Floating car data for transportation planning

Explorative study to technique and applications and sample properties of GPS data

Nick Cohn and Hugo Bischoff, TomTom

NATMEC 2012, Dallas

Technische Universiteit Eindhoven University of Technology

TomTom

Contents

- Explorative study of FCD collection, quality and applications
- Example applications
- Case studies analyzing usability
- Next steps

What is floating car data?

- Gathering of traffic information based on ‘floating’ sensors
- New collection method where the car itself functions as sensor
- Primary sources for floating car data are GPS devices, mobile phones and bluetooth

Processing steps

- Retrieve raw positions and corresponding timestamps

- Next process is that of map-matching:
  - During this process raw positions are matched to the road network
  - Smart algorithm is necessary
  - Account for:
    - Signal loss
    - Inaccuracy in urban areas
    - Uncertainty about standstills: has the vehicle stopped or is it in a traffic jam?

Quality aspects

Floating car data is always **a sample**

Main questions:

- Is the behaviour of users representative of the total traffic?
- Is the sample size large enough to deliver useful information?
- Other influences:
  - Turning device on/off
  - Routing algorithm
  - Accuracy of the map

Advantages / disadvantages

Benefits of floating car data

- Very detailed information source
- Capable of covering a vast amount of applications
- Large time period of data
- Entire network is covered
- Only solution that can deliver very detailed origin, destination and route choice information

Disadvantages of floating car data

- No direct information on traffic flow/density
- Dependent on sample properties

Potential applications

Network performance analysis

- Monitoring of the complete network
- Before/after analysis
- Bottleneck analysis
- Route choice in case of events/detours

Forecasting / estimation

- Traffic growth
- OD relations & matrices
- Route choice analysis for traffic simulation
- Emission modelling (air & noise pollution)

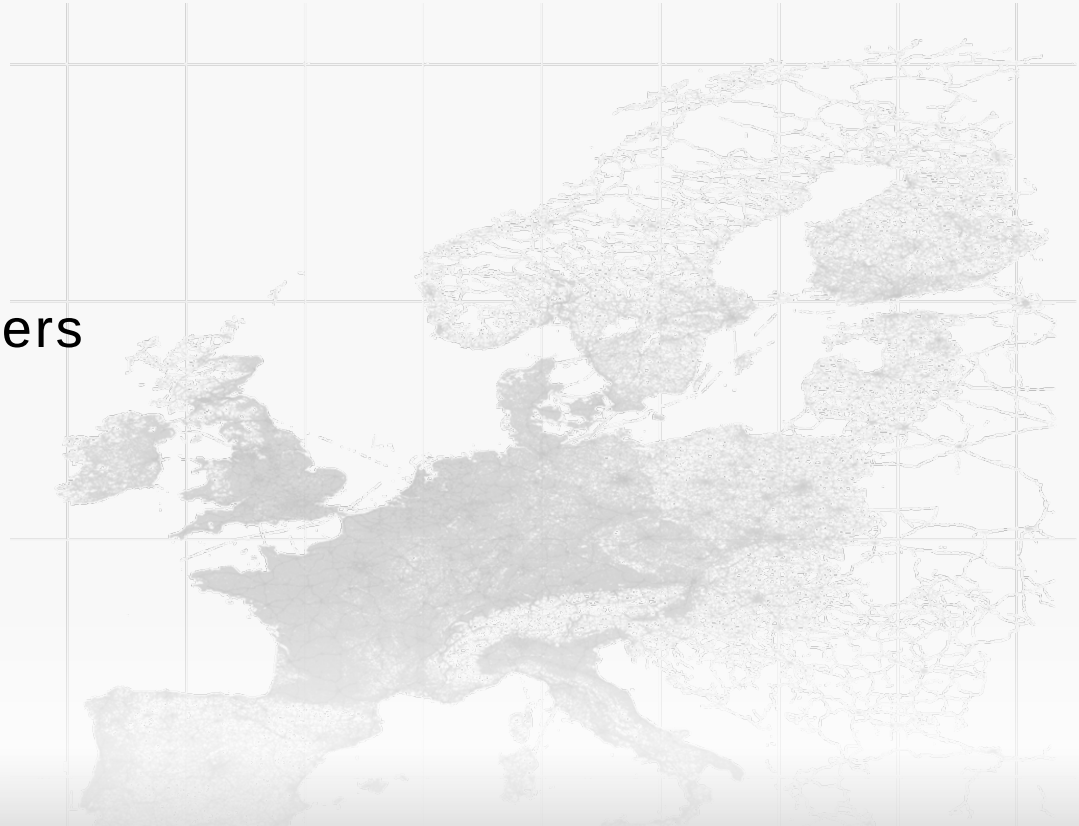
Potential applications

Road maintance / safety analysis

- Traffic volume estimation
- Speed limit exceedance

Location planning

- Regional coverage
- Location choice
- Potential visitor numbers
- Customer origins



