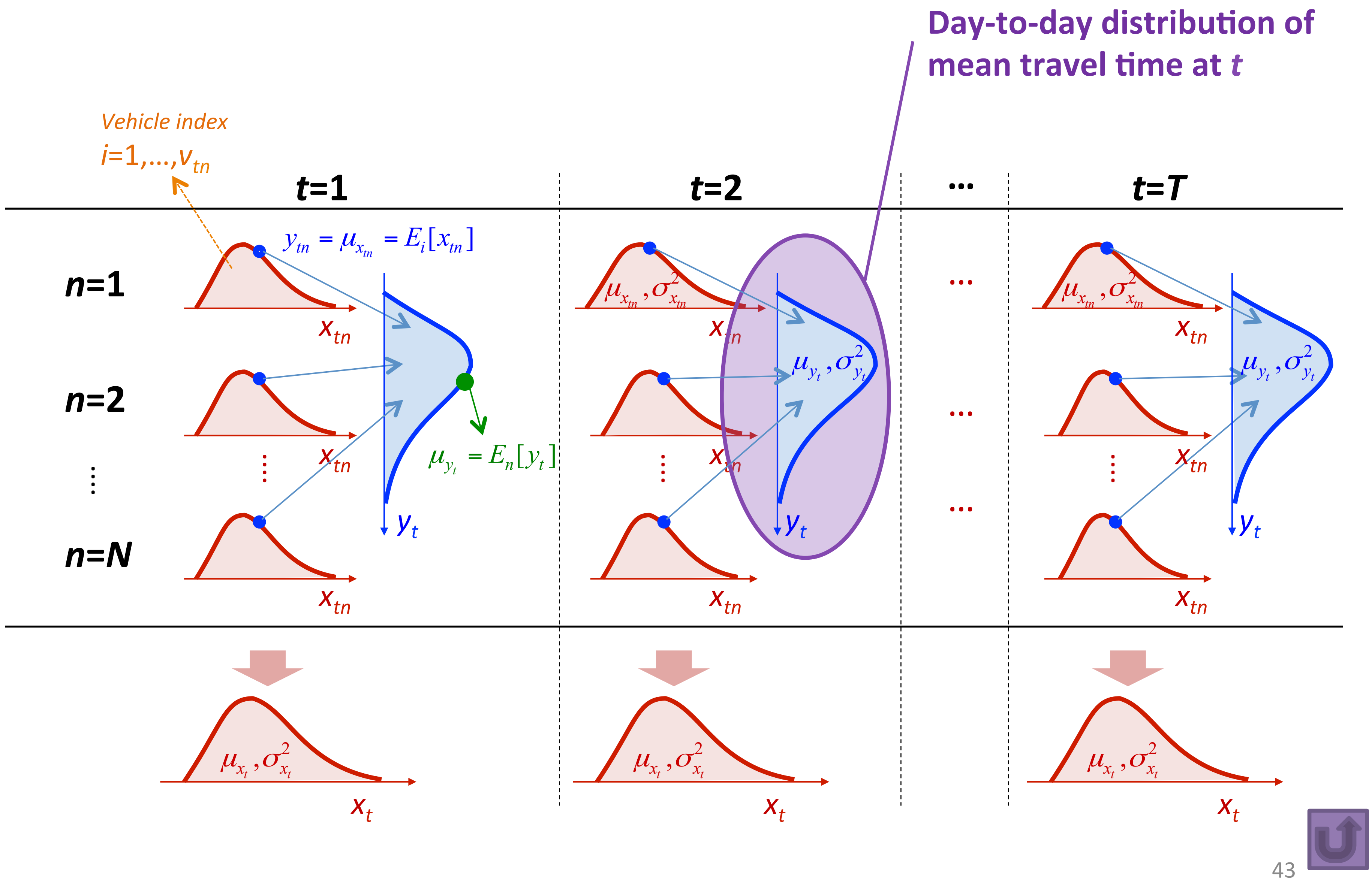
Characterizing Different Types of Travel Time Variability



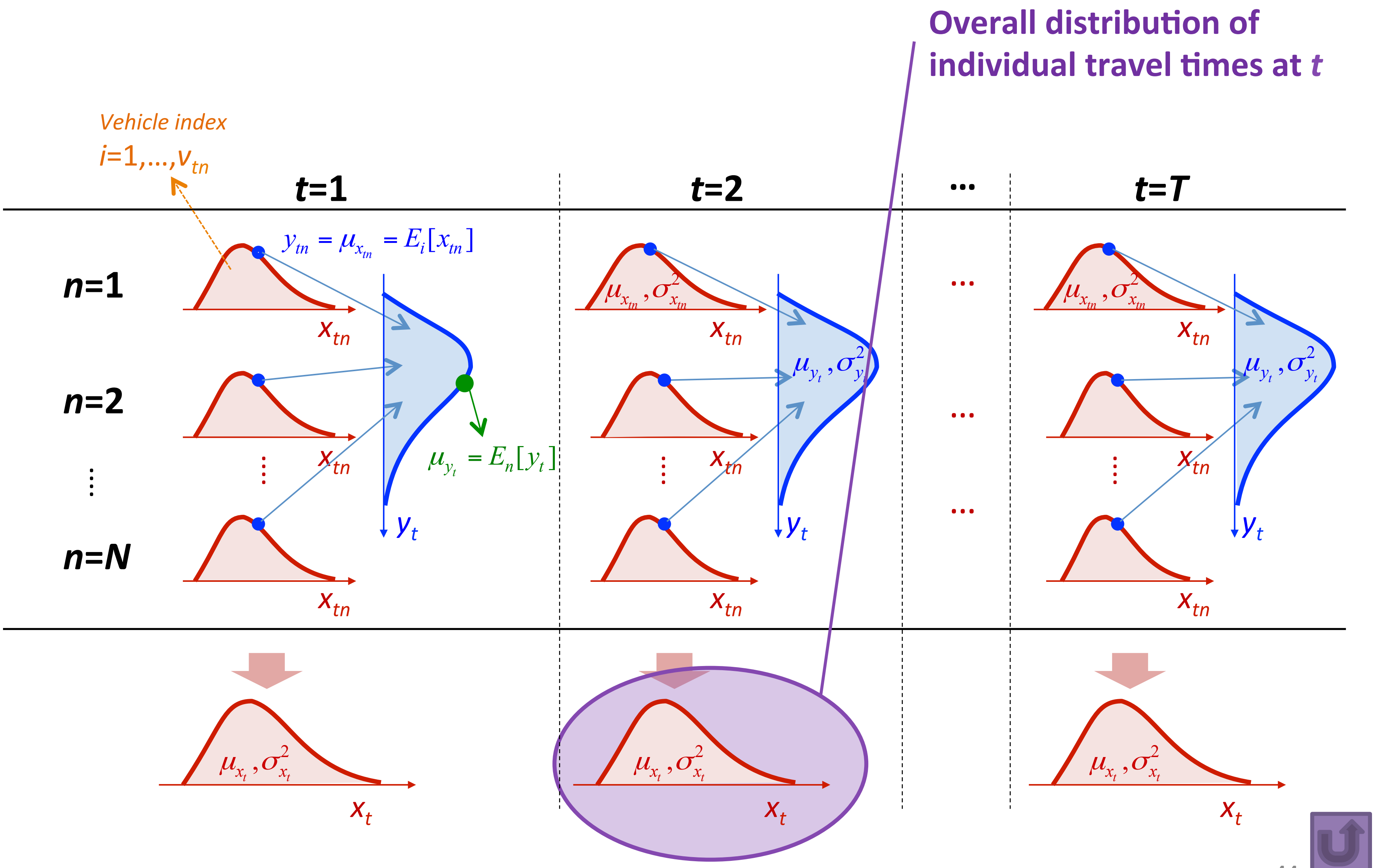
Characterizing Different Types of Travel Time Variability



