Preferred Citation for this Presentation

J. Hackney and M. Bernard (2005) A Spatial Regression Model of Traffic Speed in Zurich, IVT Seminar, ETH, December 2005.

A Spatial Regression Model of Traffic Speed in Zurich

Michael Bernard
Jeremy Hackney

IVT
ETH
Zürich

December 2005

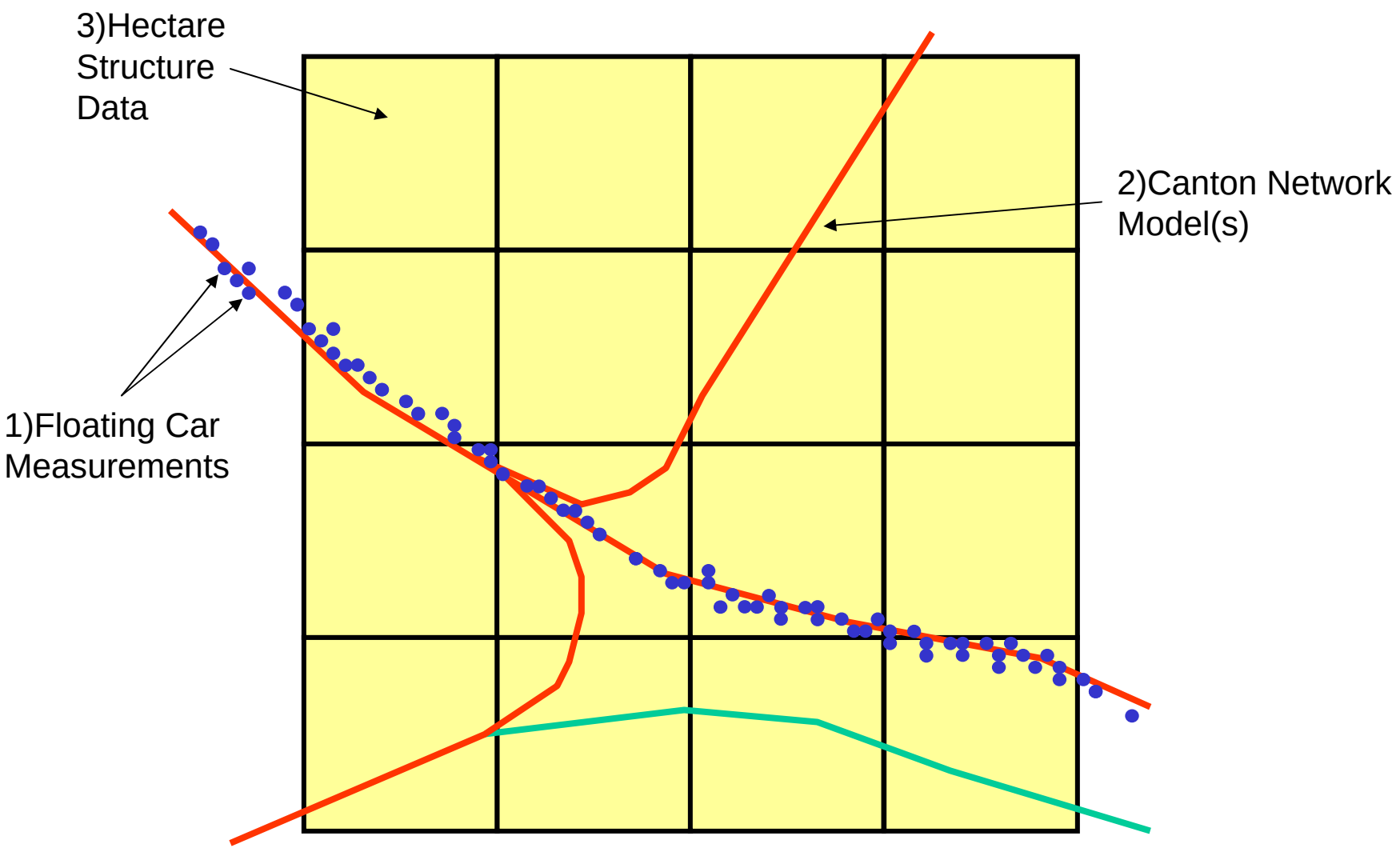
Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

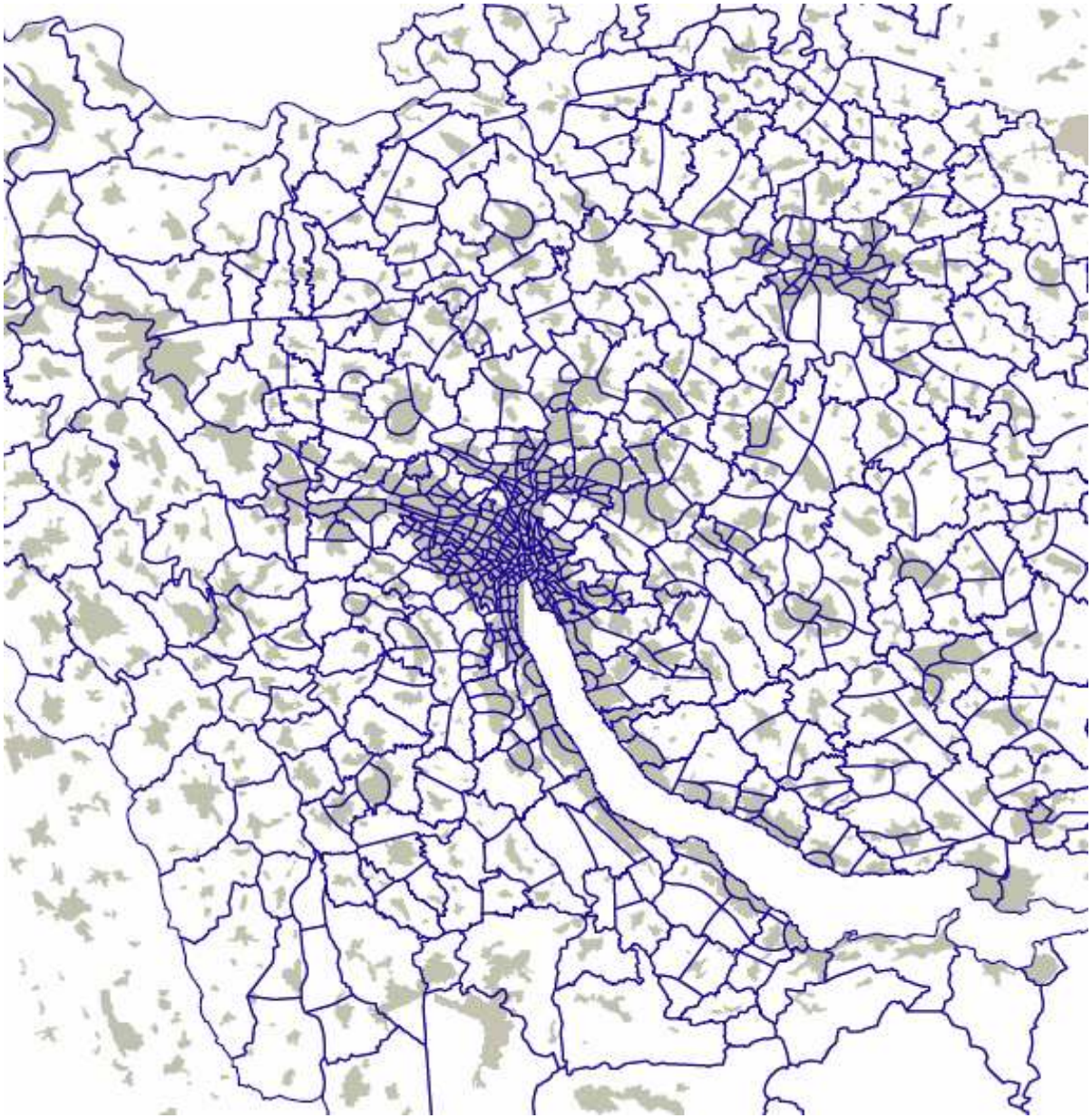
Summary of Presentation

1) GPS-based speed measurements (floating car)
2) Spatial structure data for Canton Zurich (Volkszählung BFS 1990) at hectare resolution
3) Regression of speeds on spatial structure variables
4) Observation of spatial autocorrelation of speeds
5) Methods for correction via spatial autoregression
6) First results of calculations of spatial autocorrelation term

Schematic of Spatial Structure and Speed Dataset



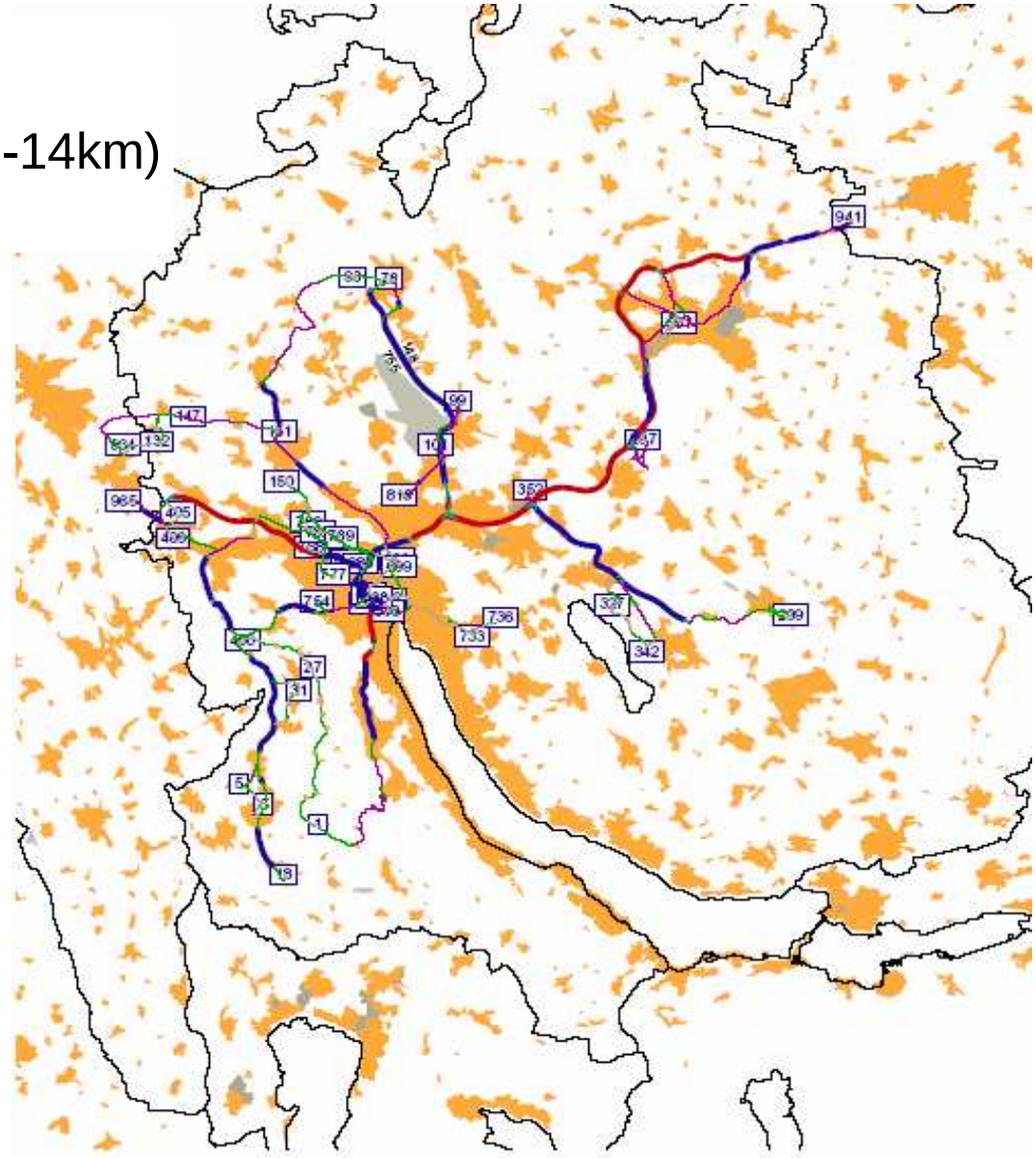
Traffic Zones in Canton Zürich: OD for Measurements