Examples of applications

Bottleneck Analysis

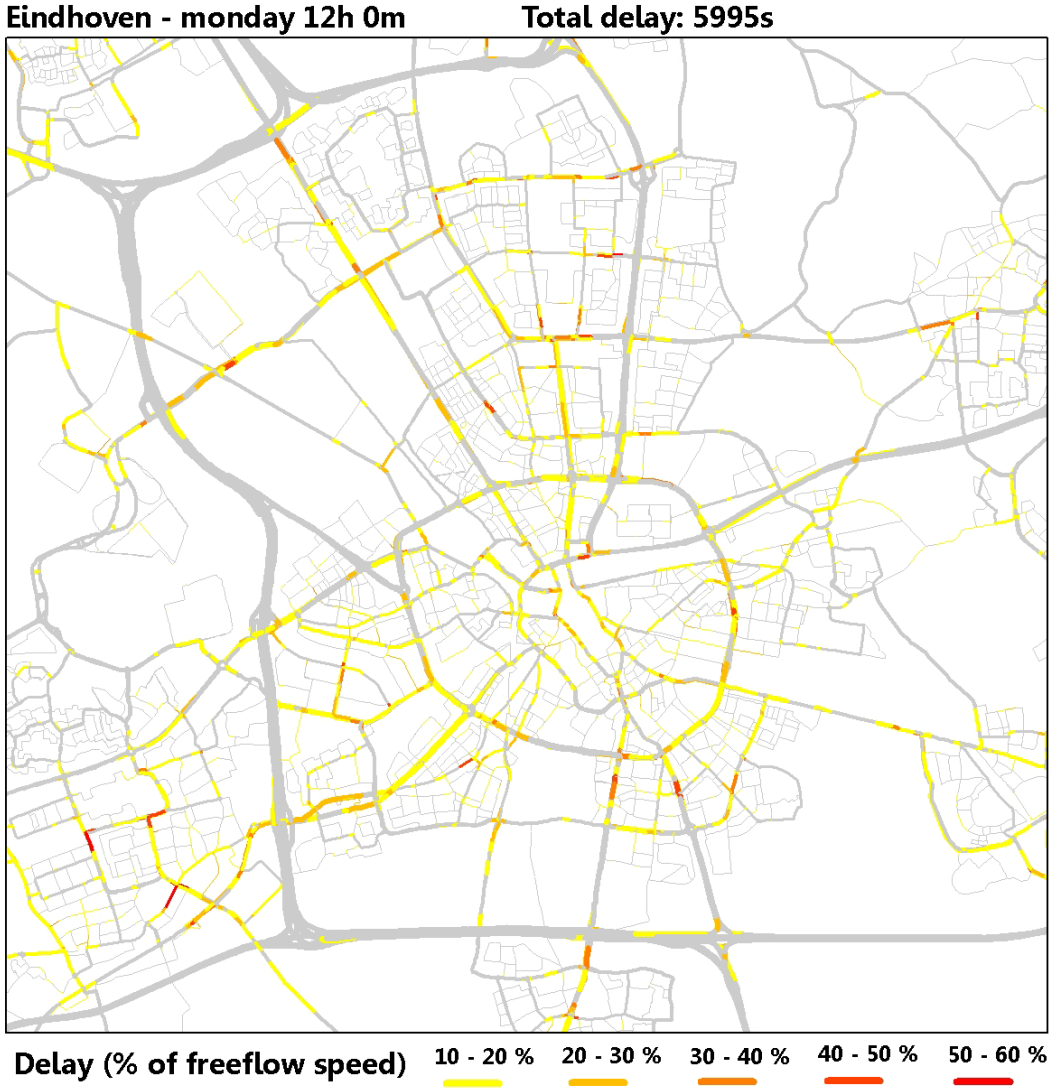
Average delay at 08:00 am

...

at 08:10 am

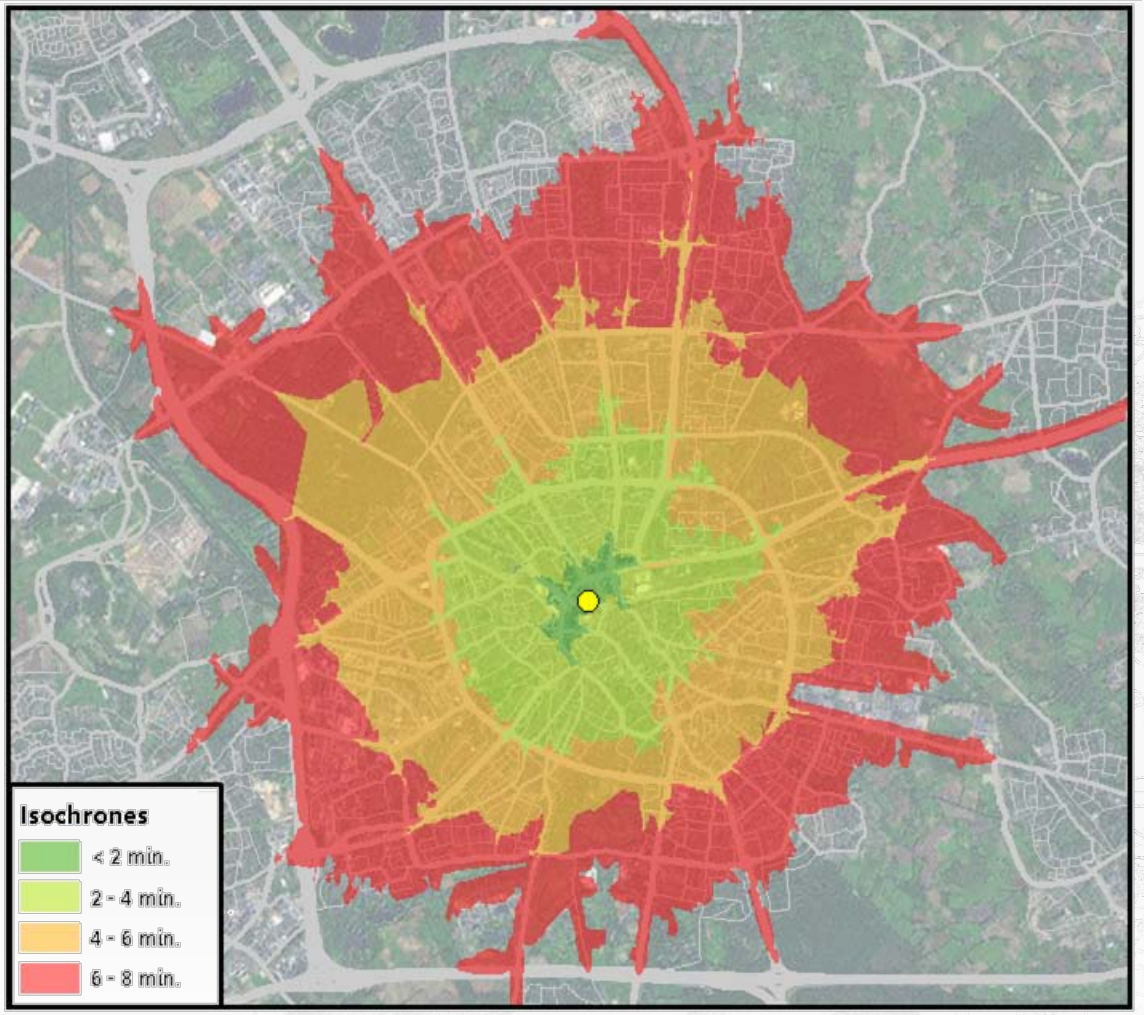
...