Characterizing Different Types of Travel Time Variability



Characterizing Different Types of Travel Time Variability

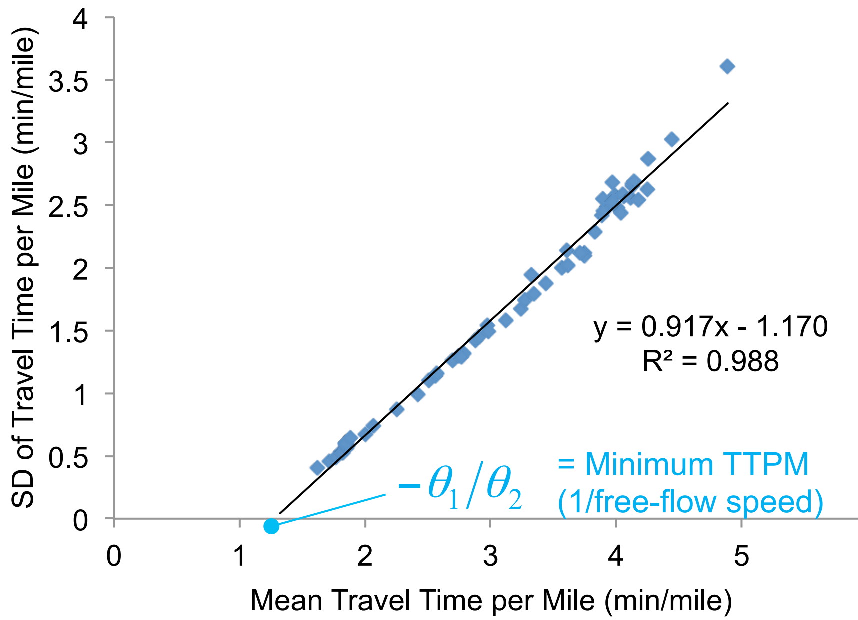


Characterizing *Vehicle-to-vehicle* Variability

- Linear Relationship between Standard Deviation (SD) and Mean
  - Jones et al. (1989) and Mahmassani et al. (2012, 2013)

$$\sigma_\tau = \theta_1 + \theta_2 \mu_\tau$$

- τ : travel time per unit distance; travel time per mile (TTPM)
- $\sigma_\tau,\ \mu_\tau$: mean and SD of τ ; θ_1, θ_2 : coefficients



Characterizing *Vehicle-to-vehicle* Variability

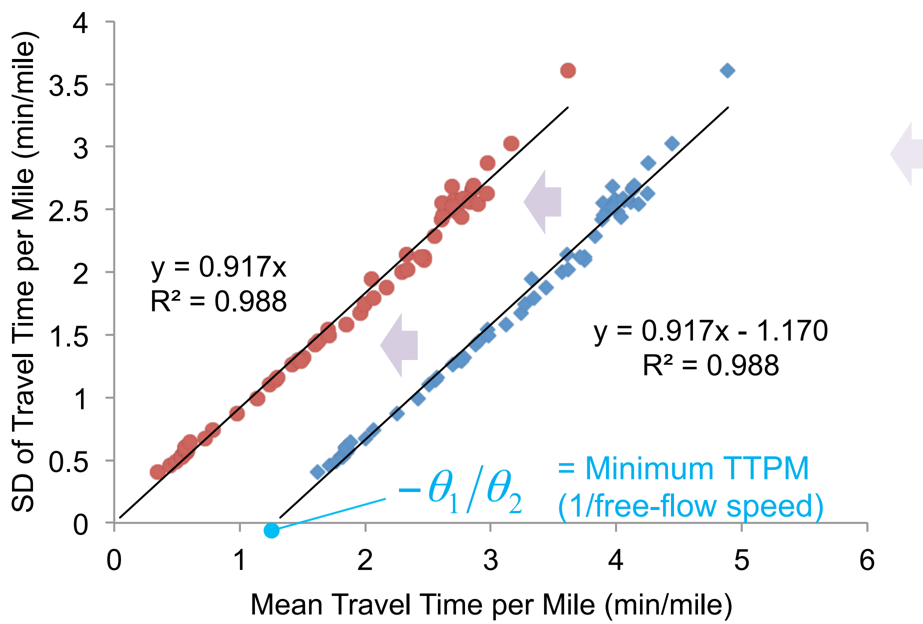
- Travel Delay per Unit Distance

TTPM – minimum TTPM

$$x = \tau - (-\theta_1/\theta_2)$$

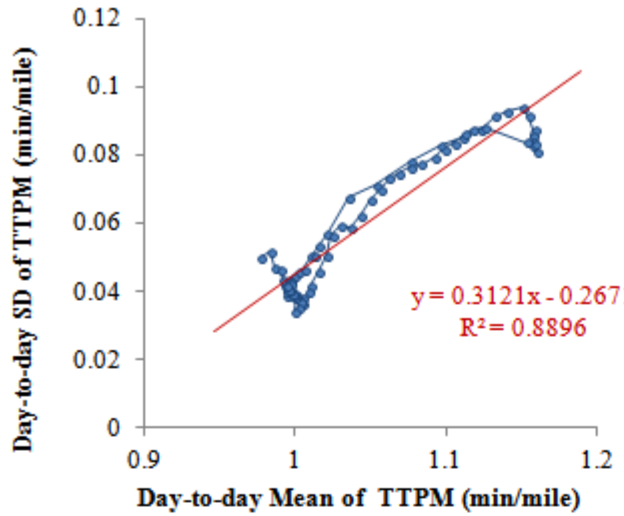
$$\Rightarrow \sigma_x = \theta_2 \mu_x$$

$$\begin{cases} E[x] = \mu_x = \mu_\tau + \theta_1/\theta_2 \\ SD[x] = \sigma_x = \sigma_\tau \end{cases}$$

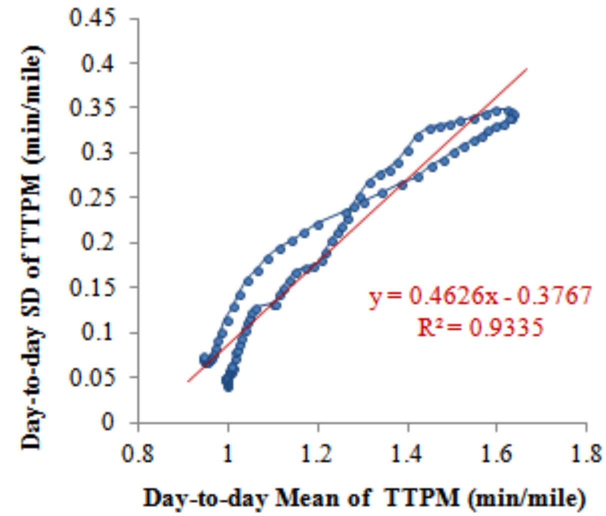


Characterizing *Day-to-day* Variability

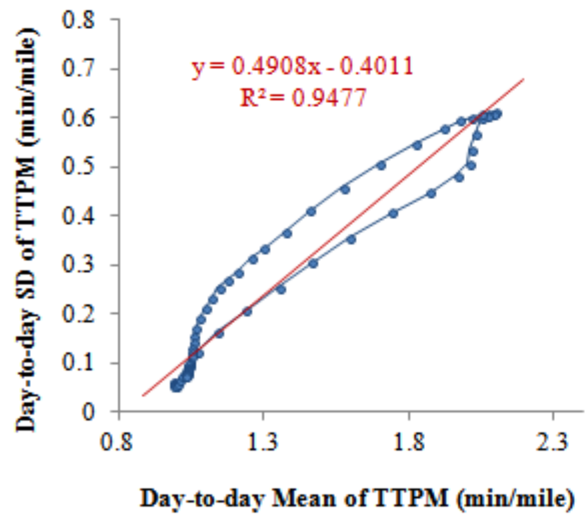
- Strong Correlation between SD and Mean
  - Herman and Lam(1974); and Richardson and Taylor (1978)
- Linear Relation between SD and Mean
  - May et al. (1989); Mazloumi et al. (2010); Yildrimoglu et al. (2013); Fosgerau (2010); and Fosgerau and Fukuda (2012)