Circuit 1

550-700km

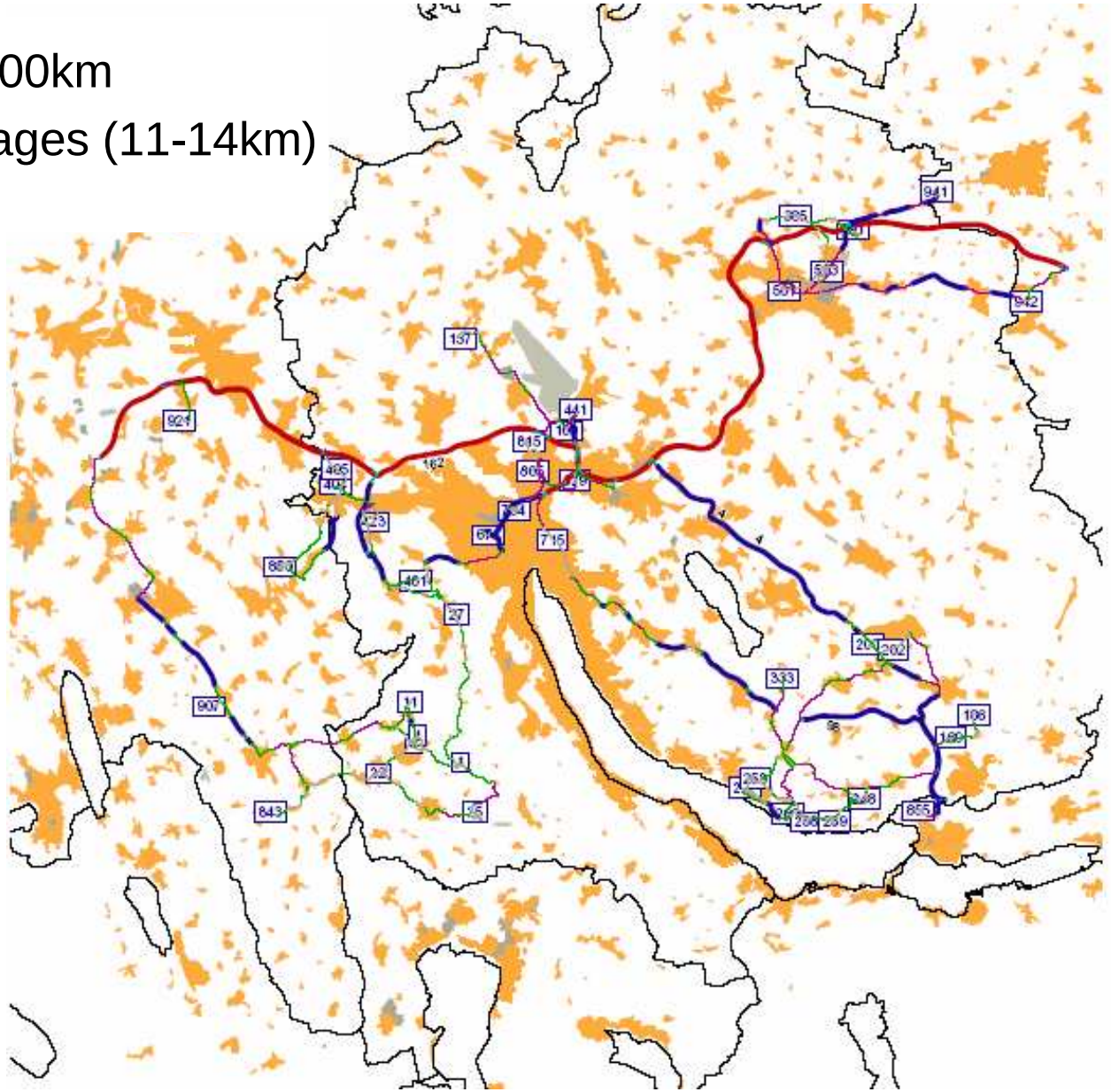
50 Stages (11-14km)



Circuit 2

550-700km

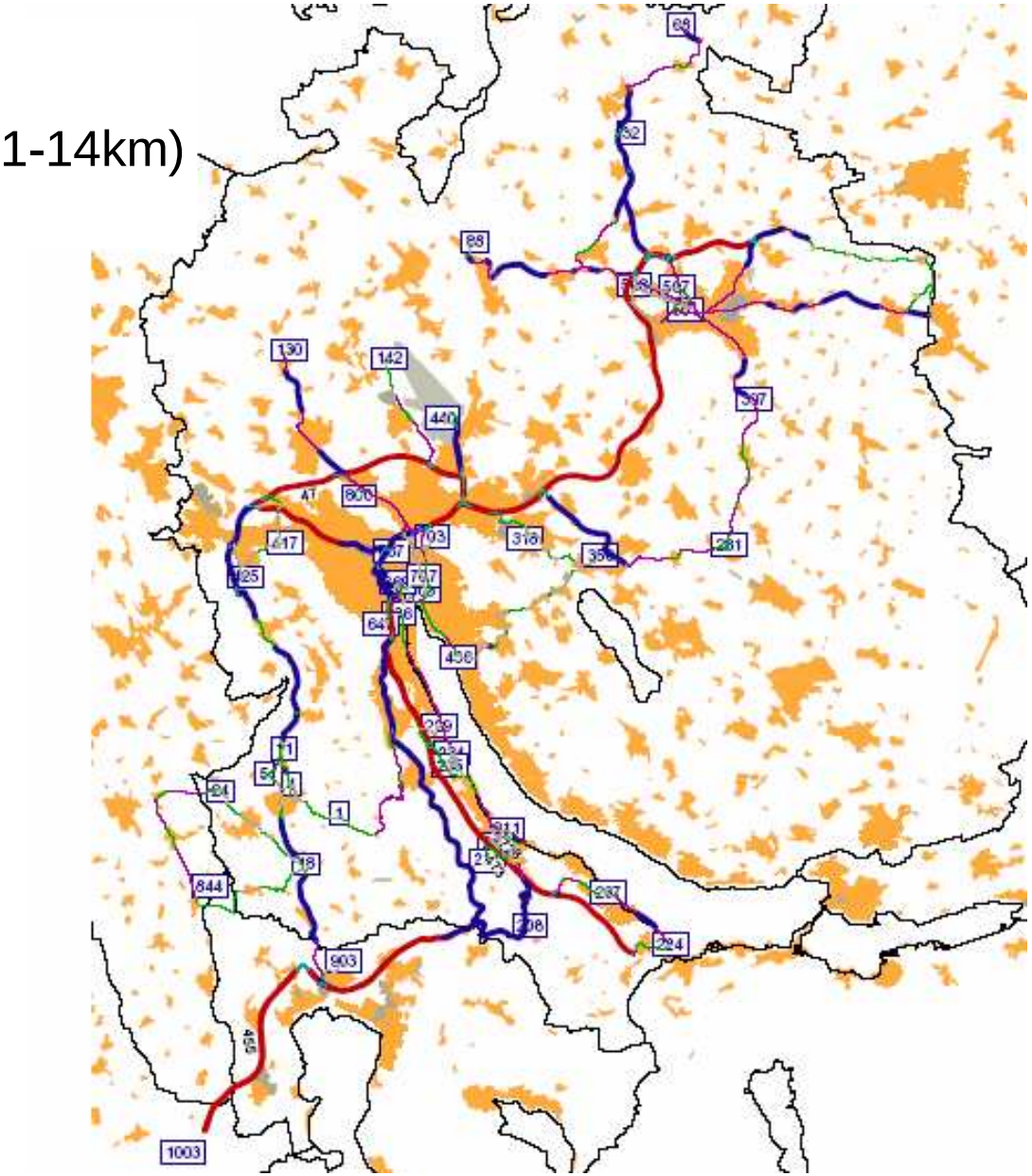
50 Stages (11-14km)



Circuit 3

550-700km

50 Stages (11-14km)



Raw GPS Data

GPS unit from Geostats records data every 1 second:

- Reception quality, speed, direction, altitude, longitude, latitude, date, time (GMT)

2.5 Mio. seconds of data on 18 days

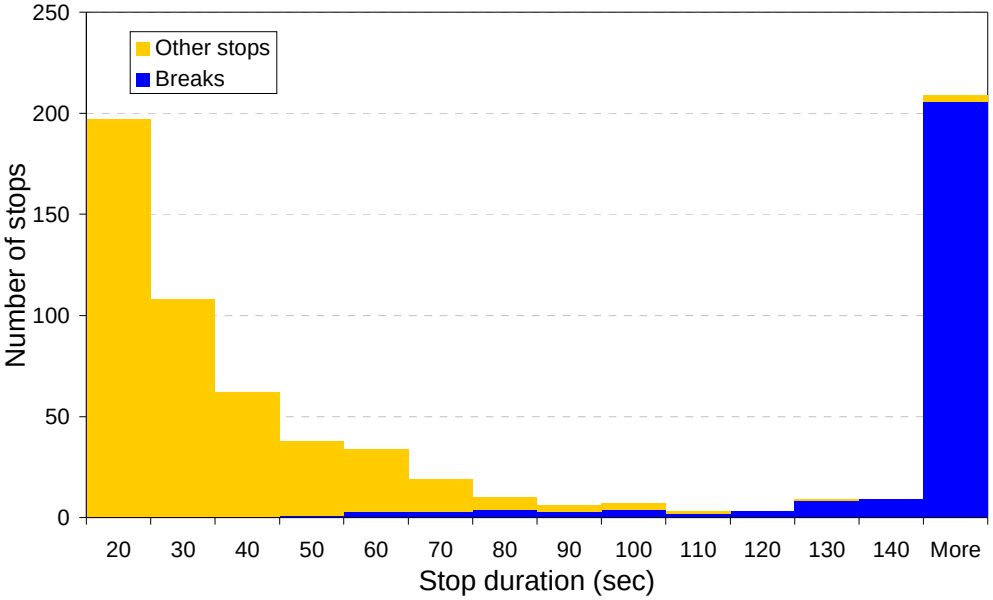
33,000 km

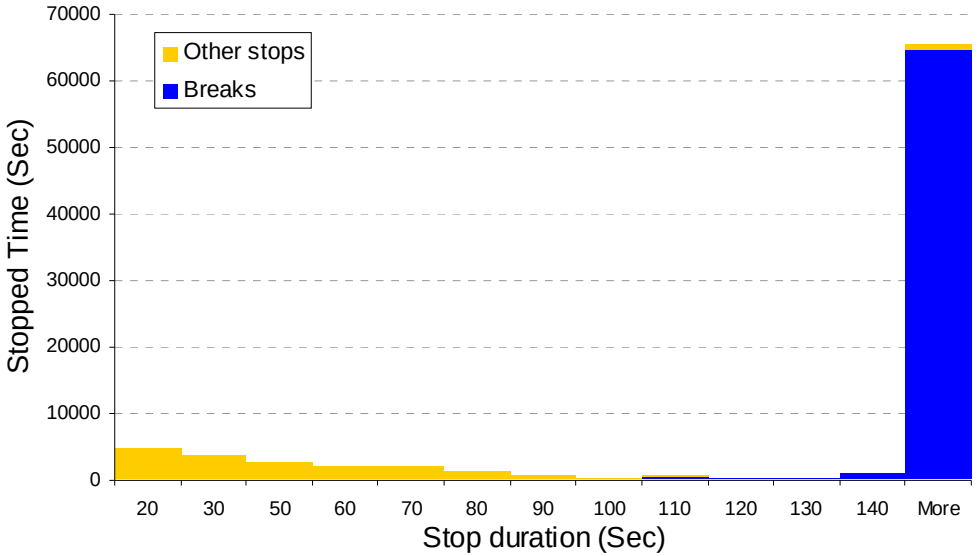
GPS system functioned extremely well

- Human error
- Reception interruptions

Ca. 33% of the GPS points (time) show speed = 0!

Cleaning the GPS Data





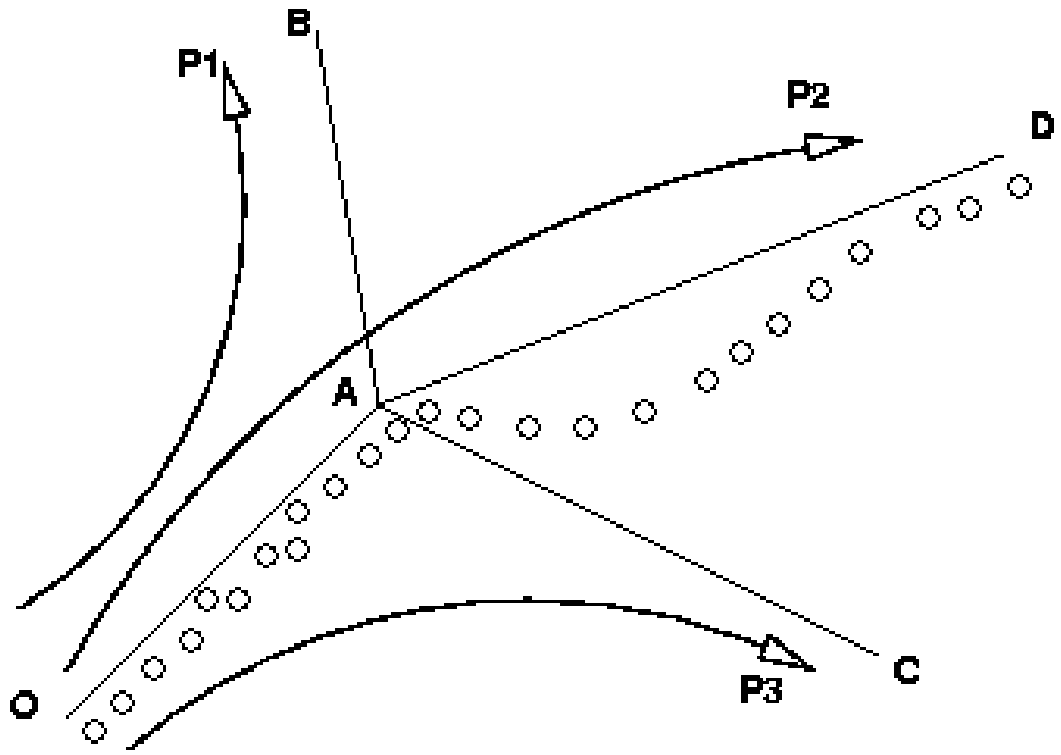
A sample of driver log books shows stops > 100 sec are driver breaks

Remove all series of GPS points with speed = 0 for longer than 100 sec.

= 5% of stops, 61% of stopped time

Stops between 0-10 seconds are omitted for clarity due to their large number

Matching GPS Data to Network



Marchal, Hackney and Axhausen (Forthcoming)

Calculation of Average Speed per Link

The matched GPS points give entry and exit times for each link

$$\bar{t} = t_{exit} - t_{enter}$$

The speed is the length of the link divided by the time to drive across it.