at 12:00 am



Examples of applications

Accessibility



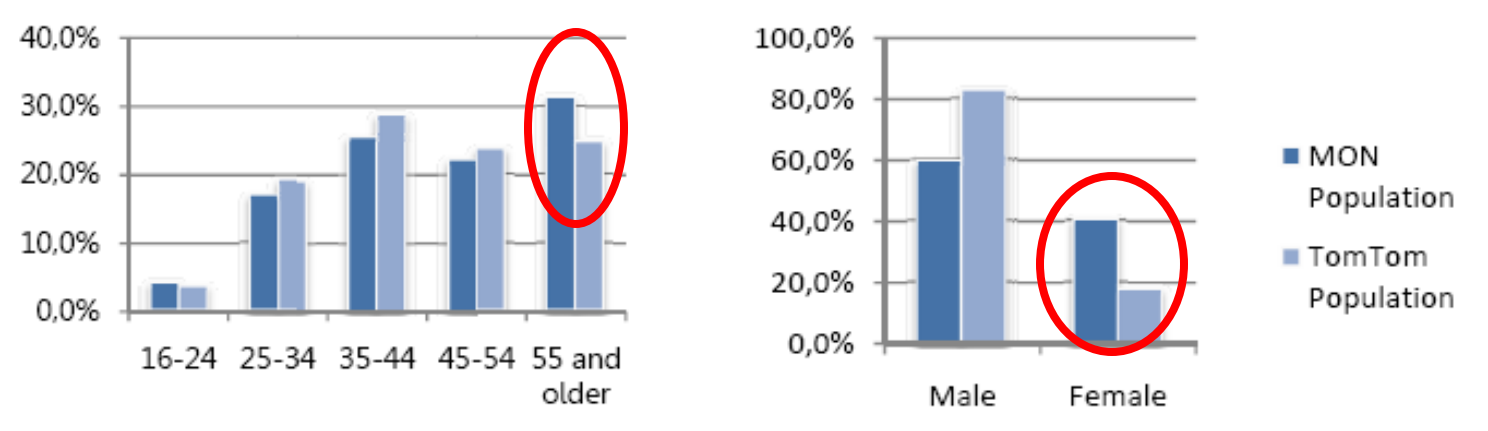
Examples of applications

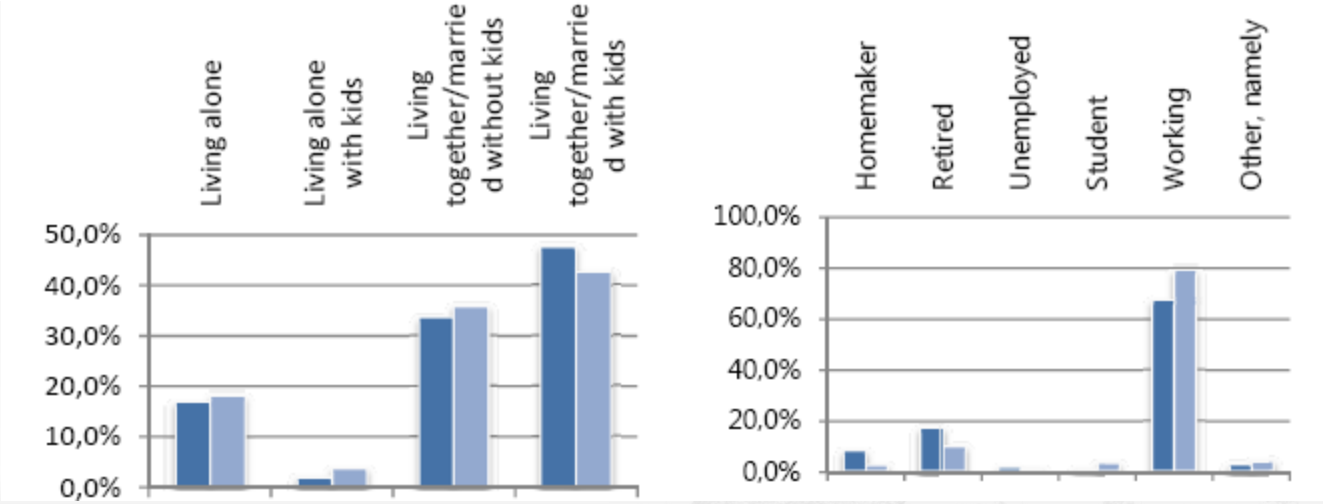
Origin/destination/route choice

	Destination									
Origin	1	2	3	4	5	6	7	8	9	Total
1	11	49	39	17						
2	63	29	16	18						
3	81	12	35	16						
4	49	101	5	1						
5	21	15	48	3						
6	102	25	47							
7	84	21	81	17						
8	125	66	65	16						
9	45	101	111	11						
Total	581	419	447	103						