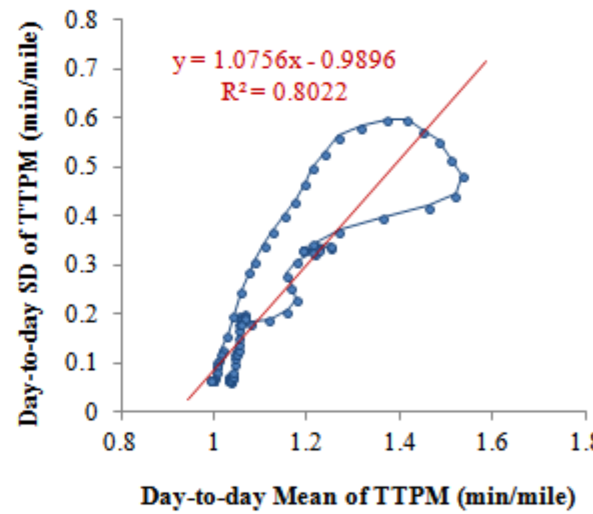
(a) I-880S: Morning (6AM – 12PM)



(b) I-880S: Afternoon (12PM – 9PM)



(c) I-405N: Morning (6AM – 12PM)



(d) I-405N: Afternoon (12PM – 9PM)

Multiplicative Error Models

Vehicle-to-vehicle Distribution

$$\sigma_{x_{tn}} = \alpha\, \mu_{x_{tn}}$$

$$x_t = y_t \varepsilon_x$$

$$\varepsilon_x \sim Gamma(\pi, 1/\pi)$$

$$SD_i[x_t] = \underbrace{SD_i[\varepsilon_x]}_{\alpha = 1/\sqrt{\pi}} y_t$$

Day-to-day Distribution

$$\sigma_{y_t} = \beta\, \mu_{y_t}$$

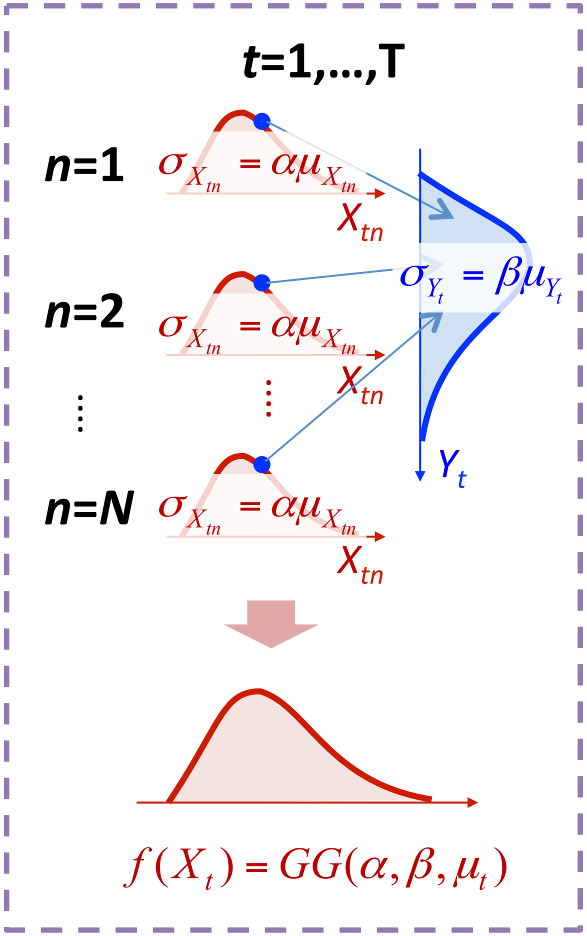
$$y_t = \mu_{y_t} \varepsilon_y$$

$$\varepsilon_y \sim Gamma(\phi, 1/\phi)$$

$$SD_n[y_t] = \underbrace{SD_n[\varepsilon_y]}_{\beta = 1/\sqrt{\phi}} \mu_{y_t}$$

Gamma-Gamma Distribution

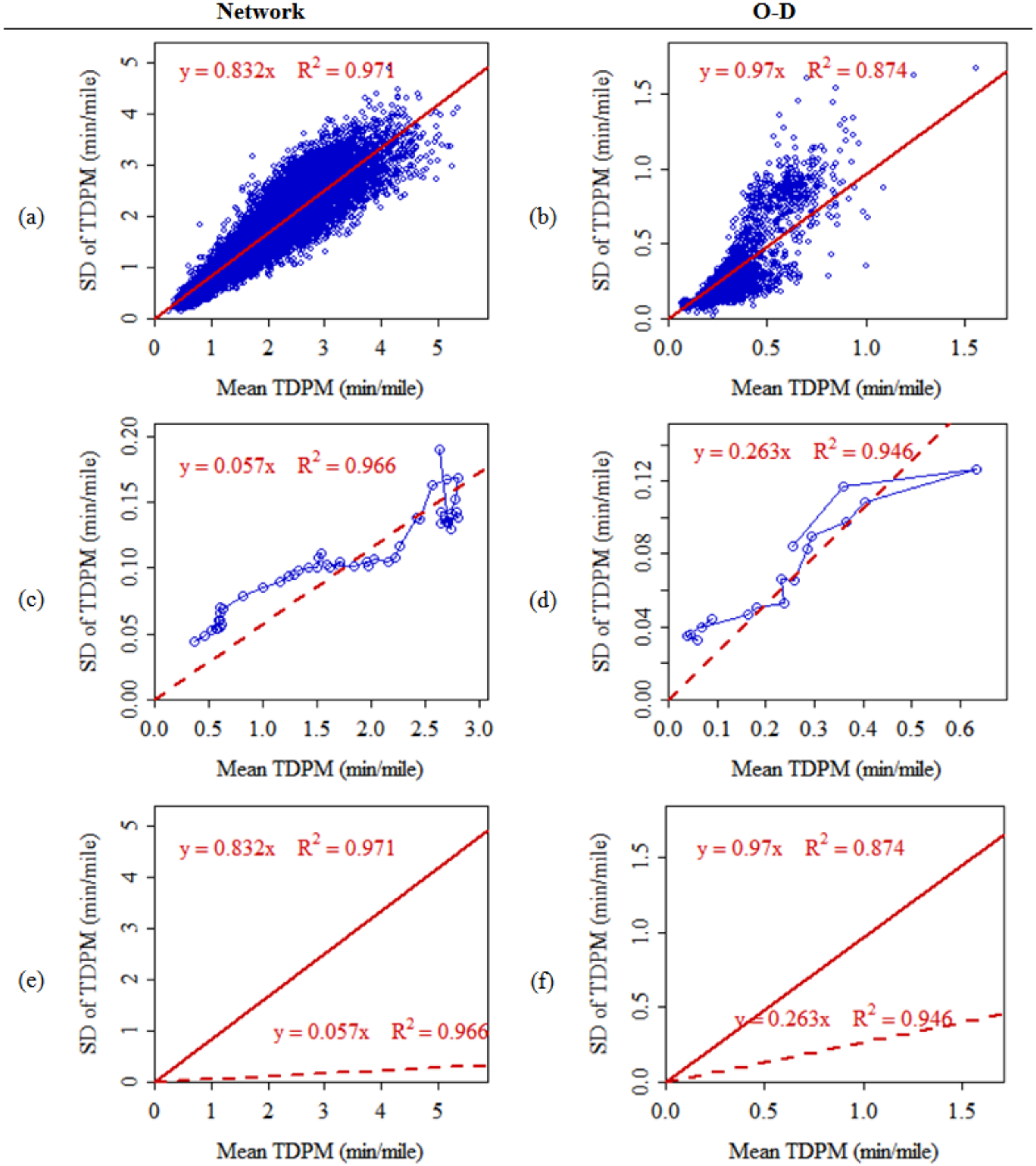
- Describe diverse scattering phenomena
  - target and clutter scattering in radar (Jakeman and Pusey, 1976; Lewinski, 1983);
  - irradiance fluctuations in optics (Al-Habash et al., 2001; Teich and Diament, 1989);
  - reverberation in sonar systems (Gu and Abraham, 2001); and
  - fading and shadowing in wireless systems (Shankar, 2004)



- For Modeling Travel Time Variability
  - Shape π reflects veh-to-veh variability (i.e., $\alpha = 1/\sqrt{\pi}$: CV of individual travel delay across vehicles)
  - Shape ϕ reflects day-to-day variability (i.e., $\beta = 1/\sqrt{\phi}$: CV of daily mean level of travel delay across days)
  - Mean μ_t represents the mean level of travel delay at time t

- Moments

$$E[X_t] = \mu_t \qquad Var[X_t] = \mu_t^2 \frac{\pi + \phi + 1}{\pi\phi} = \mu_t^2\left(\alpha^2 + \beta^2 + \alpha^2\beta^2\right)$$