Space Mean Speed for N measurements is

$$\overline{V} = \frac{(x_1 + x_2 + x_3 + \Lambda + x_n)}{(t_1 + t_2 + t_3 + \Lambda + t_n)}$$

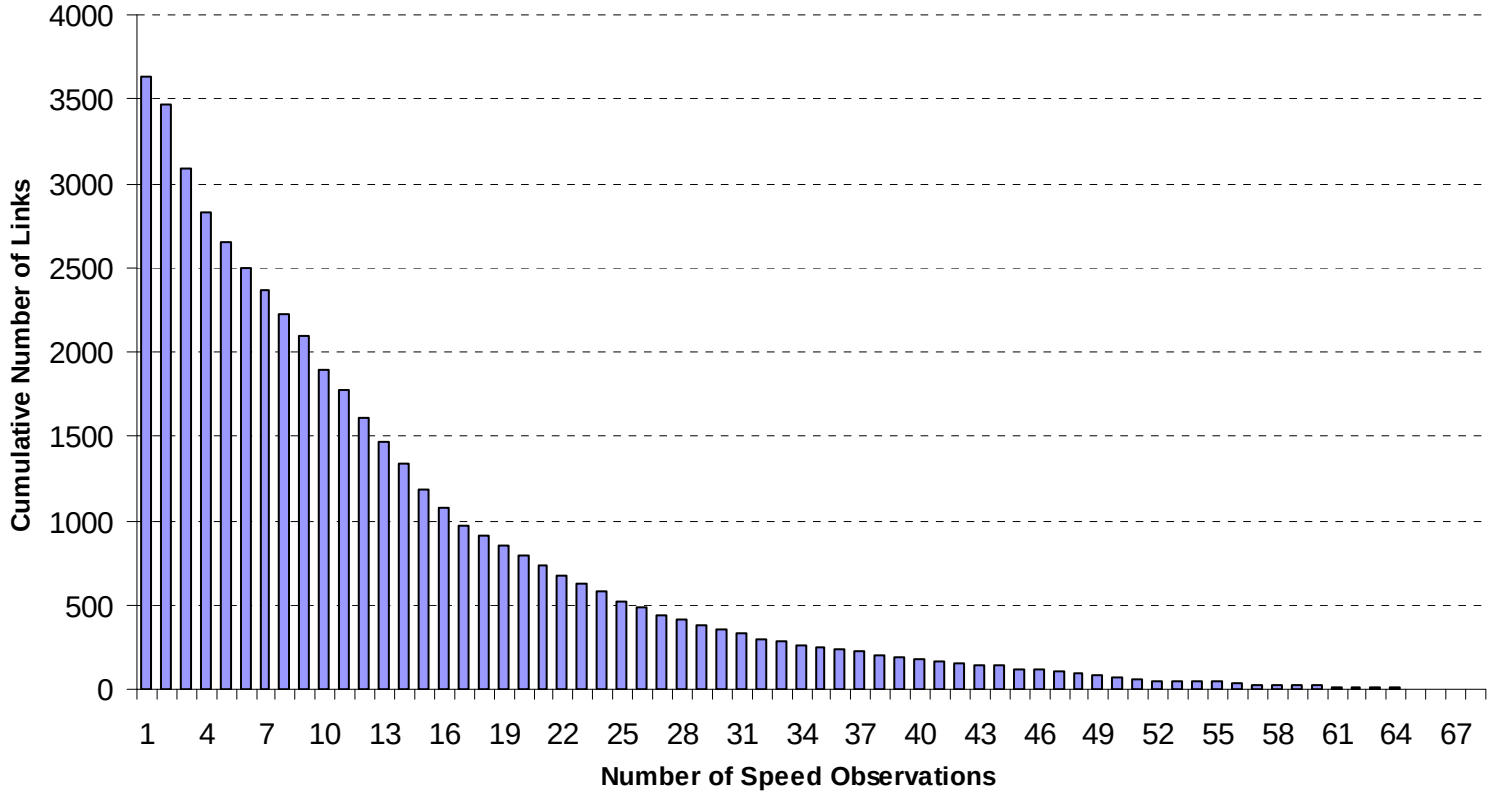
where x is link length and n is the observation (trip across the link)