Population of TomTom users

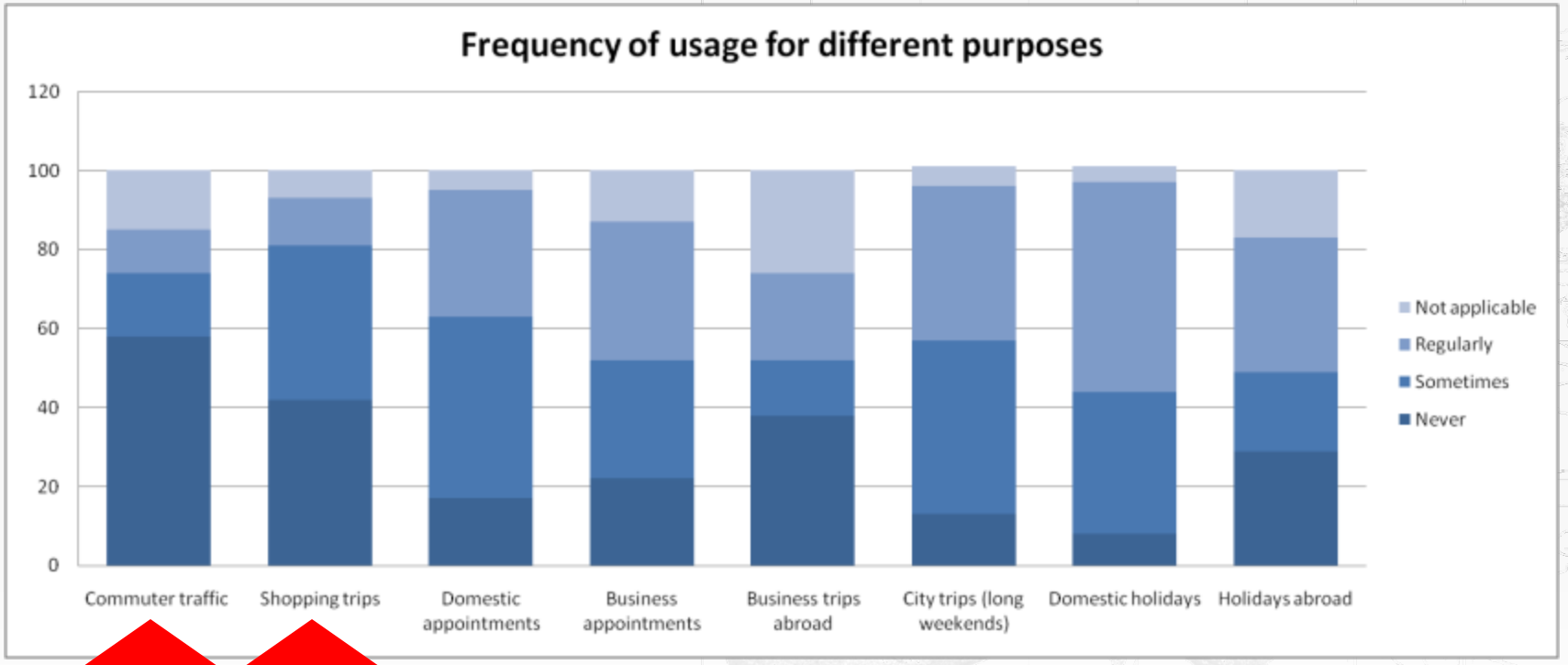
- TomTom users compared to MON 2006





TomTom users

- Quality not only dependent on user population
- Large influence by moments when TomTom device is used
- Possible bias indication:



Case study Eindhoven

Investigation of origin, destination and route choice information based on TomTom floating car data

Technische Universiteit Eindhoven University of Technology

TOMTOM

Towards results

- Important steps in coming to output:

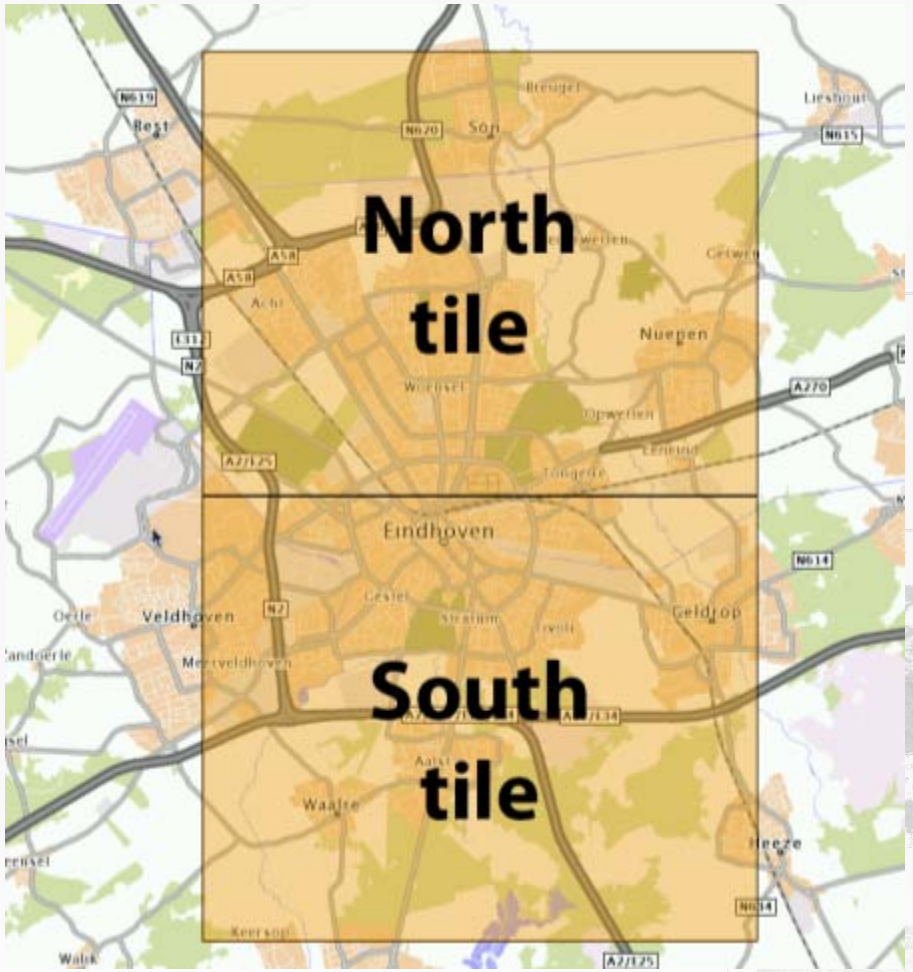
  - Converting raw floating car data for use in a GIS system with database