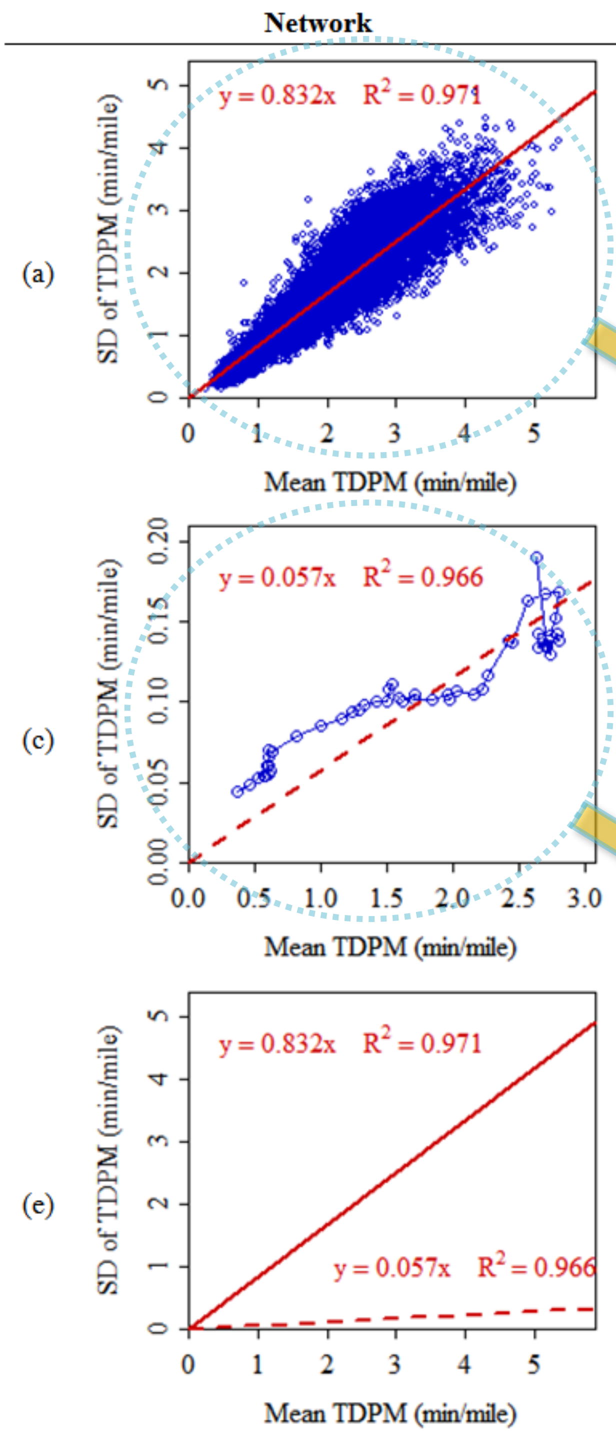


Vehicle-to-vehicle Dist.

$$\boxed{\sigma_{x_{tn}} = \alpha \, \mu_{x_{tn}}} = \frac{1}{\sqrt{\pi}} \mu_{x_{tn}}$$