Data Yield at Link Level

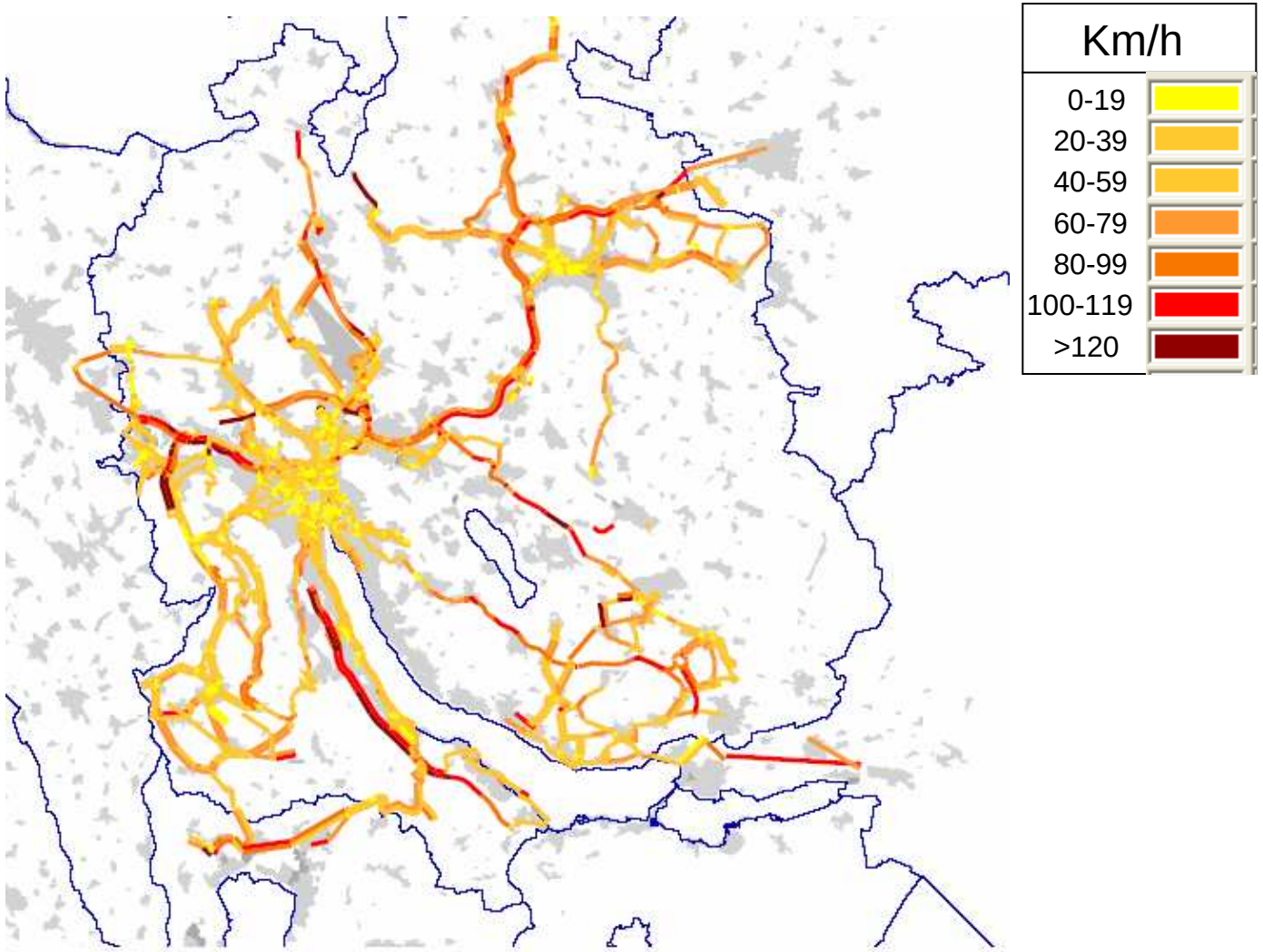
52,000 speed observations

3683 links measured

- > 16 observations for 1072 links
- 15 observations per link on average

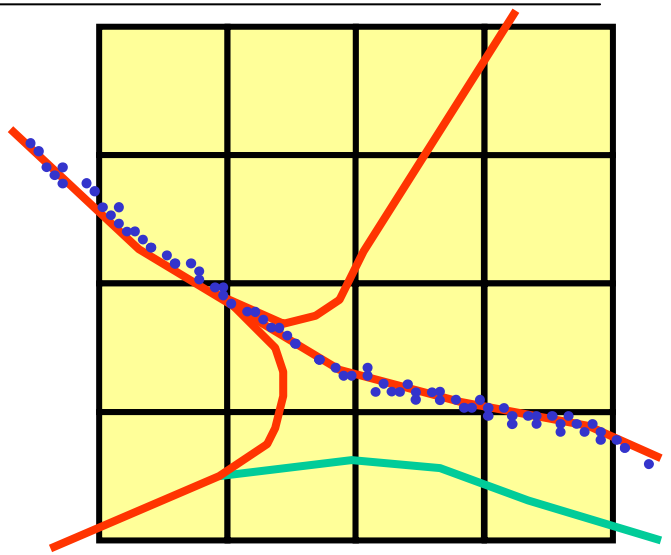


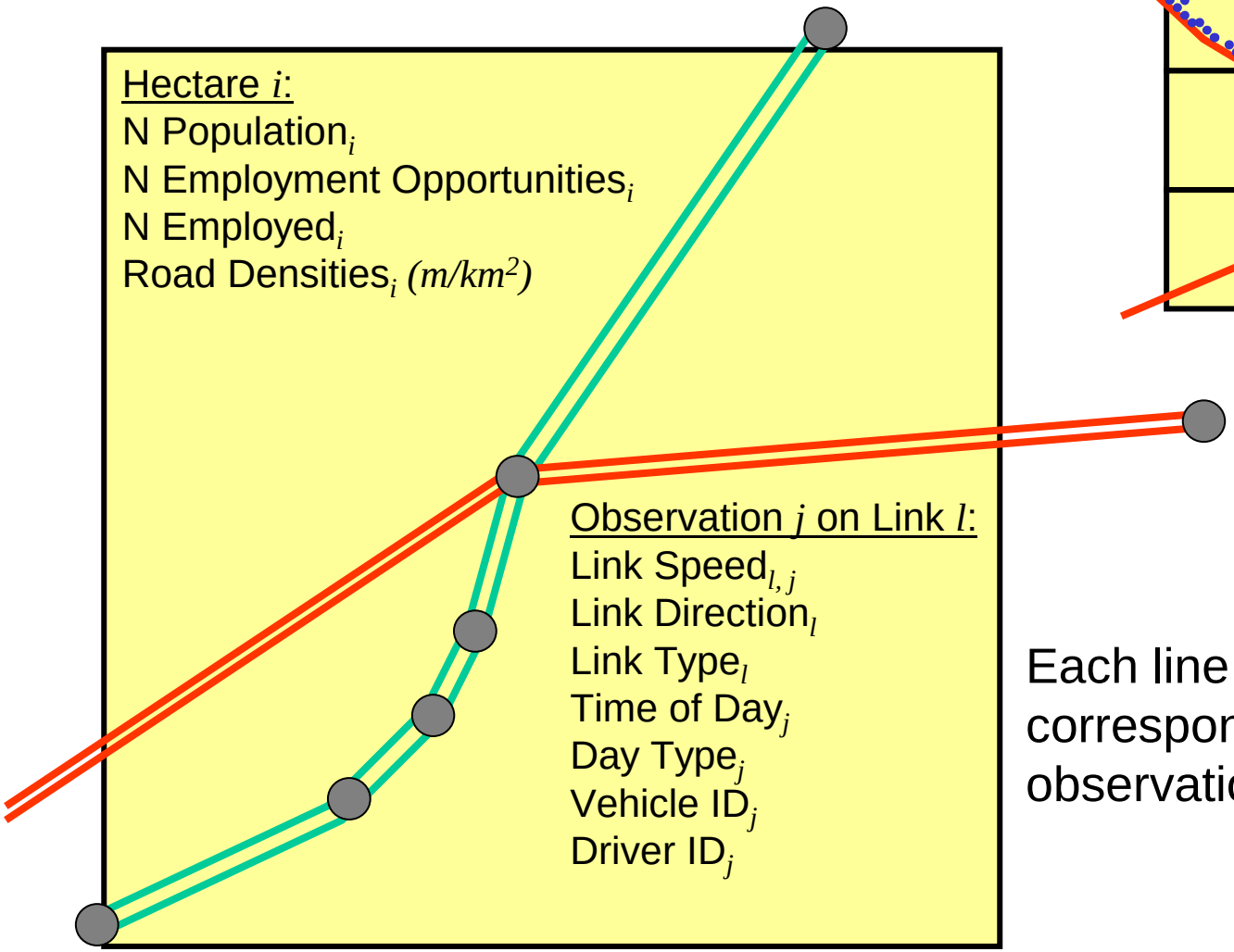
Average Speeds During Weekday Peak Time



Structure Dataset: Hectare Data and Links

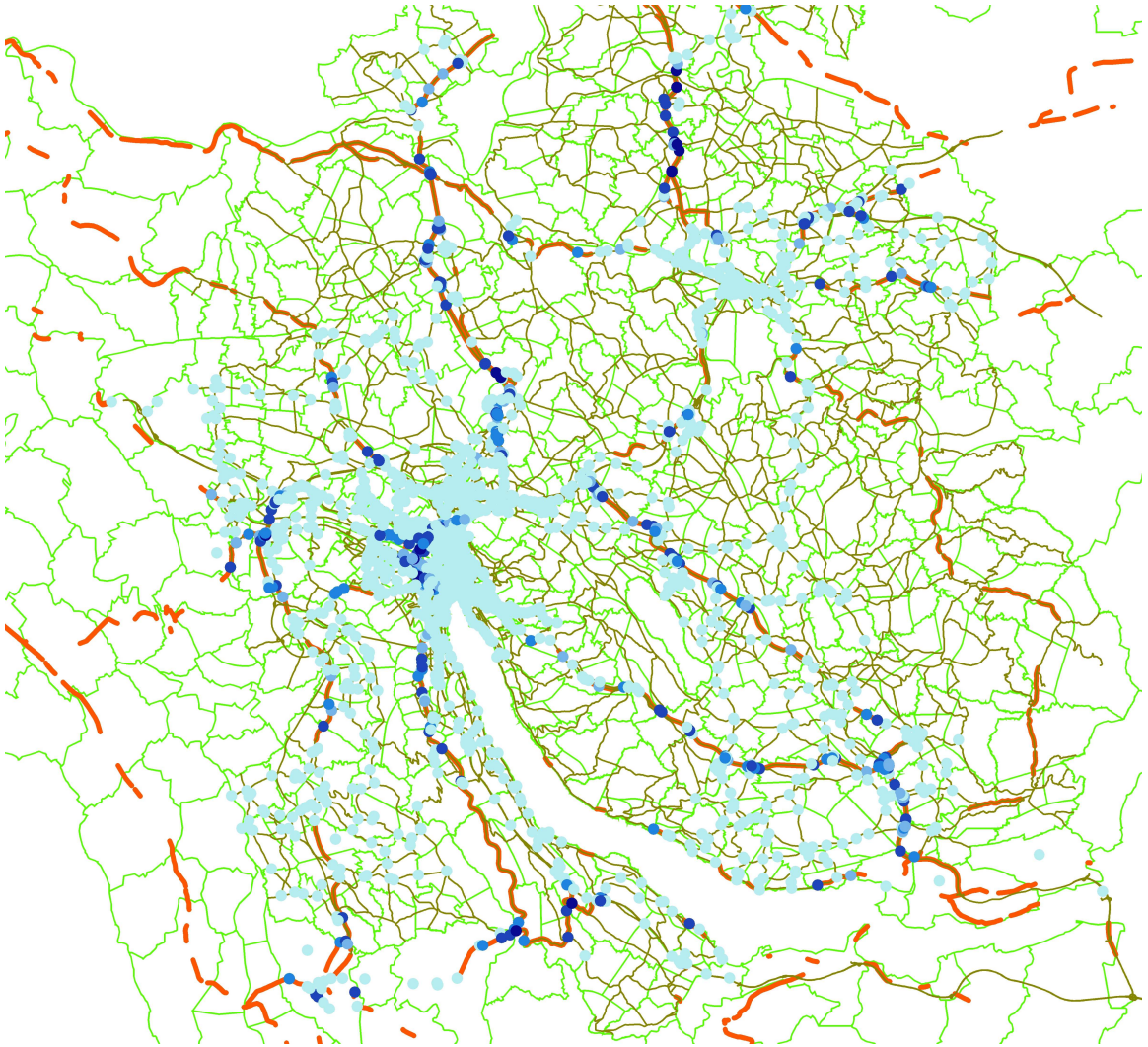
Hectare data is matched to link endpoints in GIS





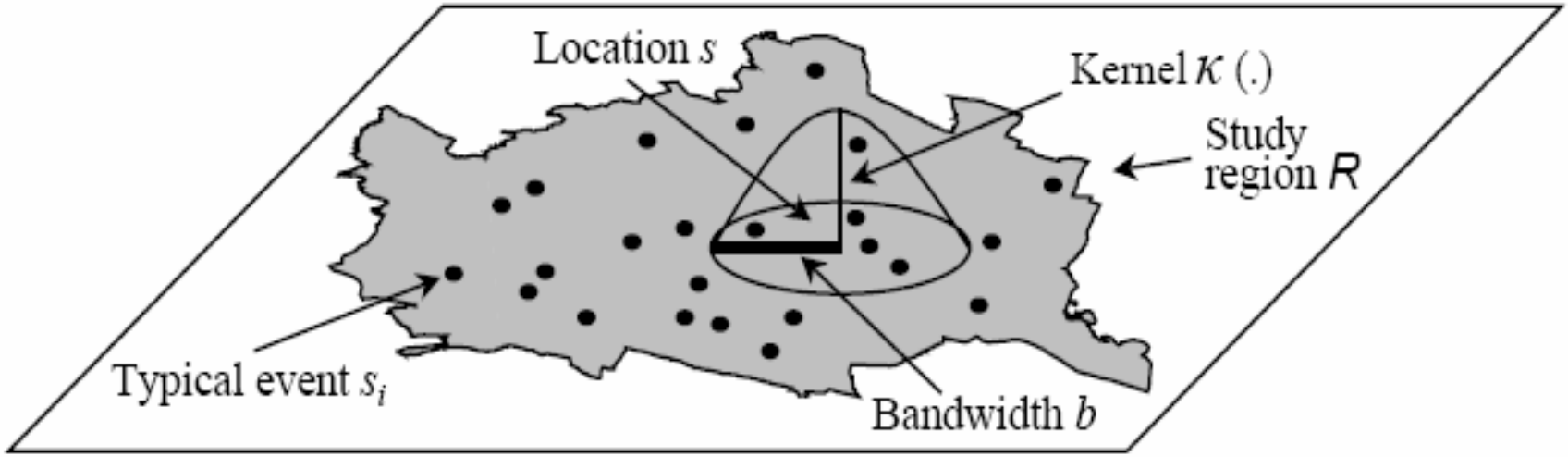
Each line in final dataset corresponds to a speed observation j on link l

Density of High Speed Roads, m/km^2 per Hectare



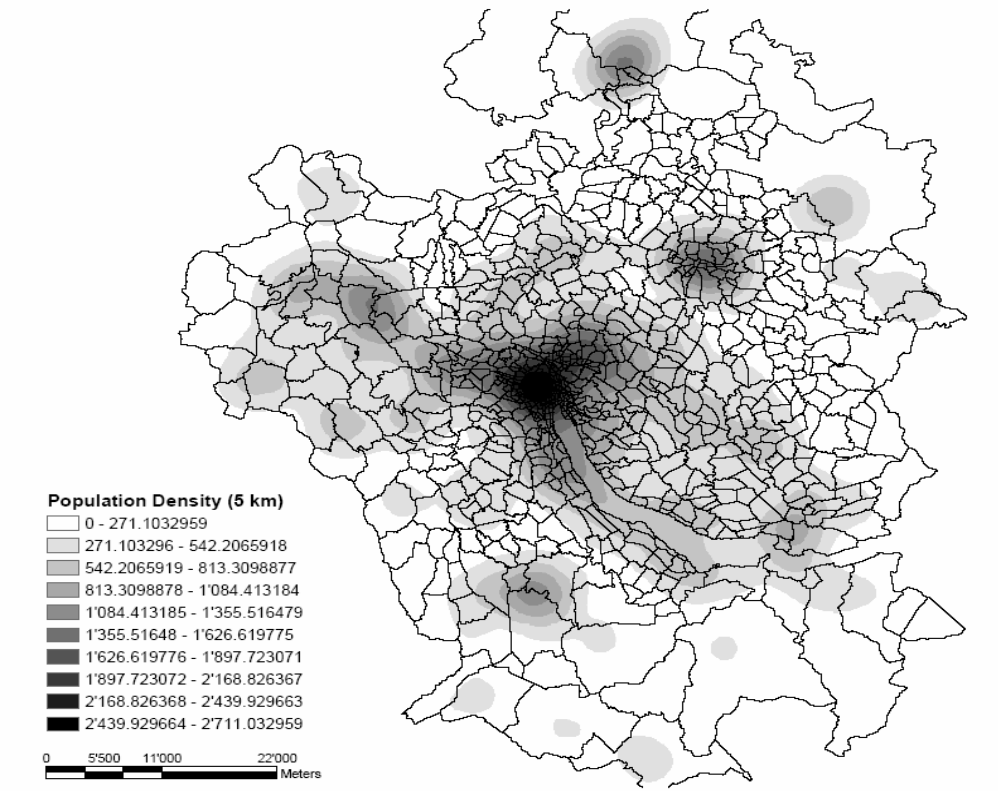
Kernel Density Averaging of Hectare Data

Radii of influence of 1km, 3km, 5km

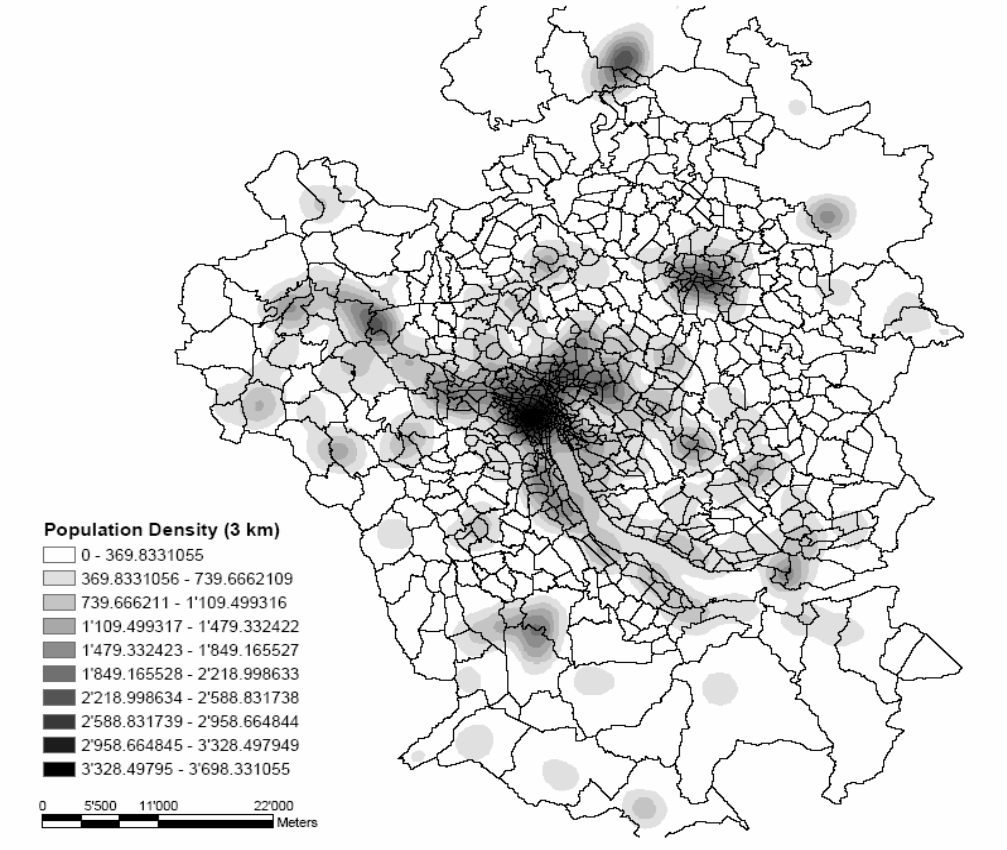


Bailey and Gatrell (1995)

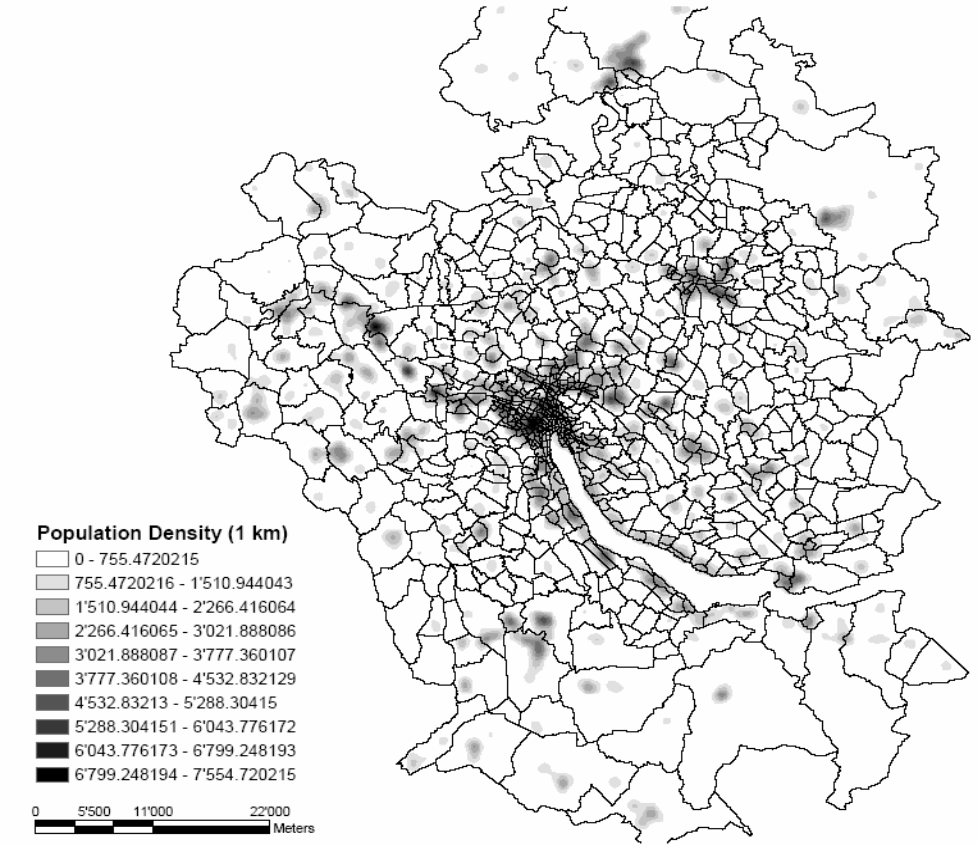
Kernel Density Averaged Population, r=5km