  - Detection algorithm to deduce trips

  - Discovering sample properties

  - Possibility of a scaling/correction algorithm for sample bias

Study area / timespan

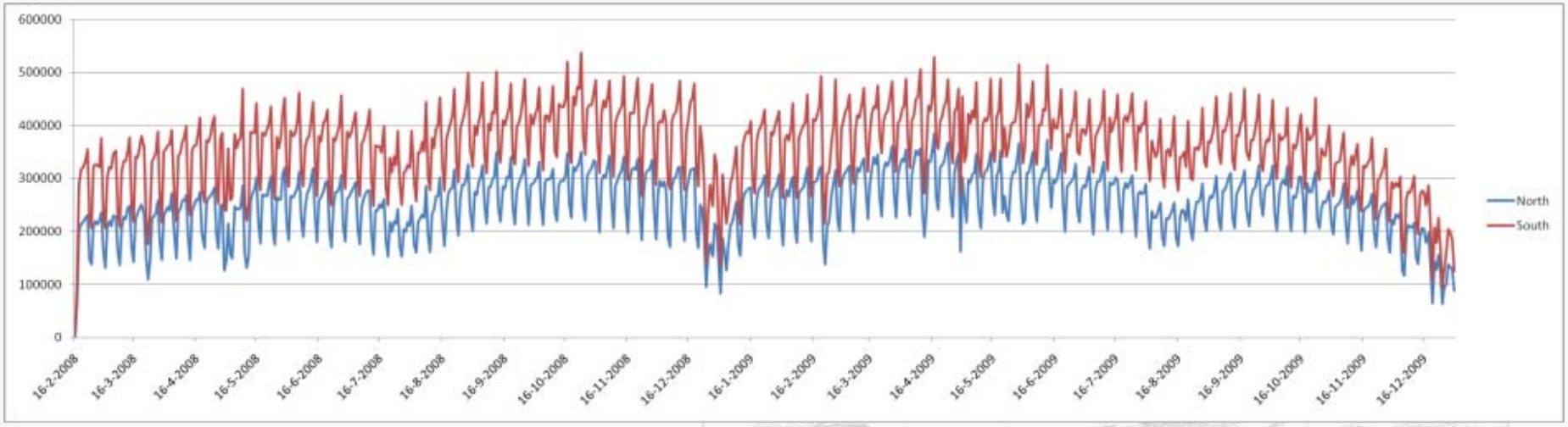


Timespan

16-2-2008 to 31-12-2009

Data points

~ 150 million measurements (~25 GB flat files)



<u>Visible patterns</u>
- Week behaviour
- Dips during holidays

Raw map-matched data

- A text file per tile per day (in total 1374 text files), containing:

Edge ID	Device ID (for this day)	Entrytime	Transittime
Number - 15 char.	*Number - 9 char.*	*Time in millisec. from epoch - 13 char.*	*Time in millisec. (no. of char. varies)*
464350001664692	68447952	1203007318000	10812
464350049531113	68447952	1203007402000	4642
464350002778471	68447952	1203007332000	1272
464350049539009	68447952	1203007385000	13445
464350001664689	68447952	1203007341000	7259

- Edge ID compares to the map
- DeviceID is a unique ID per device per day (cut-off at 00:00 UTC)
- Entrytime for that dseg and transittime are in milliseconds, where entry is from epoch (1-1-1970 00:00 UTC)

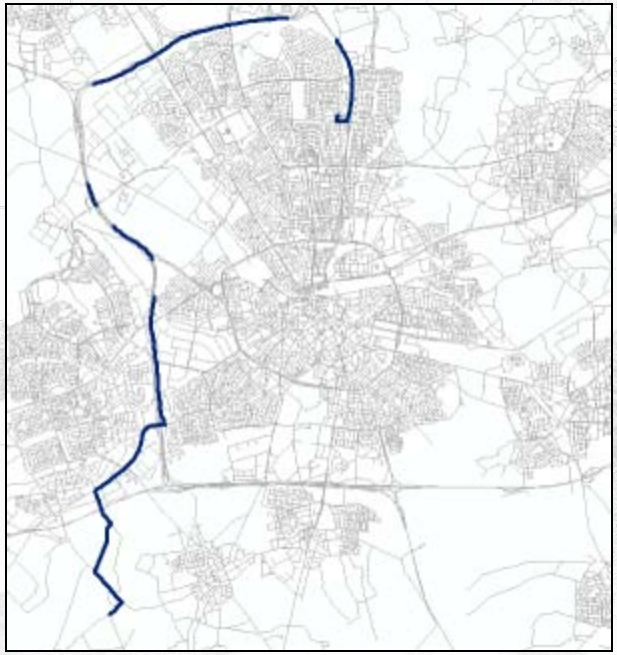
TomTom Map (TIF-R format)

- Used for matching floating car data
- Shapefile format used in ArcGIS
- Contains relevant edge information:
  - ID
  - FRC
  - FOW
  - Length

Converting the raw map-matched data

- Raw data consists of individual measurements per edge
- By using the Device ID however a vehicle can be ‘traced’ for an entire day
- Using the ingress time next to the Device ID an ordered list of travelled segments can be derived per device per day

DeviceID	EdgeID	EntryTime	TransitTime
171614373	944601	1222839627000	11266
171614373	944869	1222839639000	20561
....	....	....	....
171614373	221212	1222841334000	2183
171614373	221528	1222841336000	3530



Trip recognition

- However, more information can be derived by analyzing the time between two consecutive measurements
- When there is a considerable time gap between measurements the best assumption is that a trip had the last visited segment before the gap as its destination

DeviceID	EdgeID	EntryTime
176577013	2186594	1222870163000
		1 sec.
176577013	2290845	1222870164000
		Trip end!
		More than 2 hours!
176577013	2290845	1222879135000
		16 sec.
176577013	2186596	1222879151000