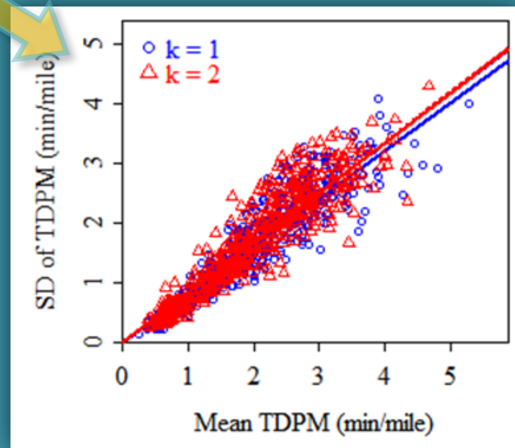
Day-to-day Distribution

$$\boxed{\sigma_{y_t} = \beta \, \mu_{y_t}} = \frac{1}{\sqrt{\phi}} \mu_{y_t}$$

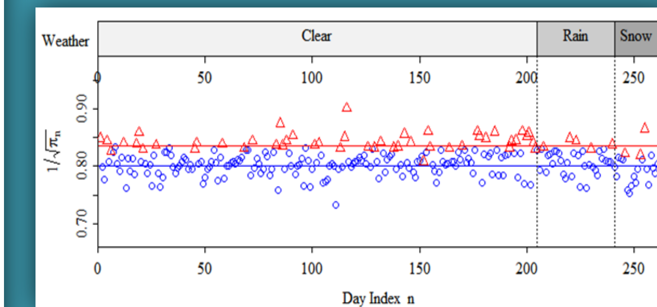


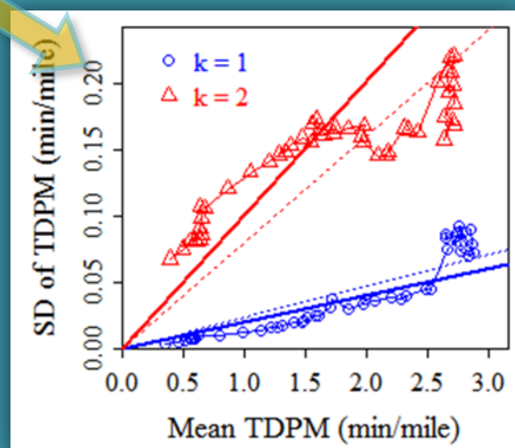
Finite Mixture Model *(Generalization)*

$$f(x_t) = \sum_{k=1}^{K} w_k \text{Gamma-Gamma}(x_t; \pi_k, \phi_k, \mu_{tk})$$

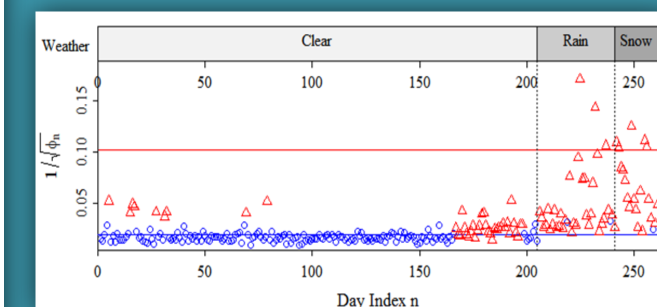


*Veh-to-veh (**K**=2)*





*Day-to-day (**K**=2)*



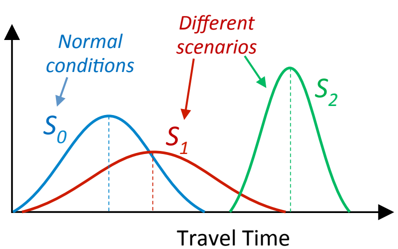
Outline

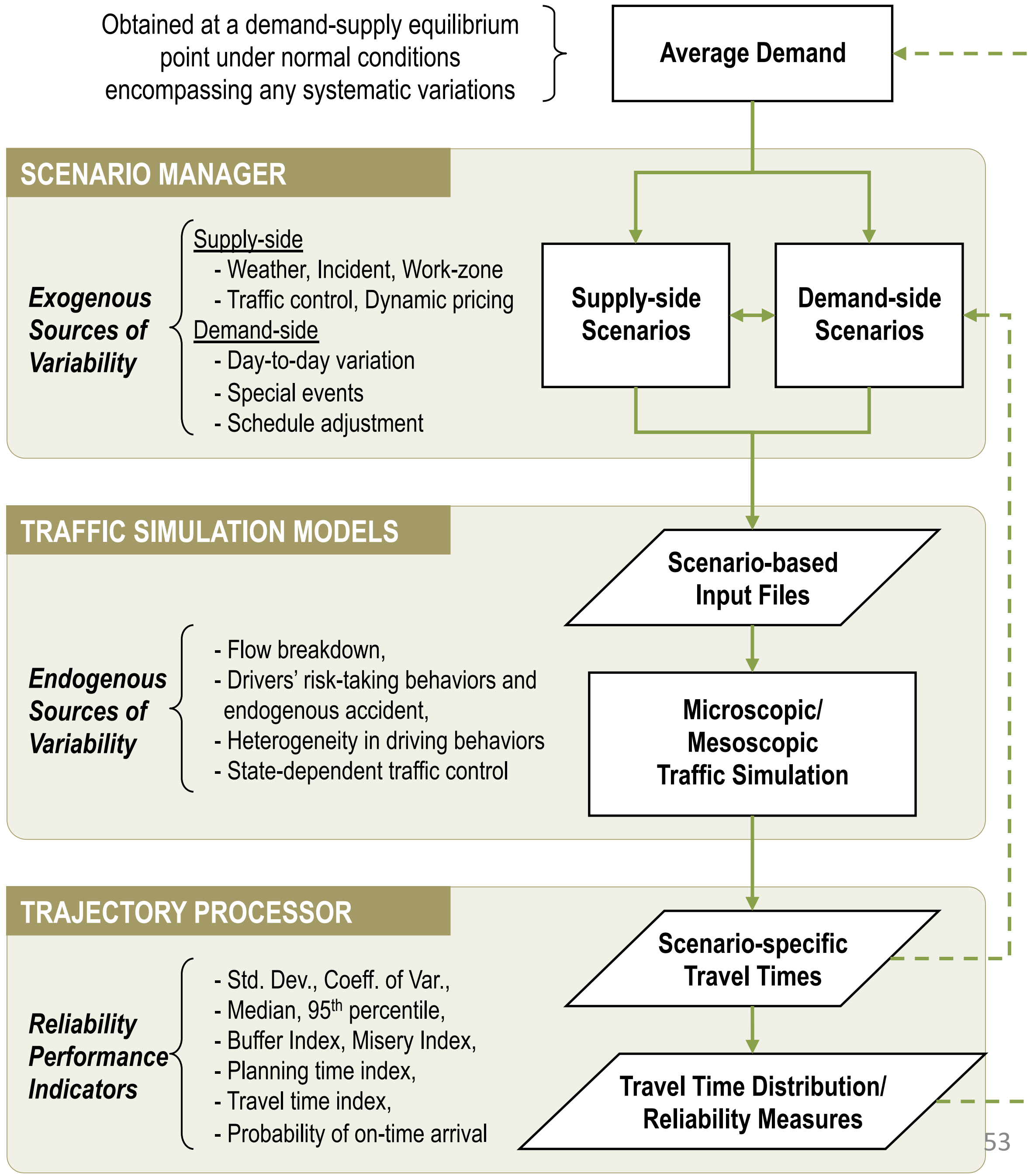
- Motivation
- From 2D trajectories to 3D trajectories
- Application to Network Flow Modeling (NFD's)
  - Vehicular networks
  - Pedestrians and crowds
- Travel time reliability
  - Signature relations and trajectory data
  - Within-day and day-to-day variability
- Scenario-based approach to reliability modeling
  - Trajectory Processor for particle-based simulators
- Takeways, Limitations and Challenges

Travel Time Reliability Analysis Framework

Trajectory Processor

Extract reliability-related measures from the vehicle trajectory output of the simulation models





Scenario-based Reliability Analysis

Define Scenario S | **Simulate Scenario S** | **Obtain Distribution of Y_t**

Events Realizations or Probabilities

Distributions of Parameters

X_S

S

$P(S)$

Scenario S with associated probability P(S)

$Y_t | S$

$f(Y_t | S)$

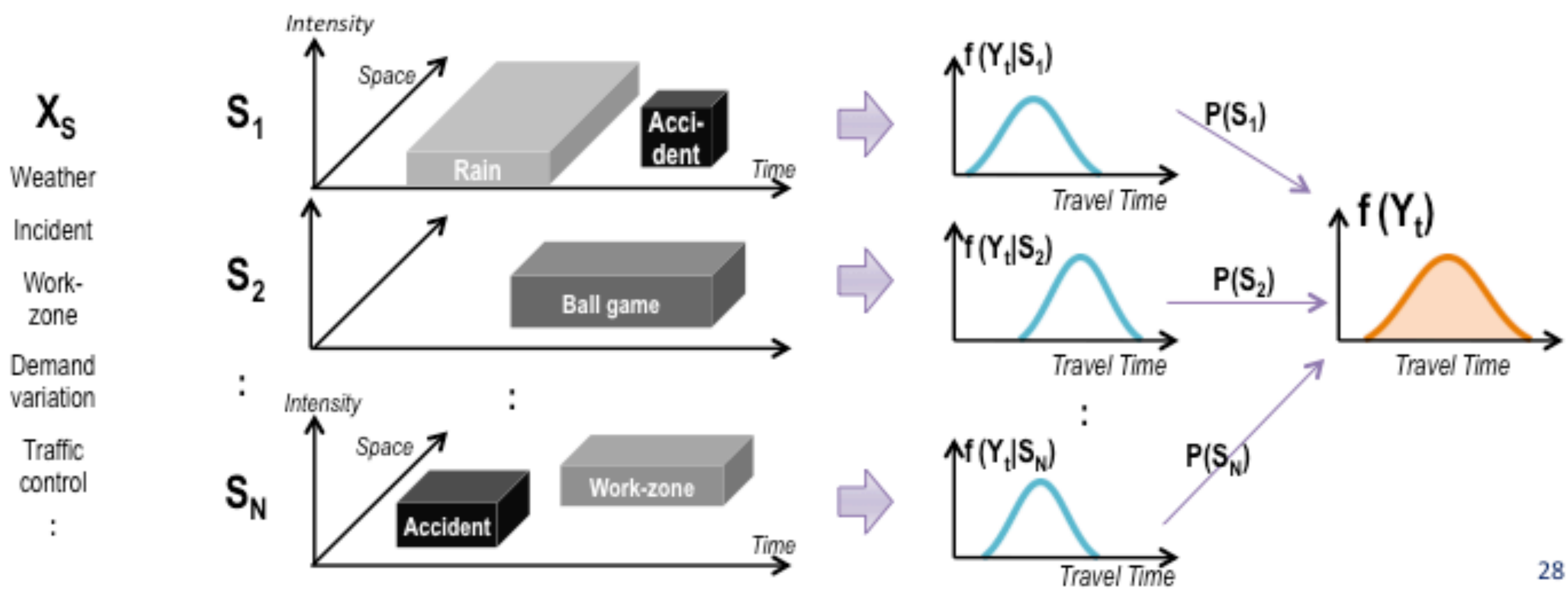
Obtain scenario-specific travel time Y|S or distribution $f_Y(Y|S)$