Kernel Density Averaged Population, r=3km



Kernel Density Averaged Population, r=1km



Linear Regression of Speed on Structure Variables R^2=0.5296

Variable	Coefficient	t-statistic	t-probability
const_1	60.53	77.54	0.00
hvz_and_HLS	44.20	23.99	0.00
hvz_and_SS	2.64	1.43	0.15
hvz_and_ES	-12.38	-5.49	0.00
hvz_and_otherRoad	-16.89	-7.31	0.00
rvz_and_HLS	56.11	27.56	0.00
rvz_and_HVS	6.06	7.10	0.00
rvz_and_SS	3.83	1.79	0.07
rvz_and_ES	-6.95	-2.72	0.01
rvz_and_otherRoad	-12.98	-3.95	0.00
nvz_and_HLS	53.10	29.82	0.00
nvz_and_HVS	2.41	3.24	0.00
nvz_and_SS	5.55	3.19	0.00
nvz_and_ES	-10.84	-4.98	0.00
nvz_and_otherRoad	-14.57	-6.76	0.00
eo3_and_HLS	-0.10	-2.86	0.00
eo3_and_HVS	-0.04	-8.16	0.00
eo3_and_SS	-0.07	-1.29	0.20
eo3_and_ES	-0.04	-3.60	0.00
eo3_and_otherRoad	-0.05	-1.67	0.10
p3_and_HLS	0.00	0.60	0.55
p3_and_HVS	-0.01	-9.64	0.00
p3_and_SS	-0.01	-1.63	0.10
p3_and_ES	0.00	-2.80	0.01
p3_and_otherRoad	0.00	-0.09	0.93

Correlations (1) Selected Measurements

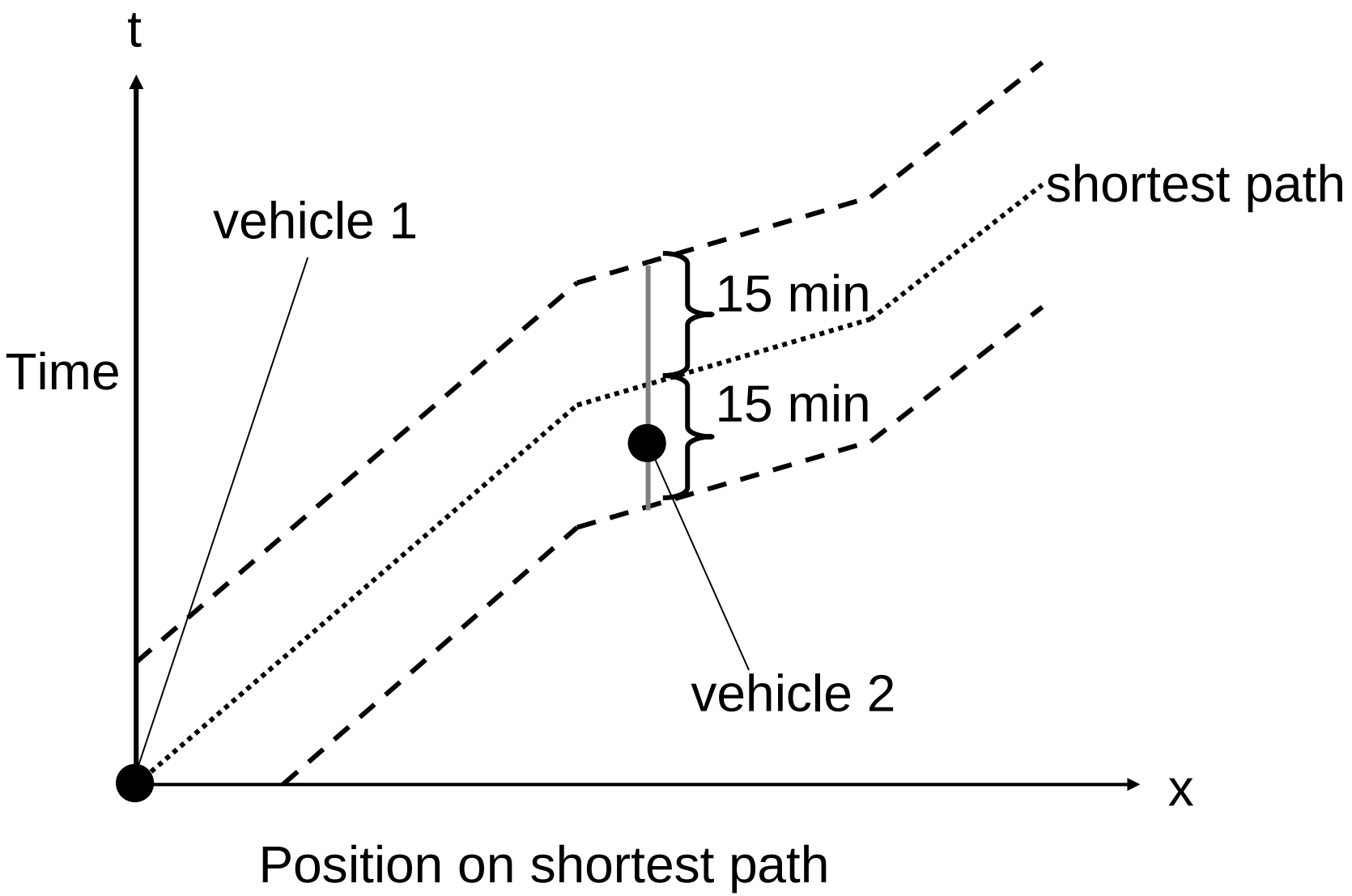
Chosen pairs of link speeds:

- Same road type (HLS, HVS, SS, ES, other roads)
- Network distance within class (1-500, 501-1000, … m)
- Difference of estimated network travel time and time gap of measurements < 15 min

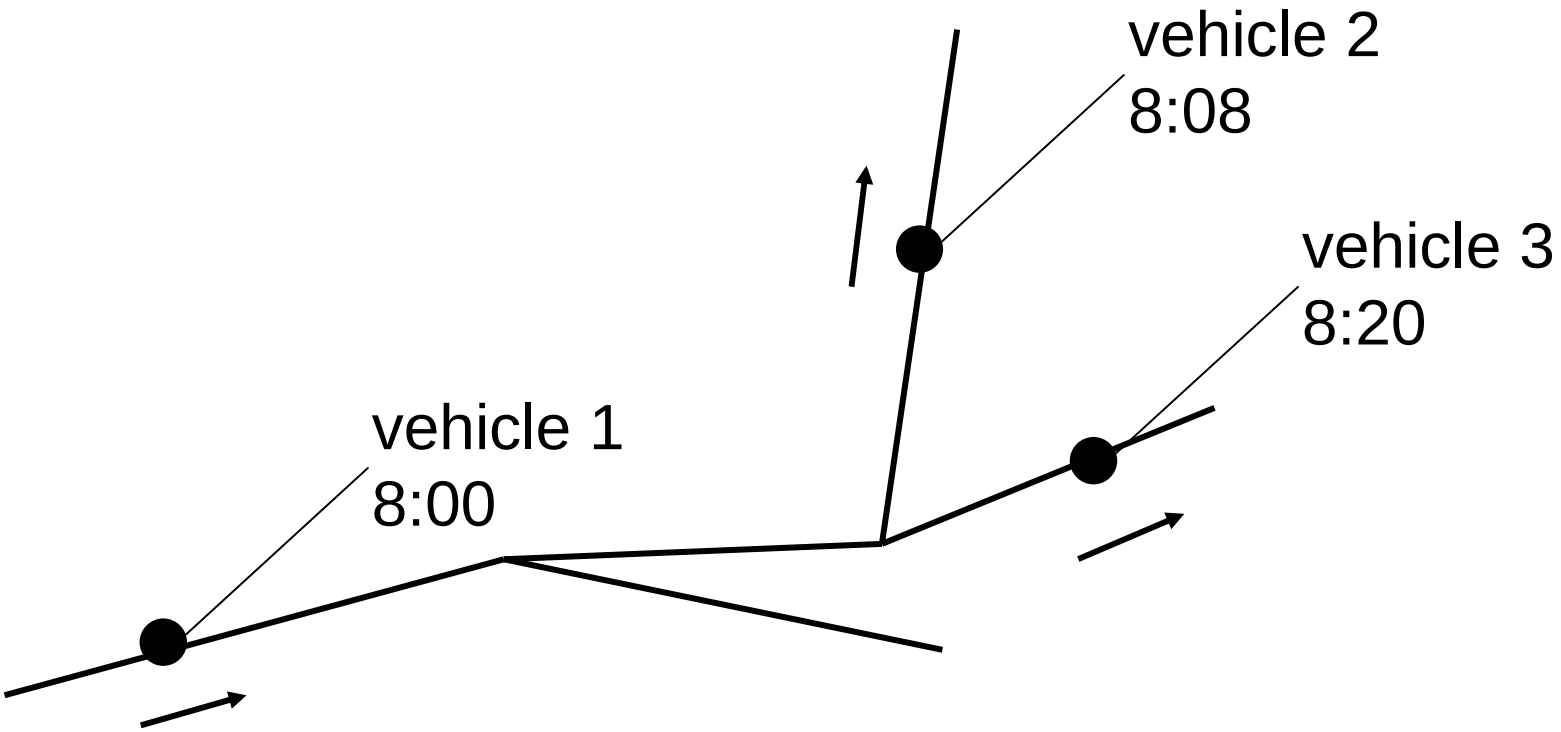
Correlation:

$$\rho = \frac{1}{N} \frac{\sum_{i=1}^{N} (v_{1i} - \overline{v_1})(v_{2i} - \overline{v_2})}{\sigma_{v1}\sigma_{v2}}$$

Correlations (2) Difference: measurements and traveltime $|\Delta t - tt|$



Correlations (3) Example valid measurments



Shortest path: tt(1→2) = 10 min; t_2 = 8:10 +- 15 min
tt(1→3) = 12 min; t_3 = 8:12 +- 15 min

Correlations (4) Example of Algorighm

• Distance class 500 – 1000 m:

Veh	Time	Link	V	Type	tt	|Δt-tt|	ok?
A	8:00	1	50	HVS			
A	8:05	2	60	HVS	tt_{1-2}= 1 min	| 4|	ok
B	8:10	3	55	HVS	tt_{1-3}=12 min	|-2|	ok
B	8:12	7	70	***HLS***			-
B	8:18	9	65	HVS	tt_{1-9}= 2 min	|16|	-
	…						

Condition:

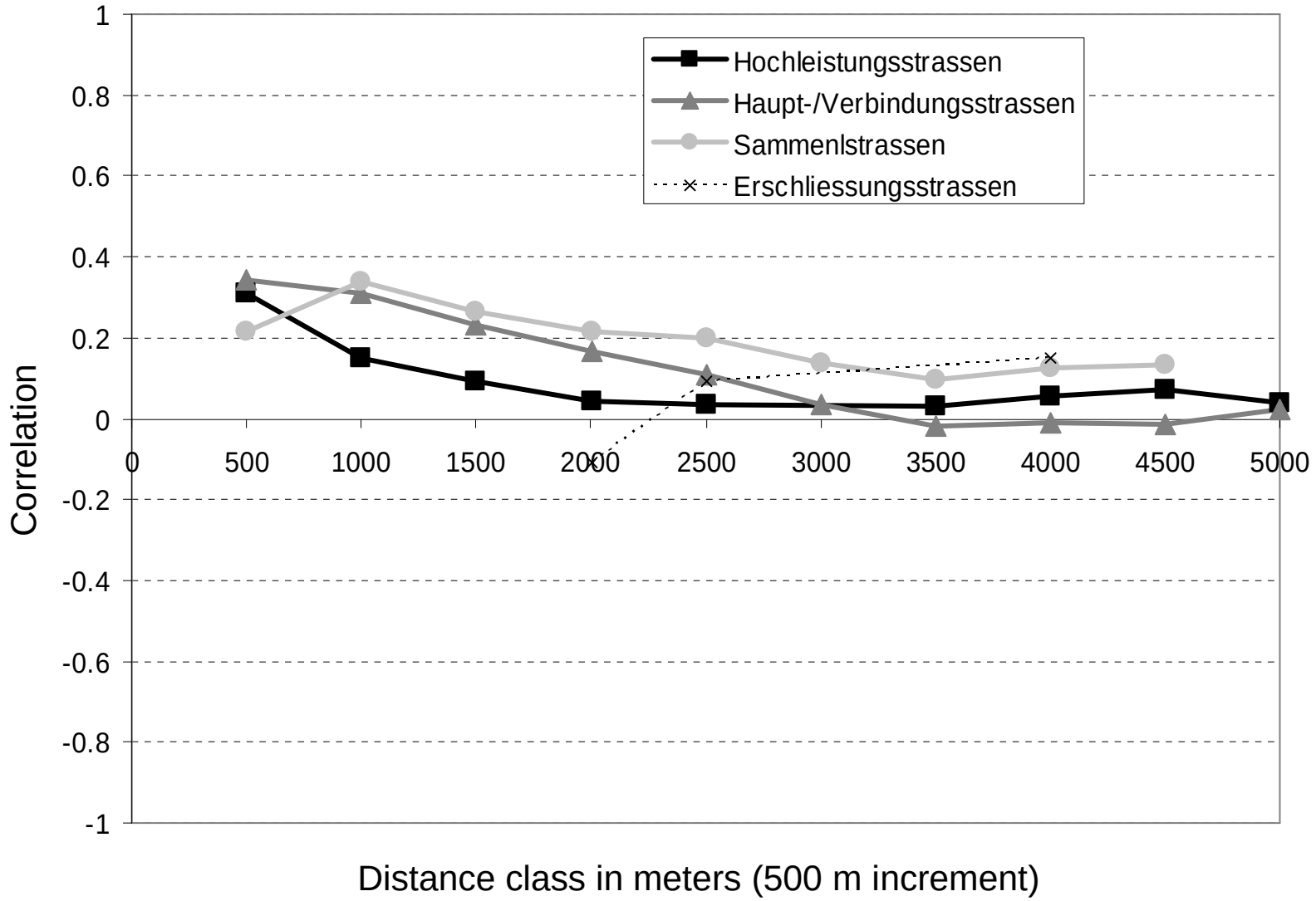
same roadtype
same distance class
$|\Delta t - tt| < 15$ min

valid pairs:

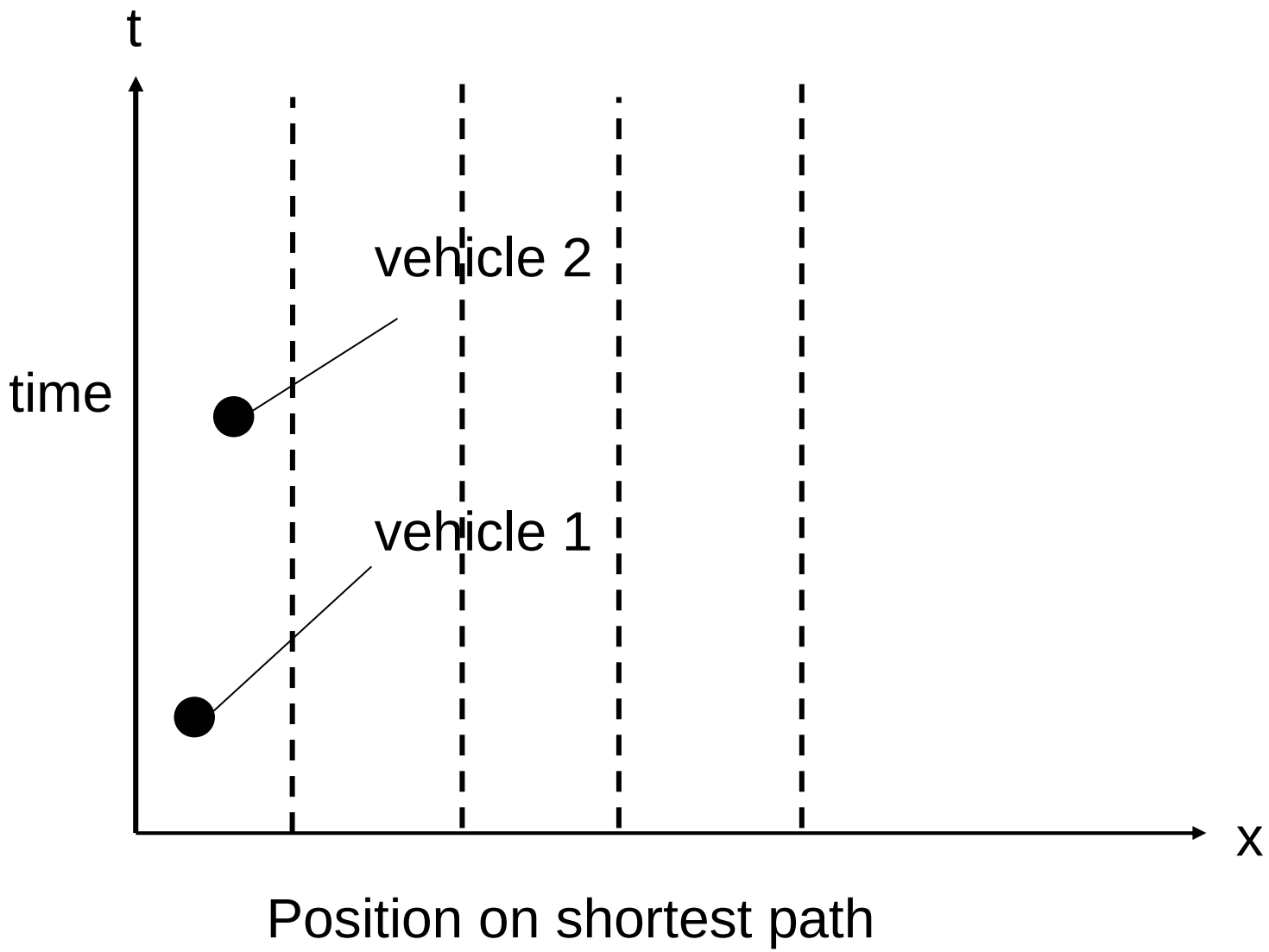
V1	V2	Type
50	60	HVS
50	55	HVS
60	…	
…		

Calculation of correlations

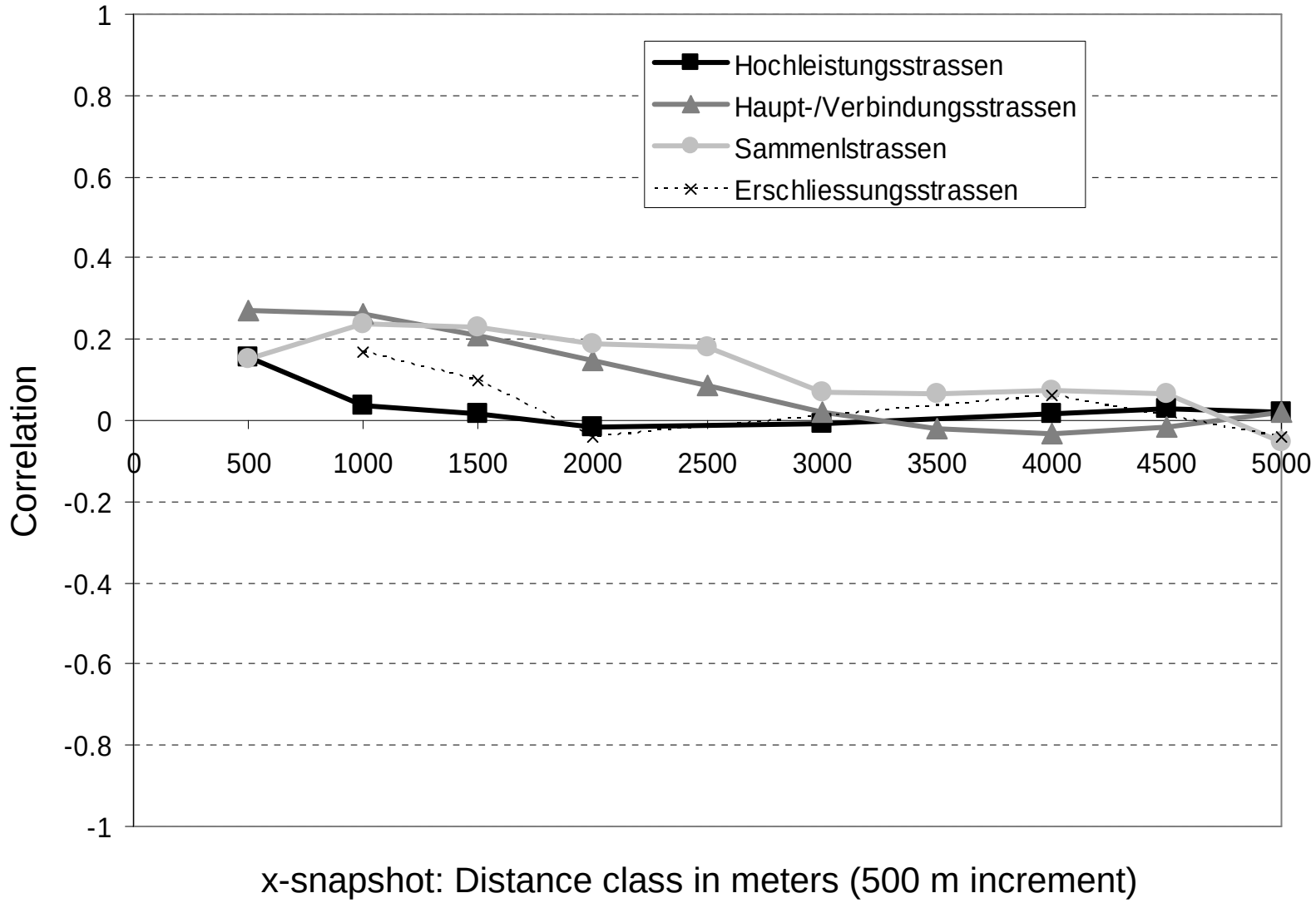
Correlation (5) weekday, HVZ: 6:30-8:30, 16:30-18:30



Correlation (8) weekday, local snapshot



Correlation (9) weekday, local snapshot, HVZ



Spatial Error Model (1)

Ordinary Least Sqares model (OLS):

$$y = \beta X + \varepsilon \qquad \varepsilon \sim N(0, \sigma)$$

Spatial Error Model (SEM):

$$y = \beta X + u$$

$$u = \lambda W u + \varepsilon \qquad \varepsilon \sim N(0, \sigma)$$

with W: contiguity matrix

λ: influence factor of spatial dependence

Spatial Error Model (2) Contiguity matrix W

$$y_i = \sum_{k=1}^{K} \beta_k X_{ik} + u_i$$

$$u_i = \lambda \sum_{j=1}^{N} W_{ij} u_j + \varepsilon_i$$

with K: number of explanatory variables

with N: number of observations → W is a (N x N)-matrix

W describes the influence of spatial neighbours

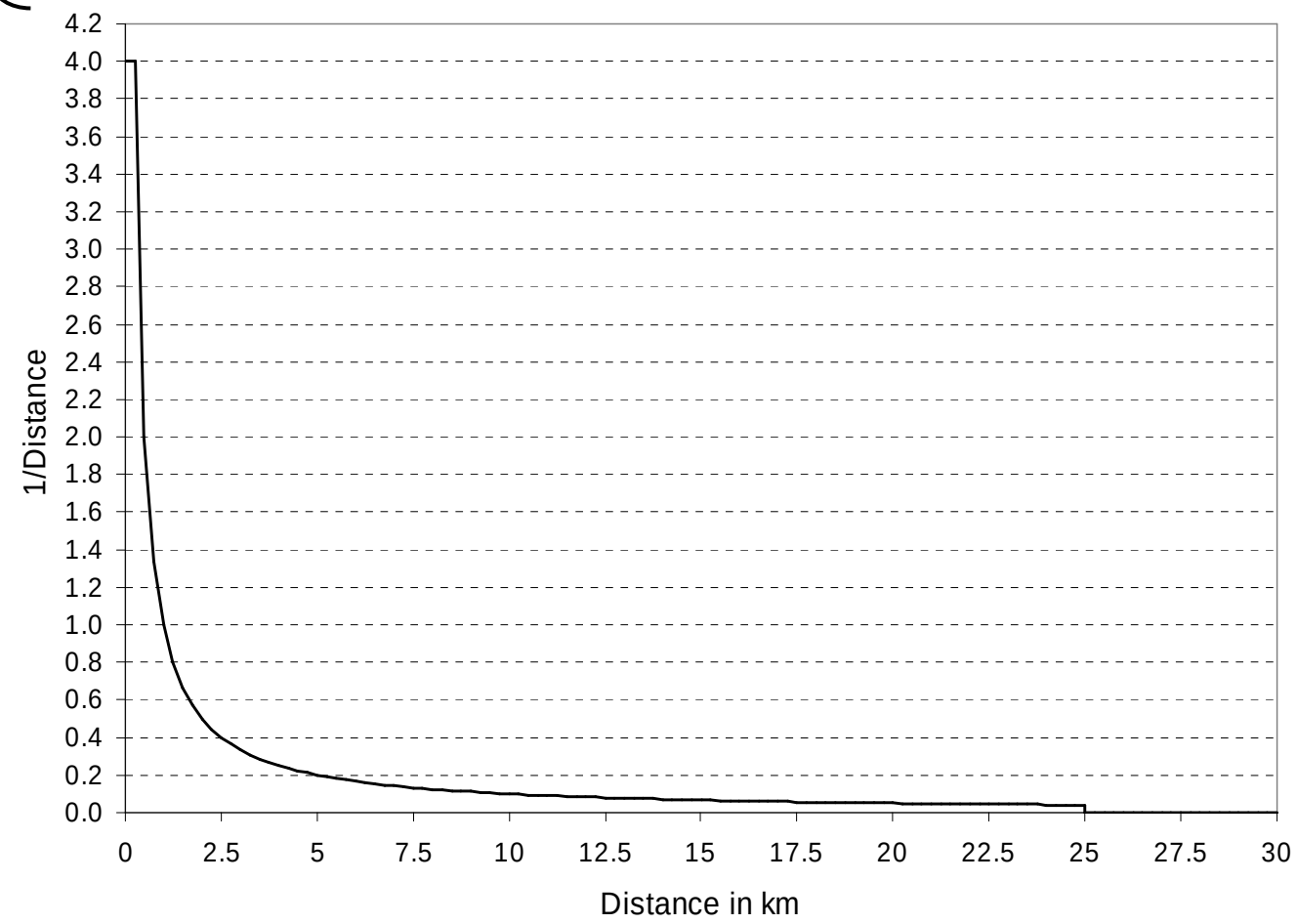
(compare autoregression: $y=\beta X+\lambda y_{t-1}+\varepsilon$)

- $W_{ij} = 0$ if i=j
- $W_{ij} = f(link_i, link_j) \sim 1/dist(i,j)$ [used here, other functions possible]
- Row norm: $\sum_{j=1}^{N} W_{ij} = 1$