Potential issues

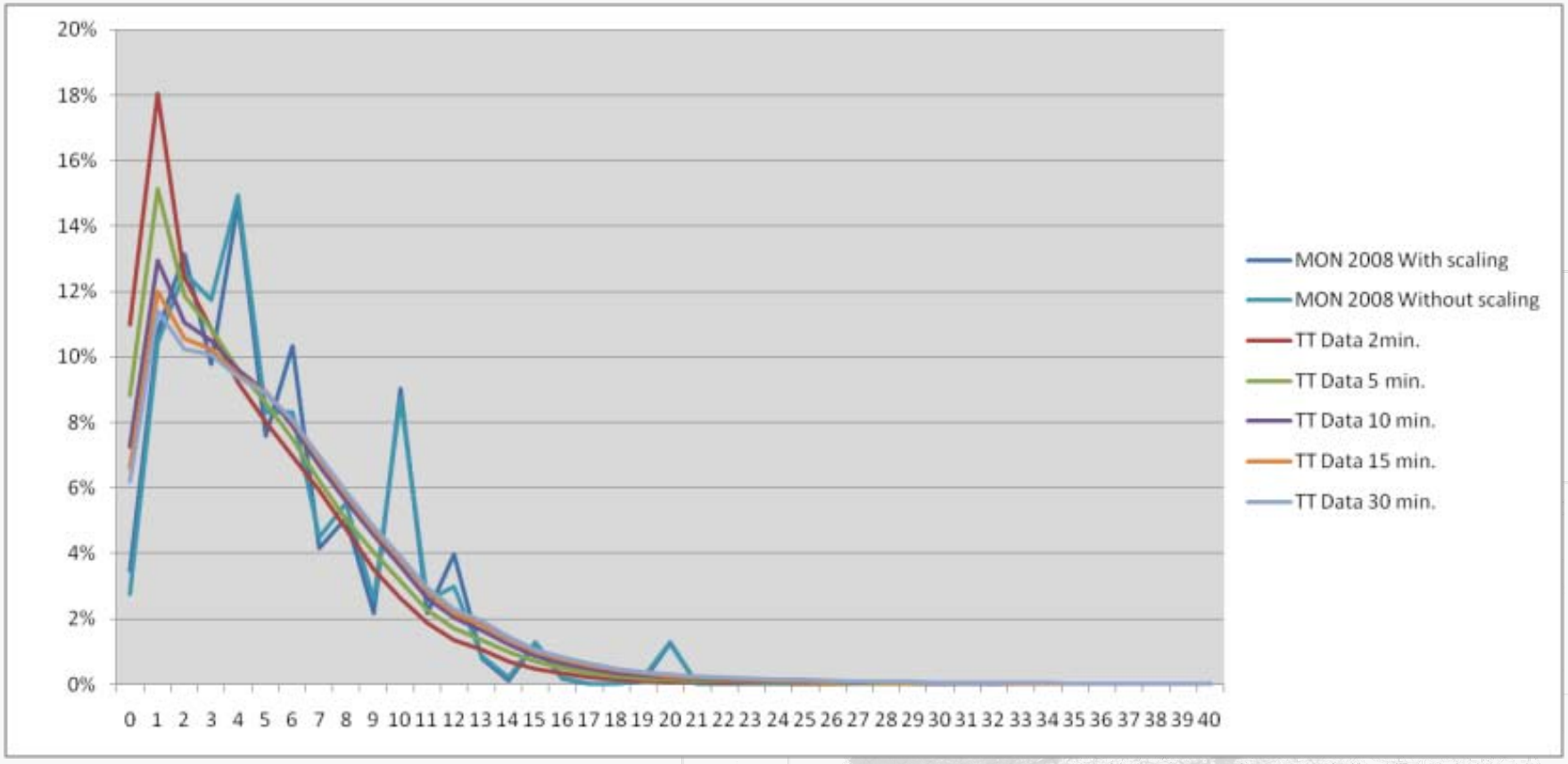
- Caution has to be taken regarding these trip starts/ends for several reasons:
  - Devices do not have to be on all the time; they can be switched on/off during the trip
  - GPS signal can be lost somewhere on the way before reaching the true destination
  - Map matching can go wrong because of a too old, too new or just wrong map