$E[Y_t]$

$= \Sigma_s E[Y_t | S] \cdot P(S)$

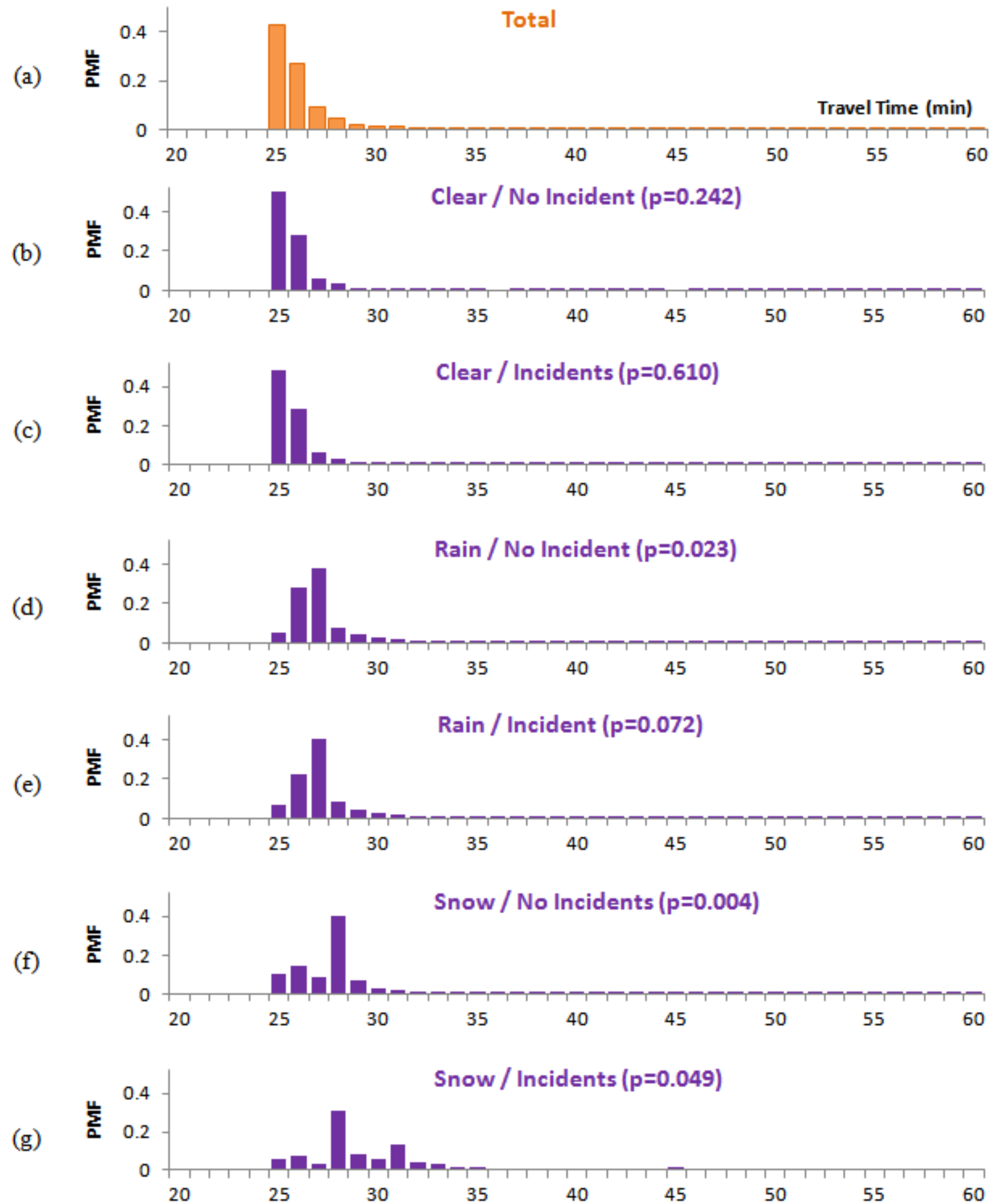
$f(Y_t)$

$= \Sigma_s f(Y_t | S) \cdot P(S)$



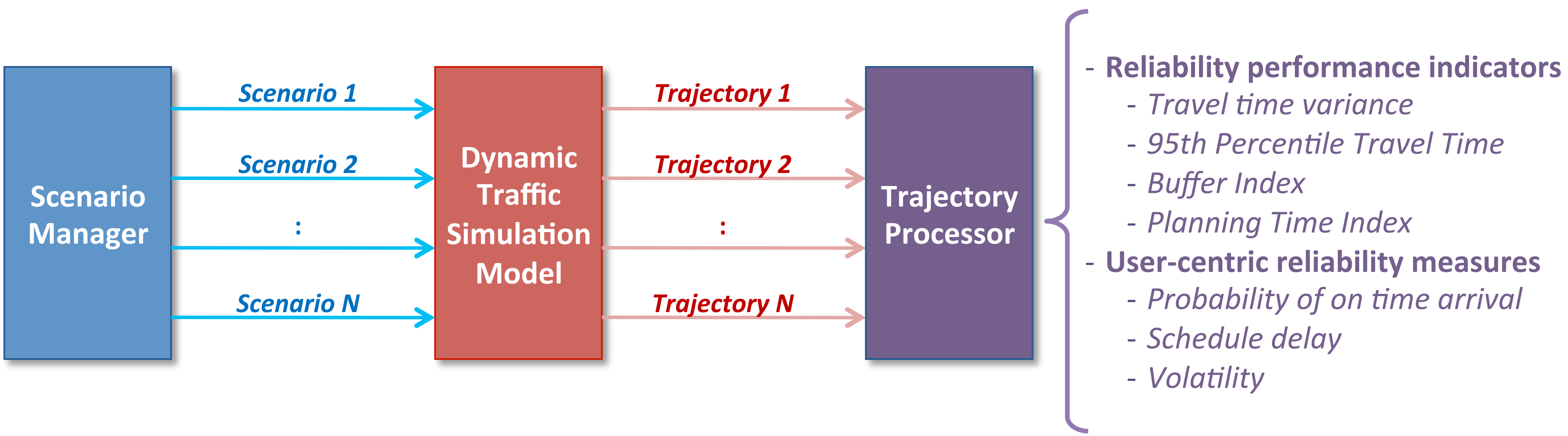
Case Study

- Understand the impact of each scenario category
- Observe the overall travel time distribution



Vehicle Trajectory Processor

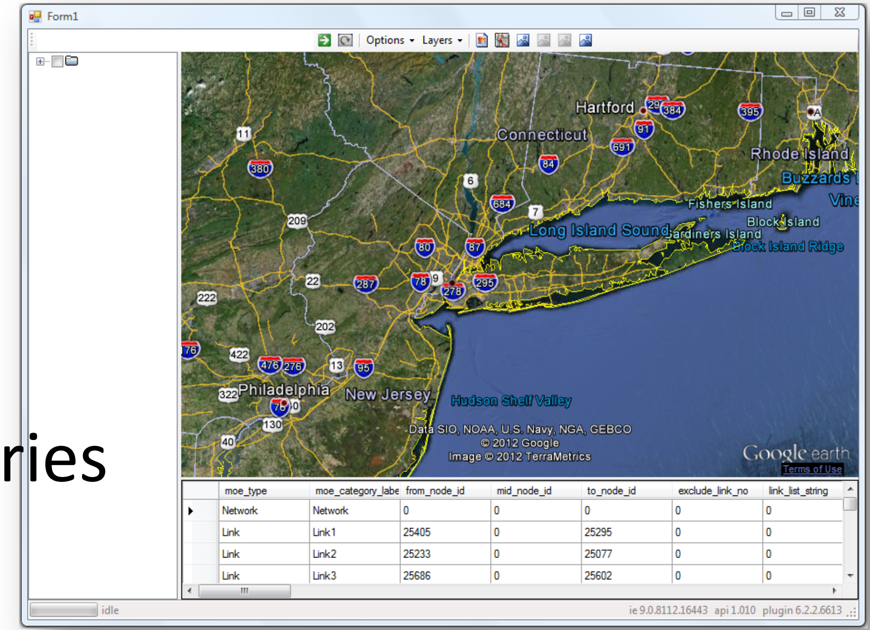
Main Role: Extract reliability-related measures from the vehicle trajectory output of the simulation models.