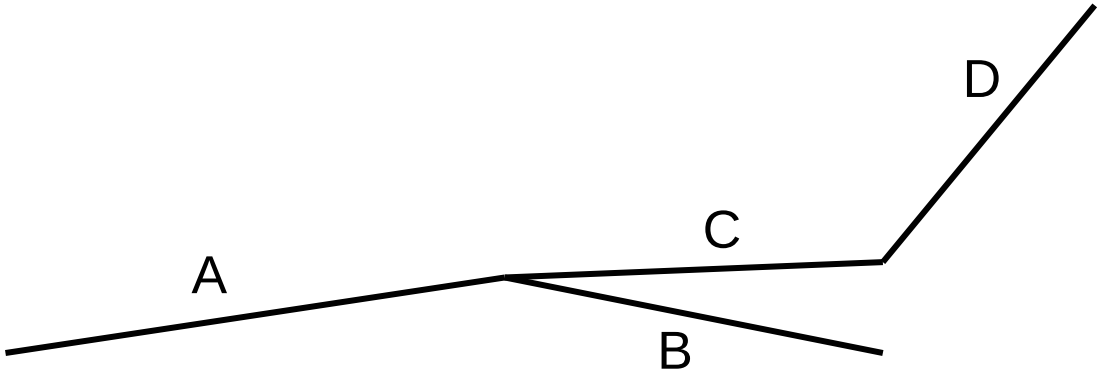
Spatial Error Model (3) Contiguity matrix W

d_{ij} =

0	if dist(i,j) = 0 (same link)
4	if 0 <= dist(i,j) < 0.250 km
1/dist(i,j)	if 0.250 <= dist(i,j) <= 25 km
0	if dist(i,j) > 25 km



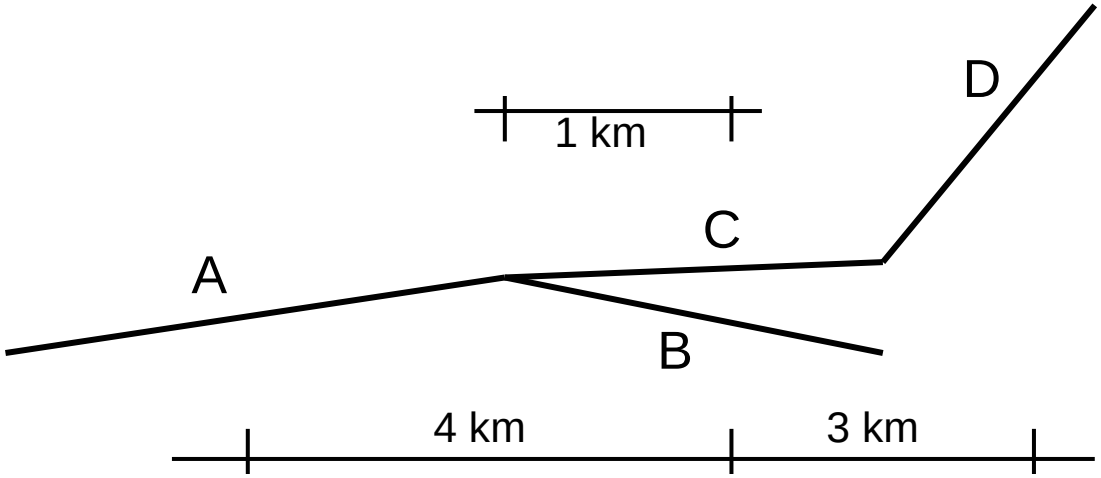
Assembly of W-matrix (1)



simple contiguity:

C	A	B	C	D
A	0	1	1	0
B	1	0	1	0
C	1	1	0	1
D	0	0	1	0

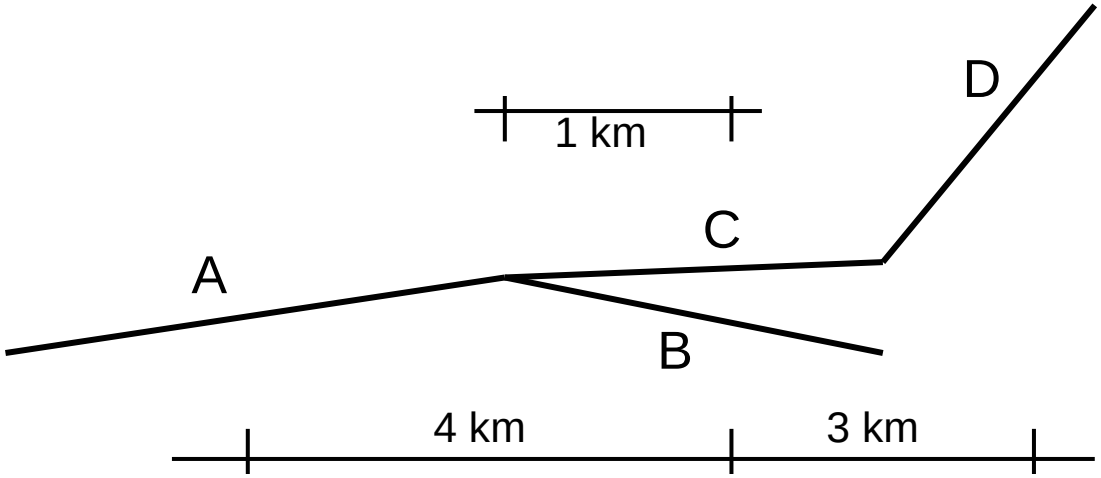
Assembly of W-matrix (2)



distances:
(directions must be separated)

d	A	B	C	D
A	0	4	4	7
B	4	0	2	5
C	4	2	0	3
D	7	5	3	0

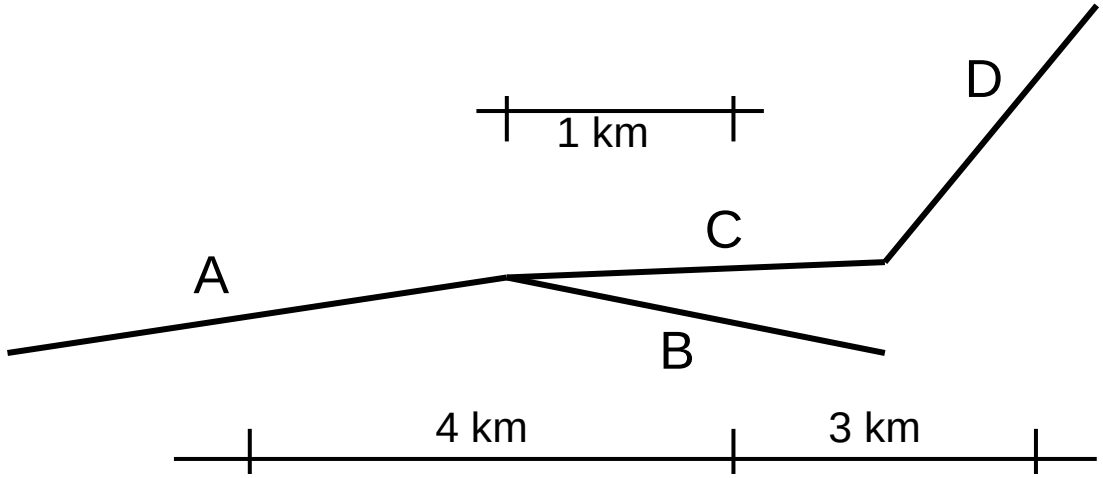
Assembly of W-matrix (3)



weights:

1/d	A	B	C	D
A	0	0.25	0.25	0.14
B	0.25	0	0.5	0.2
C	0.25	0.5	0	0.33
D	0.14	0.2	0.33	0

Assembly of W-matrix (4)



normed weights:
(row sum = 1)

W	A	B	C	D
A	0	0.39	0.39	0.22
B	0.26	0	0.53	0.21
C	0.23	0.46	0	0.31
D	0.21	0.30	0.49	0

Speed Estimation (1) Data

For each link the average measured speed was calculated

separated by direction

separated by three time windows (hvz, nvz, rvz)

→ 7827 aggregated measurements (only weekdays)

61.2 Mio. elements in W-matrix

Data sample:

link_ID	v_mean	hvz	nvz	rvz	road type	…
1001	91.33	1	0	0	10	
1001	93.50	0	1	0	10	
1001	94.00	0	0	1	10	
2002	61.25	1	0	0	20	
2002	64.00	0	1	0	20	
2002	64.75	0	0	1	20	

…

Sample Spatial Error Model of Speed on Structure Variables

	SEM R^2=0.5504		OLS R^2=0.5282	
Variable	Coefficient	t-statisic	Coefficient	t-statistic
const_1	89.70	1.03	60.53	77.54
hvz_and_HLS	42.47	21.50	44.20	23.99
hvz_and_SS	2.69	1.37	2.64	1.43
hvz_and_ES	-11.03	-4.86	-12.38	-5.49
hvz_and_otherRoad	-15.76	-6.71	-16.89	-7.31
rvz_and_HLS	38.22	0.51	56.11	27.56
rvz_and_HVS	-9.53	-0.13	6.06	7.10
rvz_and_SS	-11.17	-0.15	3.83	1.79
rvz_and_ES	-22.06	-0.29	-6.95	-2.72
rvz_and_otherRoad	-26.46	-0.35	-12.98	-3.95
nvz_and_HLS	52.52	0.92	53.10	29.82
nvz_and_HVS	3.39	0.06	2.41	3.24
nvz_and_SS	6.31	0.11	5.55	3.19
nvz_and_ES	-8.33	-0.15	-10.84	-4.98
nvz_and_otherRoad	-12.42	-0.22	-14.57	-6.76
eo3_and_HLS	-0.12	-3.19	-0.10	-2.86
eo3_and_HVS	-0.04	-5.61	-0.04	-8.16
eo3_and_SS	-0.06	-1.06	-0.07	-1.29
eo3_and_ES	-0.04	-3.58	-0.04	-3.60
eo3_and_otherRoad	-0.04	-1.45	-0.05	-1.67
p3_and_HLS	0.00	0.81	0.00	0.60
p3_and_HVS	-0.01	-6.04	-0.01	-9.64
p3_and_SS	-0.01	-1.72	-0.01	-1.63
p3_and_ES	0.00	-2.56	0.00	-2.80
p3_and_otherRoad	0.00	-0.62	0.00	-0.09
lambda	0.99	26.97		

End of slide show, <Esc> to exit.

Literature

Marchal, F., J. Hackney and K.W. Axhausen (Forthcoming). Efficient map-matching of large GPS data sets - Tests on a speed monitoring experiment in Zurich. *Transportation Research Record*.

Floating Car Speed Measurement in Canton Zurich

Goal:

Characterization of system (space mean) speed in the Canton:
Average and Variance, other statistical models

Method:

Drive a representative sample of roadway in the Canton with GPS-equipped test vehicles.

Floating Car procedure

3 independent vehicles and routes

Time Sample:

3 weeks of workday + Saturday measurements in November 2003, 0600-2100

Spatial Sample: Measurement Routes

3 circuits (550-700km length) of 50 stages (11-14km) each

Stage ODs are selected from traffic planning zones in Canton Zurich + bordering zones of surrounding Cantons

The routes are the shortest-paths between 50 OD pairs

The OD pairs were drawn representatively from the Canton demand matrix (Monte Carlo draw) and linked together end to end to form circuits

1 measurement vehicle per circuit

Structure Dataset: Hectare Data and Road Densities

Values calculated per hectare:

Density of Road Type (NavTech): *meters of road of specific type / km^2*

- *Density of Motorways*
- *Density of Trunk Roads*
- *Density of Collectors (Rural and Urban)*
- *Density of Distributors (Urban)*
- *Density of Other Rural Roads*
- *Density of Ramps*
- *Density of Other Roads*

Values averaged over larger area to simulate spatial influence:

Density of Structure Variables (Hectare Dataset): N / km^2

- *Employment opportunities*
- *Employed*
- *Population*
- *Motorway Access Points*

Temporal Categories

Time of Day:

HVZ: Morning- (0630-0830) and Evening Peaks (1630-1830)
NVZ: Day (0830-1630) and (1830-2030)
RVZ: Night and early morning (2030-0630)

Type of Day:

Workday/Weekend (Saturday) all hours

Road type:

Code	Name	Nr.
HLS	Hochleistungsstrassen (High Speed Network)	10
HVS/HVSII	Hauptstrassen/Verbindungsstrassen (Main Trunk Roads)	20
SS	Sammelstrassen (Collectors)	30
ES	Erschliessungsstrassen (Distributor Roads)	40

EXTRA System Speed

Road type	Data	Saturday	HVZ (Peak)	NVZ (Off-peak)	RVZ (Other)	Mo.-Fr.	Total
HLS (motorway and similar)	Average	101.2	85.3	101.5	93.0	96.1	97.1
	Standard Deviation	8.3	5.5	8.3	7.0	6.8	8.1
	Number of Observations	2495	2235	6306	991	9532	12027
	Km driven	1475.9	1359.4	3703.2	628.4	5691.0	7166.9
HVS/HVSII (trunk roads)	Average	40.5	35.0	39.0	41.6	38.1	38.6
	Standard Deviation	2.1	1.5	1.9	2.8	1.8	2.1
	Number of Observations	5696	5246	14732	2648	22626	28322
	Km driven	2159.5	2244.9	5808.1	846.8	8899.7	11059.2
SS (collector)	Average	50.9	41.7	46.1	48.3	45.0	46.1
	Standard Deviation	3.0	2.2	2.1	4.7	1.9	1.8
	Number of Observations	1118	1195	3052	258	4505	5623
	Km driven	753.0	728.1	2057.0	185.5	2970.6	3723.5