OD Matrix comparison

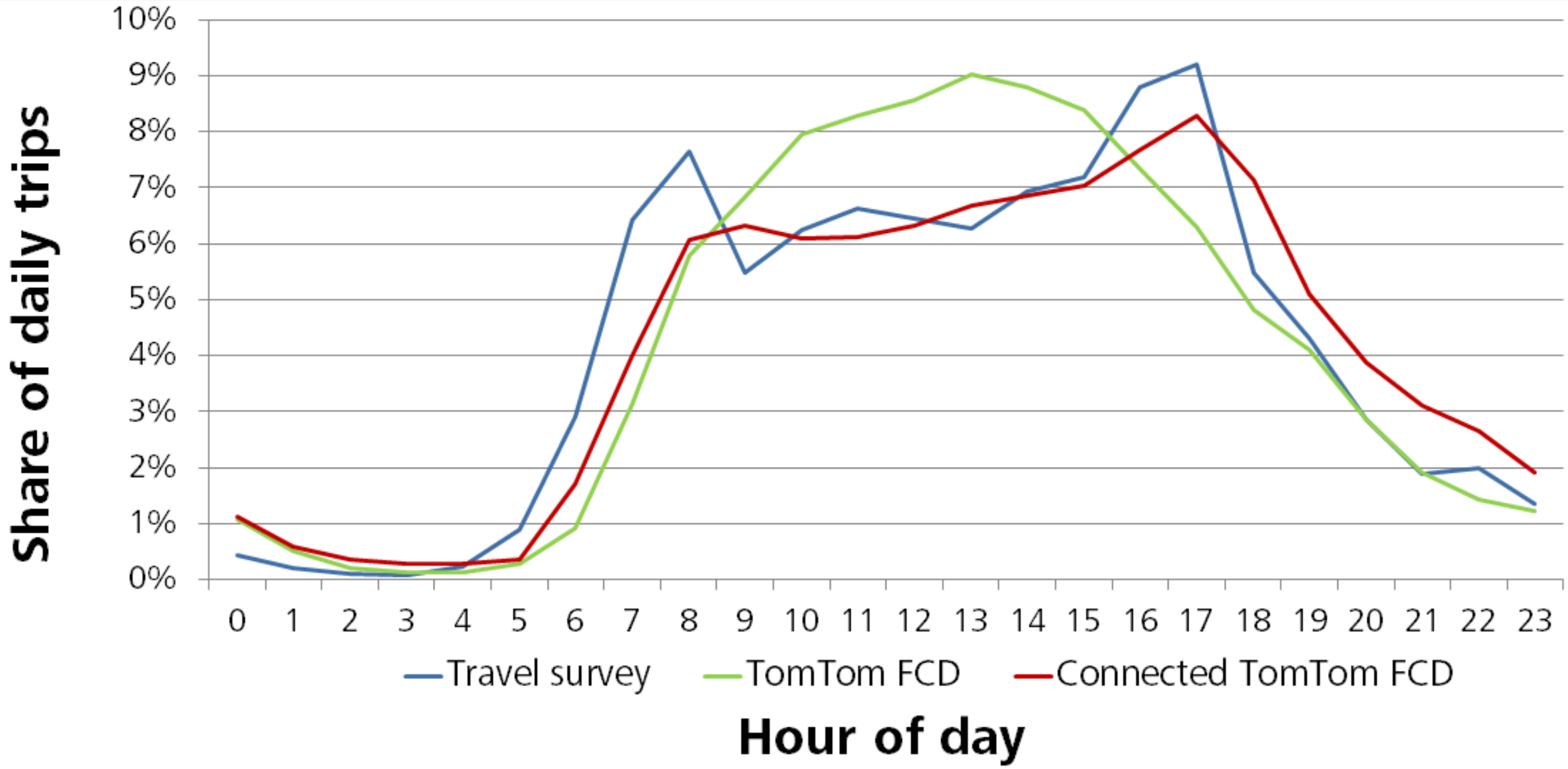
OD Matrix MON	550	558	561	562	563	564	565	566	567	568	569	999	Origins
550	0,08	0,00	0,51	0,24	0,00	0,07	0,30	0,08	0,10	0,19	0,09	3,45	5,13
558	0,00	1,69	0,00	0,07	0,00	0,07	1,09	0,15	0,00	0,00	0,00	1,22	4,29
561	0,67	0,00	3,30	1,82	0,39	0,44	0,43	0,46	0,67	0,00	0,23	4,52	12,93
562	0,24	0,00	2,10	3,65	0,73	0,58	0,70	0,32	0,53	0,04	1,15	5,06	15,10
563	0,00	0,00	0,47	0,79	0,25	0,33	0,12	0,06	0,09	0,00	0,00	1,38	3,50
564	0,12	0,07	0,56	0,76	0,22	0,56	0,00	0,58	0,21	0,09	0,00	1,54	4,71
565	0,19	1,16	0,43	0,50	0,07	0,00	0,60	0,32	0,33	0,00	0,20	3,05	6,85
566	0,08	0,06	0,15	0,33	0,06	0,63	0,32	2,81	0,26	0,00	0,00	2,00	6,70
567	0,10	0,00	0,37	0,53	0,09	0,30	0,33	0,22	1,71	0,10	0,11	1,15	5,02
568	0,00	0,00	0,00	0,00	0,00	0,13	0,00	0,00	0,06	0,65	0,07	2,28	3,19
569	0,09	0,00	0,23	1,20	0,00	0,00	0,23	0,00	0,11	0,00	2,06	1,79	5,71
999	3,21	1,31	5,01	4,81	1,41	1,57	2,65	1,70	1,39	1,94	1,88	0,00	26,89
Destinations	4,79	4,29	13,13	14,71	3,23	4,68	6,77	6,70	5,46	3,02	5,79	27,43	100

OD Matrix TomTom	550	558	561	562	563	564	565	566	567	568	569	999	Origins
550	0,66	0,11	0,22	0,17	0,05	0,11	0,42	0,09	0,07	0,04	0,05	1,09	3,06
558	0,10	0,62	0,21	0,15	0,05	0,18	0,32	0,10	0,06	0,03	0,04	0,97	2,83
561	0,19	0,22	4,07	1,53	0,61	1,11	2,00	0,45	0,42	0,47	0,33	2,17	13,59
562	0,18	0,17	1,55	3,72	0,76	0,59	2,28	0,32	0,63	2,25	0,46	1,84	14,75
563	0,06	0,05	0,49	0,63	0,64	0,25	0,31	0,12	0,28	0,12	0,22	0,52	3,69
564	0,09	0,16	0,85	0,55	0,31	1,24	0,65	0,32	0,25	0,13	0,13	0,83	5,51
565	0,33	0,25	1,59	1,41	0,32	0,61	4,23	0,32	0,32	0,30	0,31	2,28	12,27
566	0,08	0,08	0,48	0,33	0,17	0,38	0,44	2,20	0,45	0,11	0,09	1,19	6,00
567	0,05	0,06	0,49	0,62	0,37	0,26	0,36	0,50	2,70	0,21	0,34	0,68	6,64
568	0,03	0,03	0,29	0,80	0,12	0,09	0,22	0,07	0,23	2,35	0,62	0,50	5,35
569	0,06	0,05	0,42	0,56	0,19	0,15	0,35	0,11	0,39	0,69	1,76	0,89	5,63
999	1,30	1,37	3,33	2,62	0,59	1,53	5,92	1,53	0,80	0,81	0,89	0,00	20,70
Destinations	3,13	3,18	14,00	13,08	4,19	6,49	17,50	6,13	6,58	7,51	5,24	12,96	100

Comparison of trip length distribution



Comparison of time of day distribution



Case study Richmond, VA

Investigation of origin, destination and route choice information based on TomTom floating car data

Technische Universiteit Eindhoven University of Technology

TOMTOM

I-95/I-64 Overlap Study

- Virginia DOT completed an analysis of origin-destination movements using two sets of data:
  - Extensive license-plate survey (ALPR)
  - GPS-based origin-destination data from TomTom
- The OD data sources were used to understand travel patterns in the study area
- The GPS-based OD data compared favorably in terms of its quality and cost