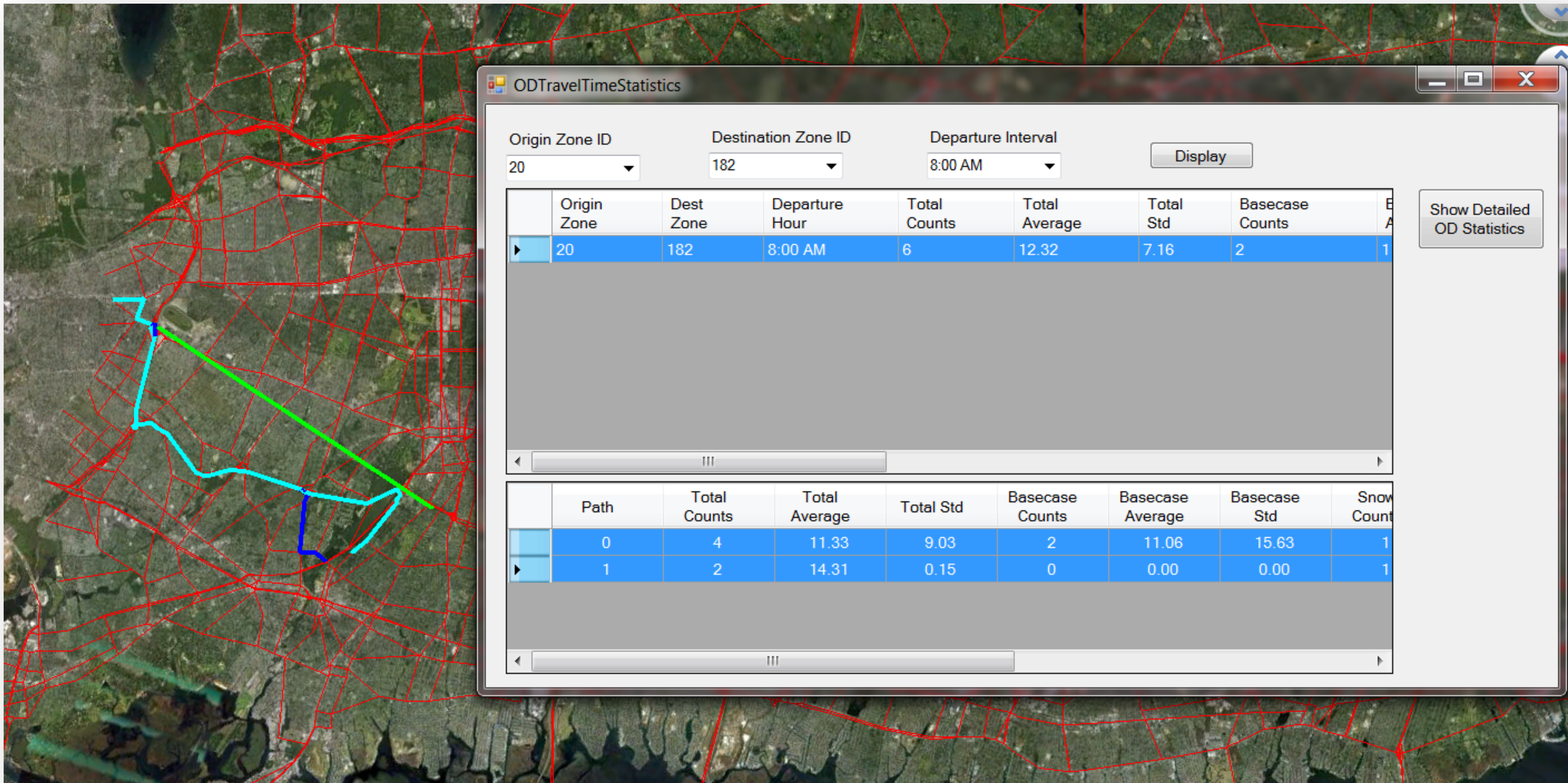
Application Main Functionalities

- Define critical O-D pairs, paths and links either via input text files of by selecting on the map
- From simulation outputs, for selected OD/path/links,
  - show scenario-specific travel time distributions
  - show combined travel time distribution (weighted by scenario probabilities)
  - extract various reliability performance measures
- Compare simulated trajectories with observed trajectories (e.g., TomTom GPS data)

↙ **Trajectory Processor**

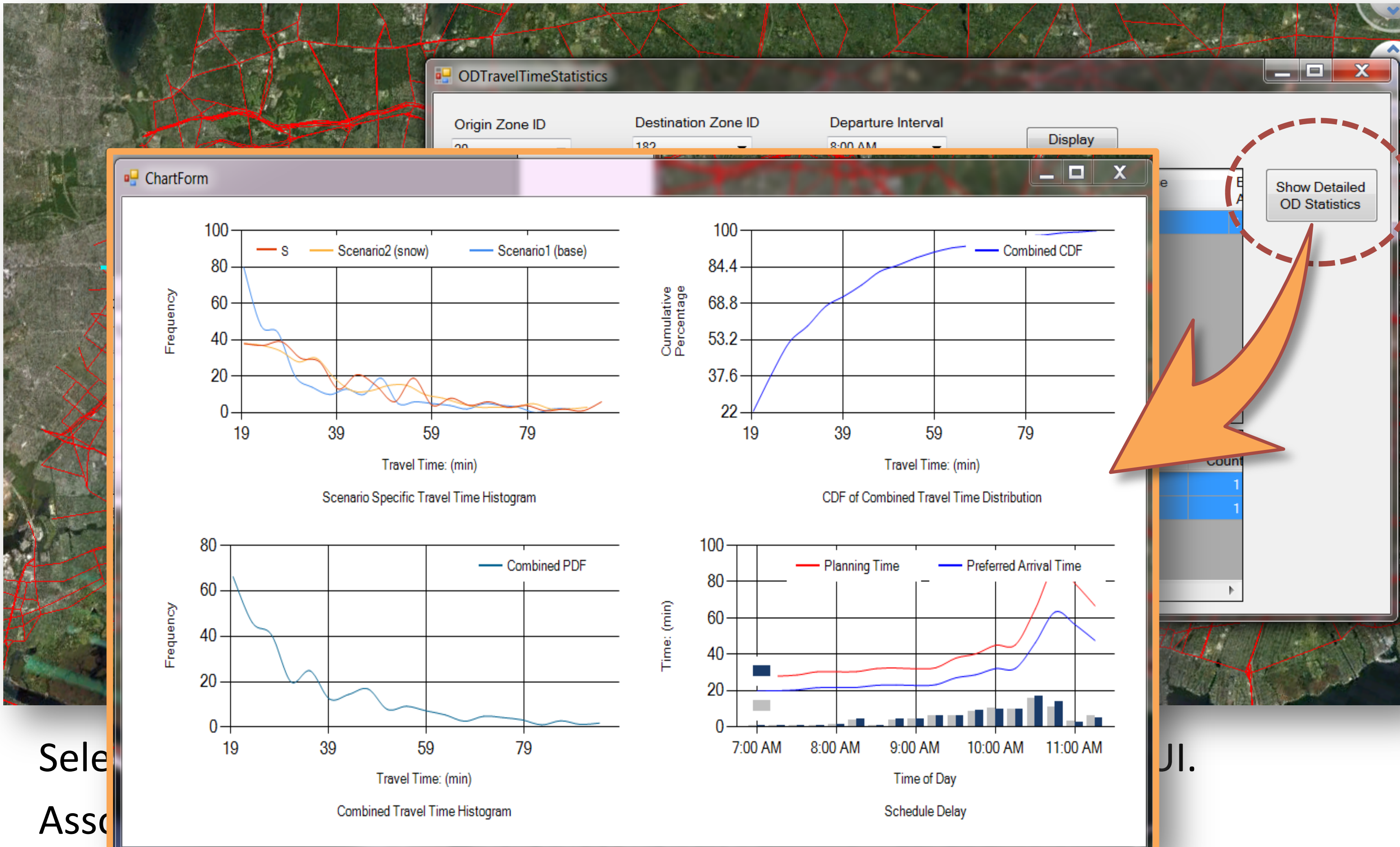


Select O-D pair from the map



- Select two points on the map by clicking on the Google Earth GUI.
- Associated O-D pair is identified (straight (green) line).
- Two alternative paths between the selected OD are identified (blue and light blue curves).