EXTRA Stichprobenziehung

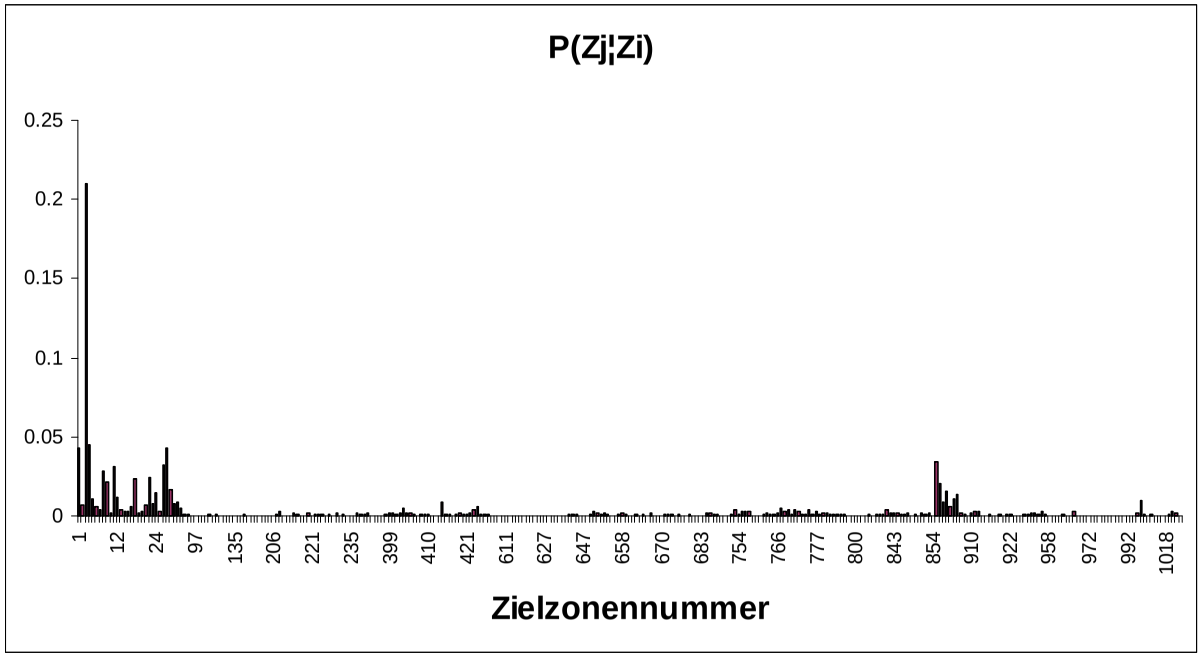
Repräsentative Verkehrsplanungszonen

Ziehung der Ziele gewichtet nach der DTV Matrix des Kantons

Monte Carlo Markov Chain:

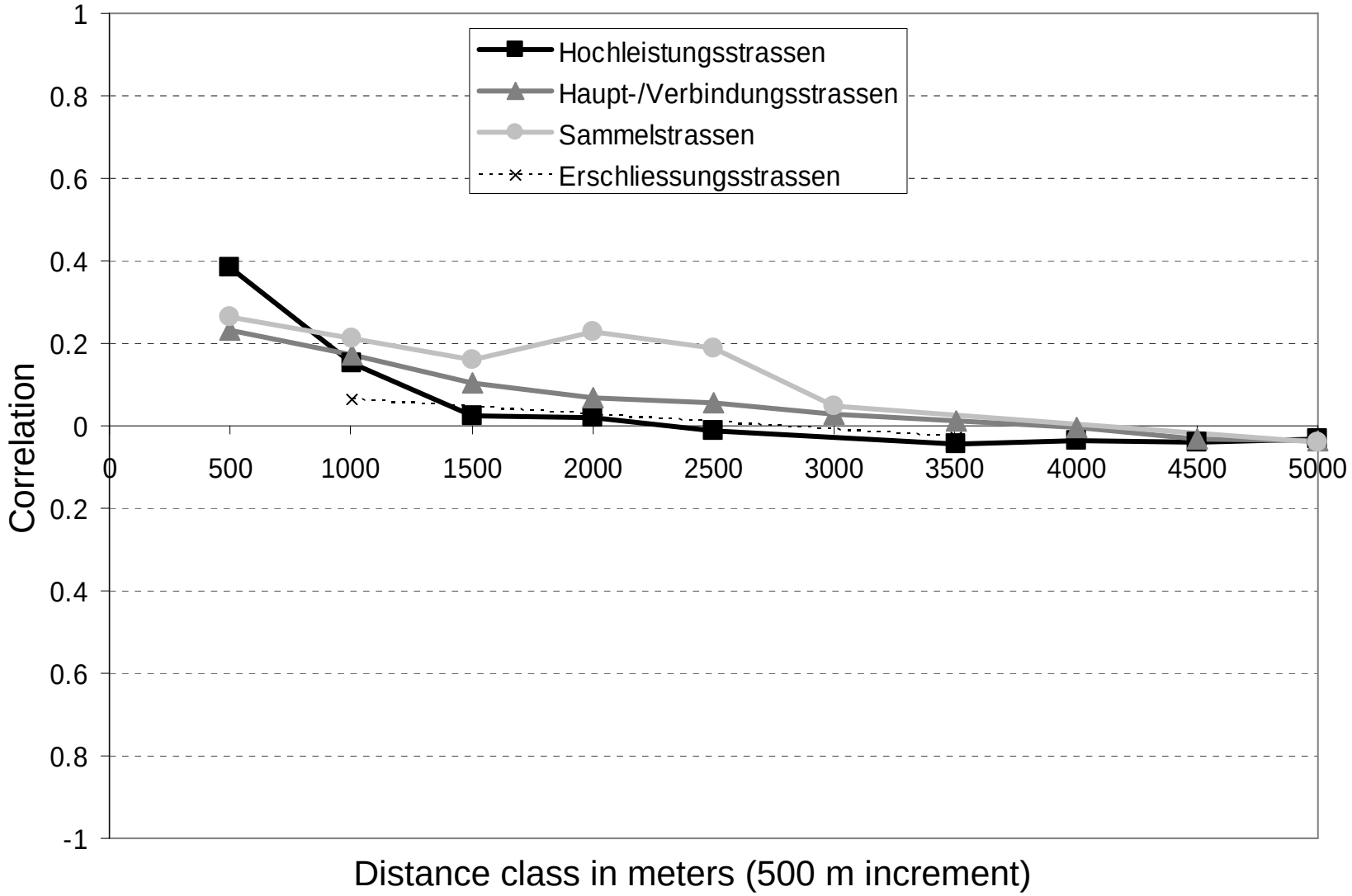
$$P(Z_i \mid Z_j) = \frac{N_{i,j}}{\sum_{l=1}^{J} N_{i,l}}; Z_i = Z_j$$

EXTRA Beispiel der Zonenziehung

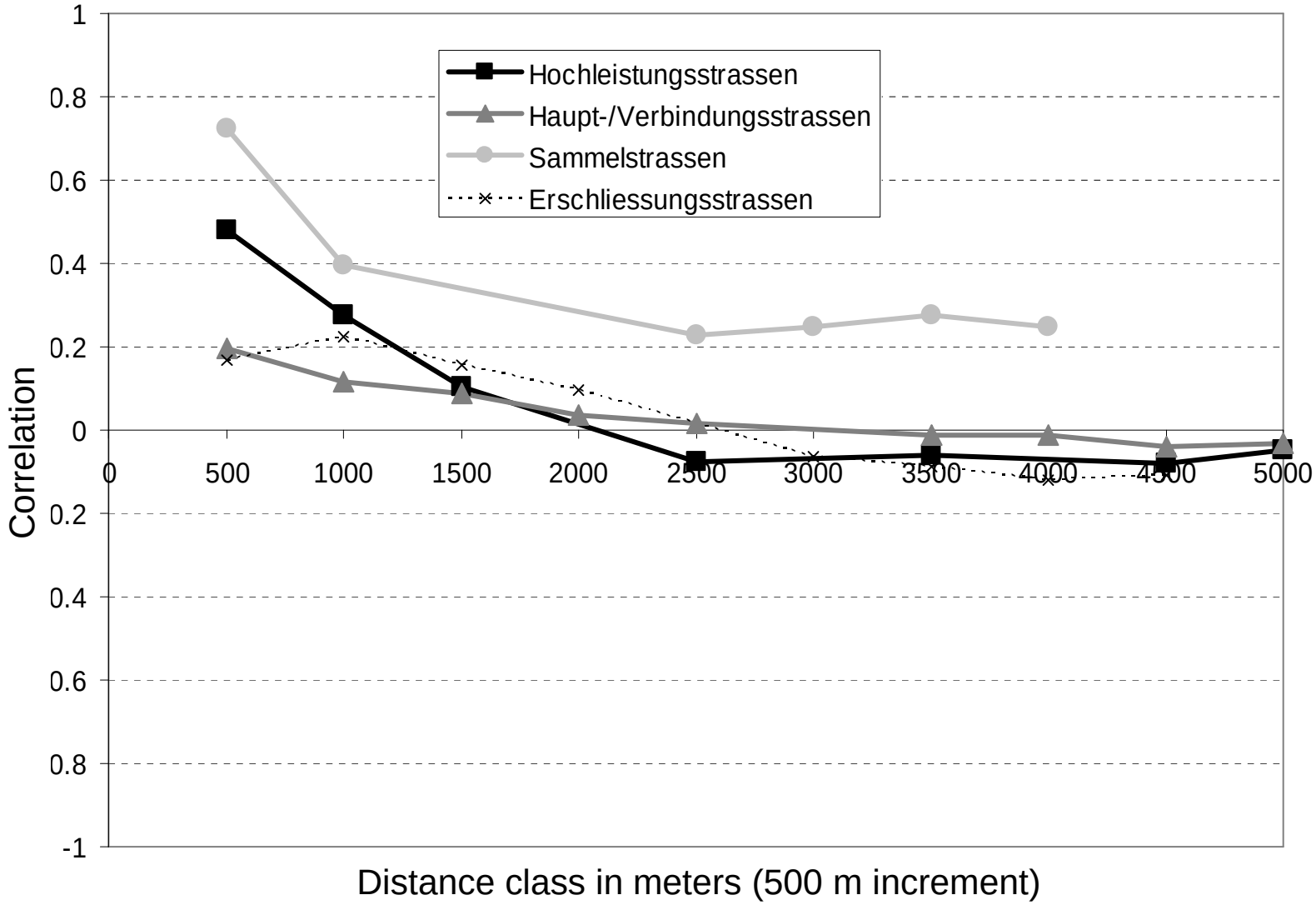


EXTRA Punkte und Netzmodell

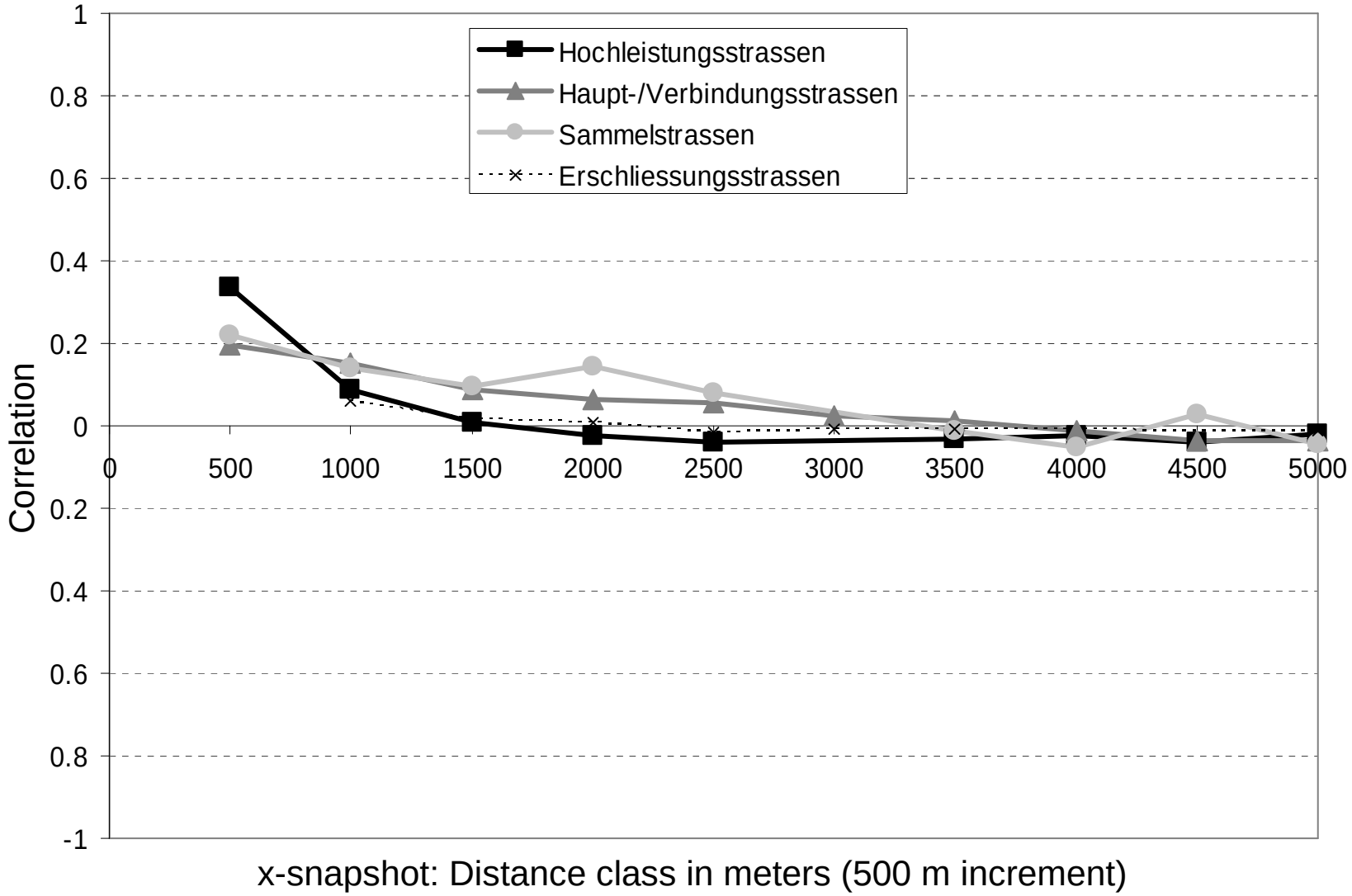
Correlation (6) weekday, NVZ: 8:30-16:30, 18:30-20:30



Correlation (7) weekday, RVZ: 20:30-6:30



Correlation (10) weekday, local snapshot, NVZ



Correlation (11) weekday, local snapshot, RVZ