I-95/I-64 Overlap Study

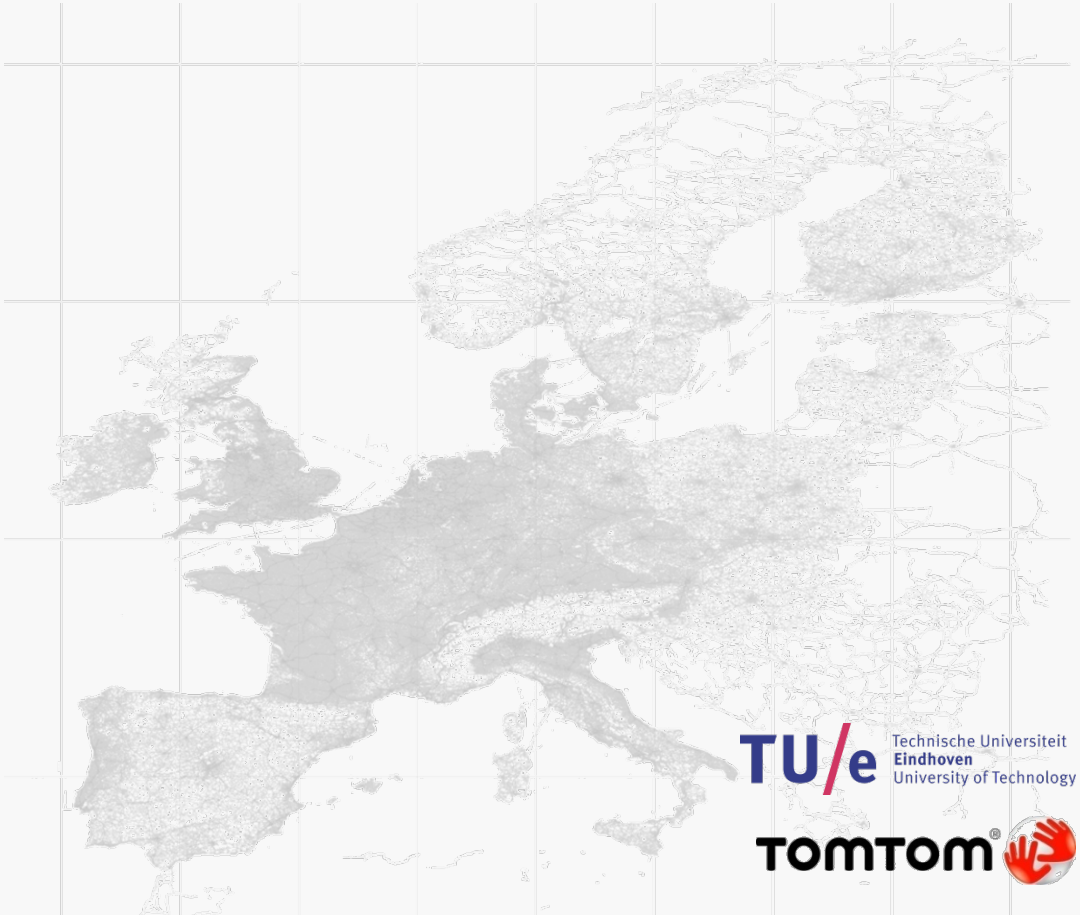
Table 4: TomTom O-D Data – Morning Peak Period (6:00 AM to 9:00 AM)

Destinations	Origins							
	A	B	C	D	E	F	G	H
A – I-64 West of Staples Mills Interchange		9.72%	23.77%	21.48%	1.08%	16.48%	2.54%	15.98%
B – I-95 North of Bryan Park Interchange	12.63%		54.27%	7.59%	2.08%	39.35%	1.41%	17.58%
C – I-195 South of Bryan Park Interchange	15.48%	30.30%		5.41%	0.43%	2.03%	0.00%	0.80%
D – I-64 East of Shockoe Bottom Bridge	25.12%	7.01%	2.62%		65.18%	11.15%	12.39%	23.78%
E – I-195 West of I-95	0.44%	1.72%	0.14%	17.37%		6.88%	1.69%	0.30%
F – I-95 South of I-195 Interchange	10.38%	18.22%	0.75%	11.73%	12.85%		21.13%	4.70%
G – Belvidere Street North of I-95/I-64	0.26%	0.19%	0.14%	0.79%	0.57%	0.99%		19.68%
H – Belvidere Street South of I-95/I-64	3.41%	4.56%	0.25%	1.09%	0.07%	0.32%	49.01%	
Number of Measurements	5,020	5,762	4,426	5,320	1,393	5,875	355	1,001

Table 5: TomTom O-D Data – Midday Peak Period (11:00 AM to 2:00 PM)

Destinations	Origins							
	A	B	C	D	E	F	G	H
A – I-64 West of Staples Mills Interchange		12.20%	25.52%	29.75%	2.78%	18.10%	3.19%	17.29%
B – I-95 North of Bryan Park Interchange	14.66%		52.59%	10.19%	3.57%	37.52%	3.09%	21.19%
C – I-195 South of Bryan Park Interchange	16.93%	29.76%		4.13%	0.56%	2.22%	0.43%	1.17%
D – I-64 East of Shockoe Bottom Bridge	28.32%	6.93%	3.54%		59.84%	14.51%	15.00%	25.59%
E – I-195 West of I-95	0.68%	0.98%	0.25%	13.62%		6.87%	2.34%	0.53%
F – I-95 South of I-195 Interchange	12.02%	22.83%	1.87%	12.06%	16.40%		13.40%	5.43%
G – Belvidere Street North of I-95/I-64	0.41%	0.28%	0.09%	1.70%	0.23%	1.52%		17.11%
H – Belvidere Street South of I-95/I-64	3.35%	4.80%	0.33%	0.33%	0.32%	0.67%	54.57%	
Number of Measurements	9,227	9,859	6,320	9,008	2,159	7,477	940	2,817

Conclusions and Next steps



Final conclusion

- Floating car data highly accurate and usable for travel times and speeds

- Also usable for OD and route choice, however:
  - Larger areas and time periods are necessary for good results
  - Usability depends on application

Recommendation for next steps

- Put more research into a scaling method
- Keep validating with more accurate and better reference sources
- When delivering to businesses and governments, be clear about potential bias in end results

Thanks for your attention

Questions?

Technische Universiteit Eindhoven University of Technology

TOMTOM