View travel time distributions for selected OD



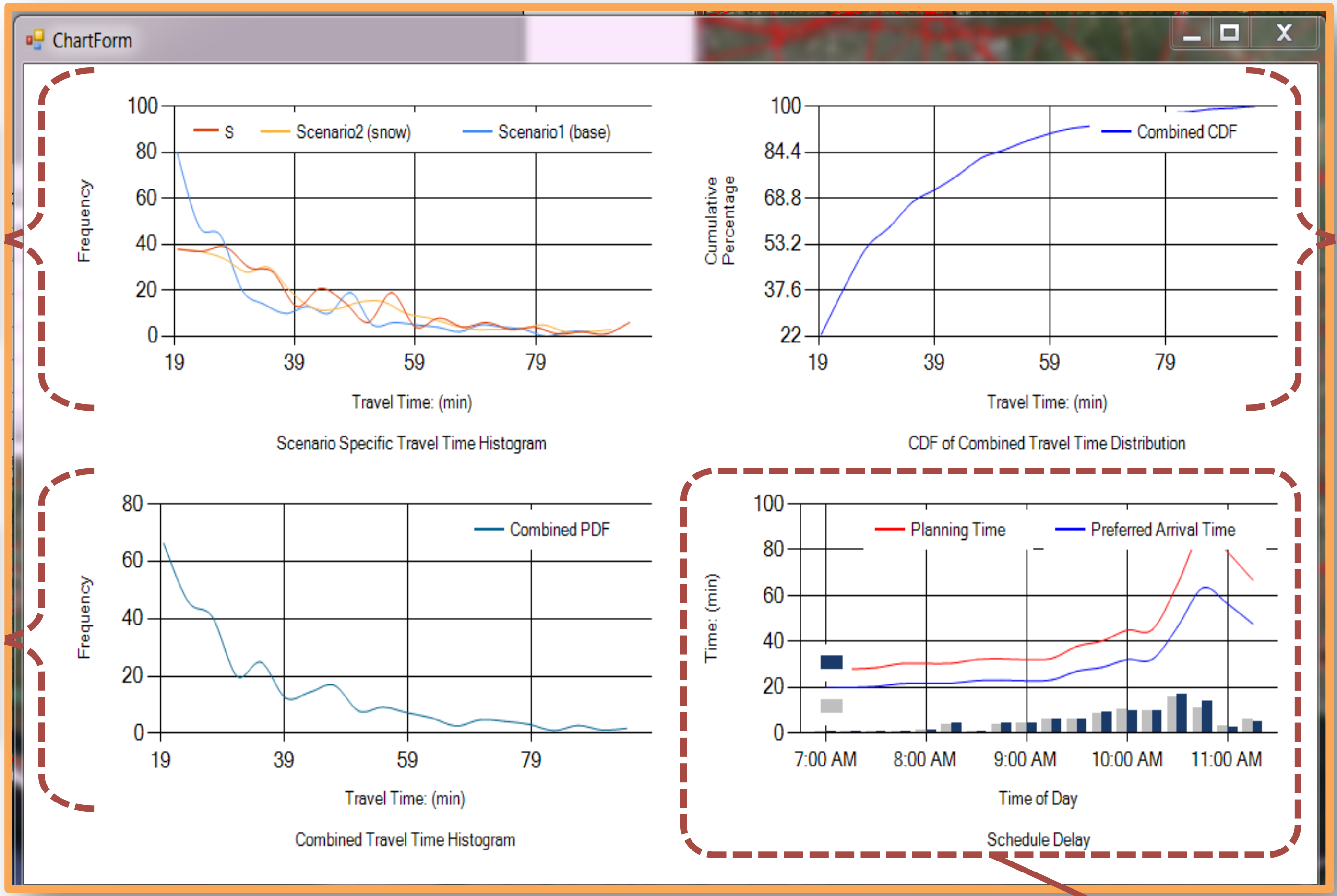
- Sele … UI.
- Ass …
- Two alternative paths between the selected OD are identified (blue and light blue curves).

Trajectory Processor

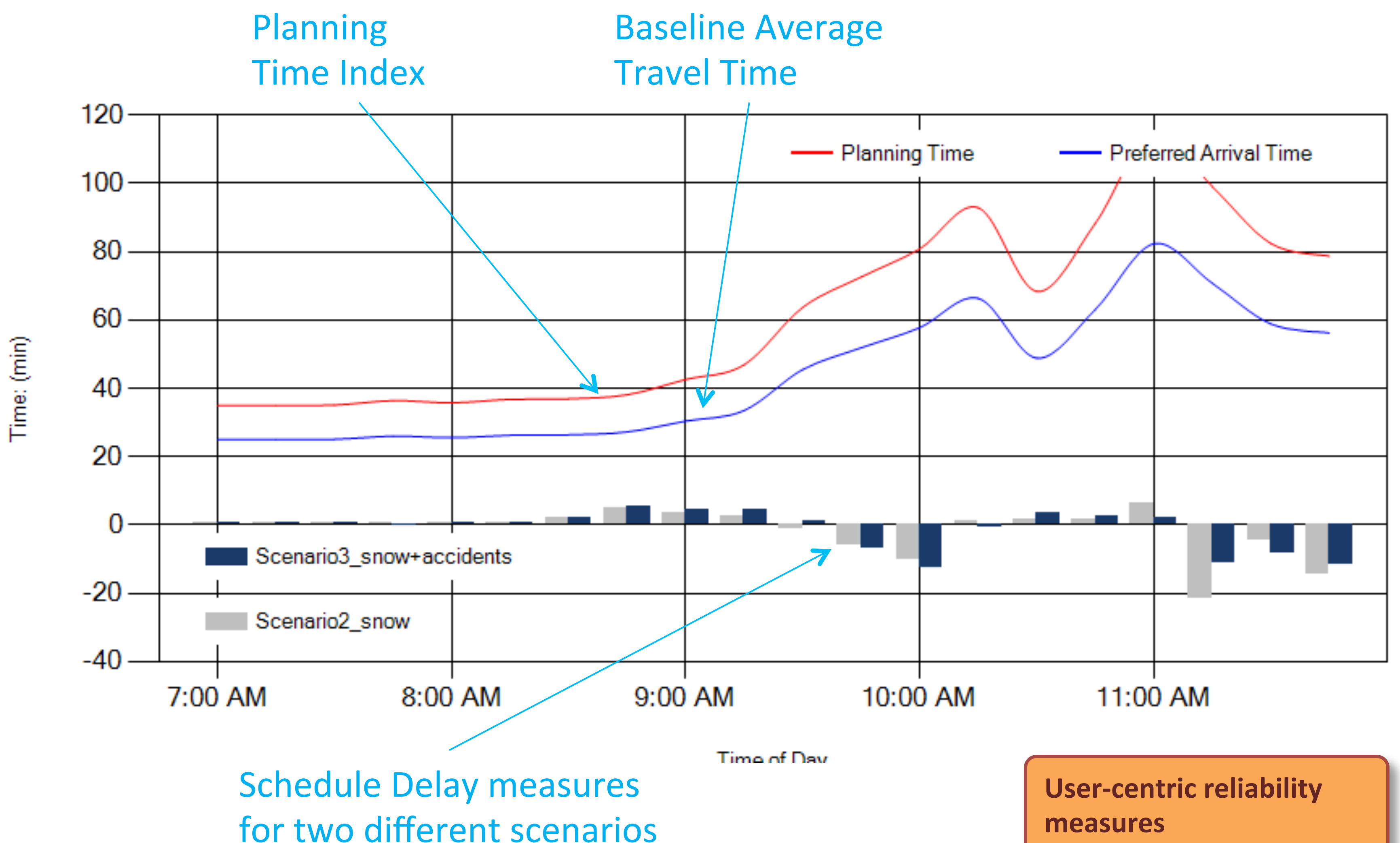
Scenario-specific travel time distribution (PDF)

Combined CDF

Combined travel time distribution (PDF)



User-centric reliability measures



User-centric reliability measures

2D Trajectories

Calibrate microscopic traffic relations (e.g. NGSIM)

Extract point measurements for mesoscopic and macroscopic model calibration and validation

3D Trajectories

MICROSCOPIC LEVEL

In addition to 2D capabilities, could model travel behavior choices

route choice; responses to information, pricing, controls

trip timing

Reliability characterization at individual vehicle level; both within day and across day variability

Ideal Revealed Preference manifestation

When based on smart phones or personal devices, not limited to car trips

When complete, allow capturing social influences and information

When coupled with transaction data enables wider range of behavioral responses

3D Trajectories

MESOSCOPIC and MACROSCOPIC

Retains ability to do same things as with point data

Time-dependent O-D demand estimation as input to dynamic network models

Takes validation to new levels, from patterns at points to spatio-temporal variation and user experience over entire travel; from facility-level to network-level.

Network-level relations

- Network fundamental diagram
- Reliability signature relations

Takeaways?

New era of trajectory-driven traffic and network performance analysis:

- More complete and compact description of system state
- Capture all aspects of individual actions (most complete record of actual behavior), with no loss of ability to characterize systems at any desired level of spatial and temporal aggregation/disaggregation
- Retain ability to extract stochastic properties of both individual behaviors and performance metrics
- Enable better model formulation/specification at all levels of resolution, and model calibration
- Most promising hope to recognize and capture collective effects and interaction mechanisms.

Limitations and Interesting Methodological Issues

- Partial trajectories
  - Censoring: incomplete trajectories of individual particles
  - Estimation of partially observed state variables: flow, density
  - Recognizing and correcting for selection bias
- Sampling trajectories
  - Sample designs and implications for estimation and model validation
- Non-uniqueness of underlying set of individual trajectories corresponding to observed aggregate (point-based) measurements
- Is it fair/reasonable to expect simulation tools to replicate individual vehicle trajectories, or only to replicate flow patterns in some aggregated fashion on links and
- Person trajectories, multimodal travel and activity-based modeling
- "Big data" aspects of very detailed trajectories
- ***Make trajectories your friend!***

THANK YOU!
QUESTIONS?

masmah@northwestern.edu