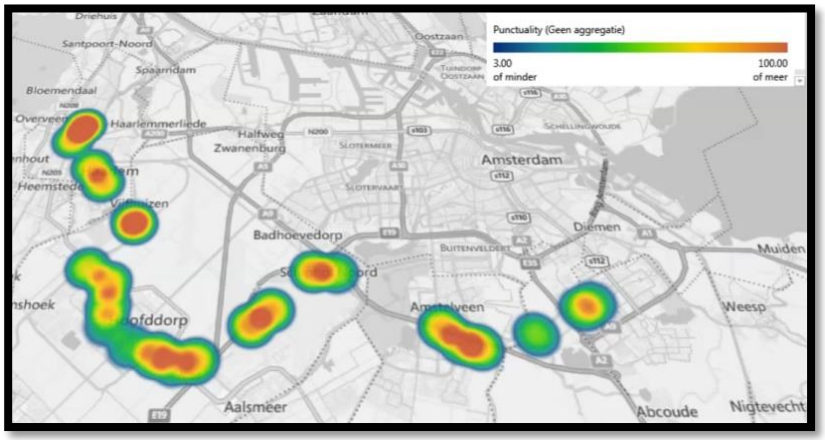
Applying multiple Big Data sources

to improve public transport planning and operations



dr. ir. N. van Oort

Assistant professor public transport

Smart Public Transport Lab

Delft University of Technology

www.SmartPTLab.TUDelft.nl

UITP INDIA
SEMINAR ON URBAN RAIL NETWORK – BUILDING SUSTAINABLE CITIES

December 2018

@Niels_van_Oort



Intermodal Hubs

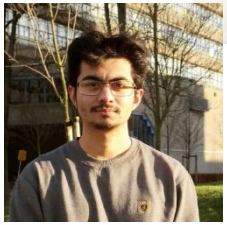

Demand Responsive Transit


Passenger flows


Choice modelling

Network Vulnerability

Network assignment



Demand Forecasting

Prediction & Information

North-South metro line

CO_2 modelling

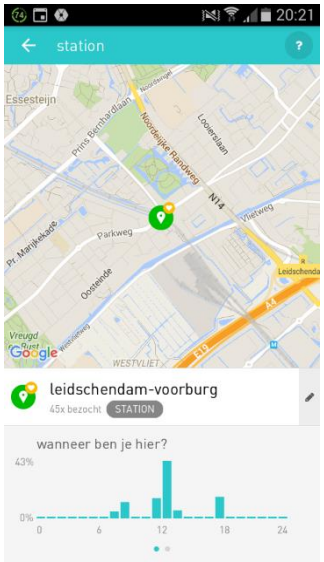






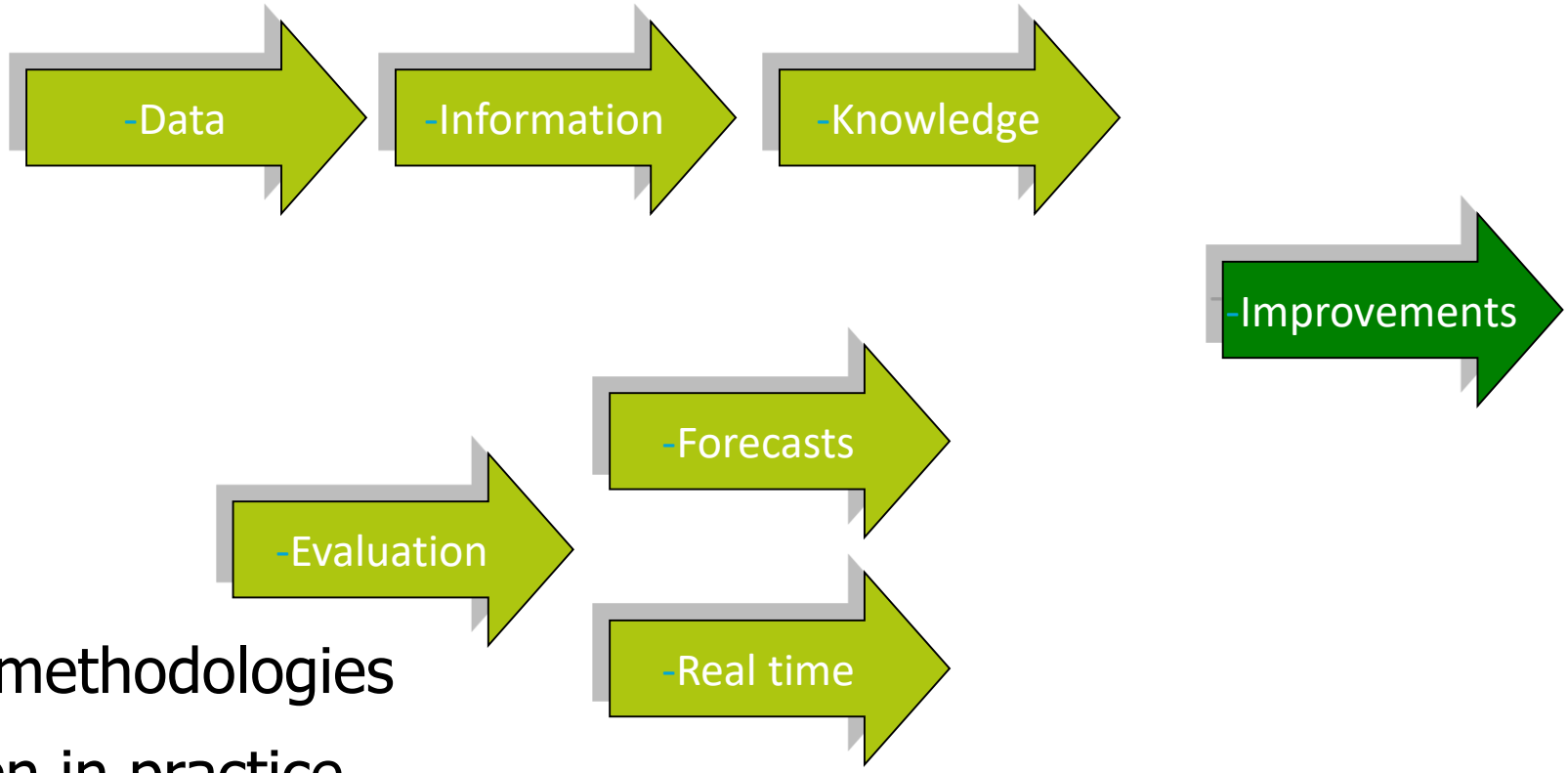

TU Delft Smart Public Transport Lab

Improving public transport



Data + Trends → Understanding → Predicting → Improving

The challenge



- New methodologies
- Proven in practice

Promising data sources

Vehicle monitoring

- Traditional: Manual countings / Selection of vehicles equiped with automatical vehicle location (AVL)
- New: Board computers (AVL)

Passenger monitoring

- Traditional: Manual countings /Selection of vehicles equiped with automatical passenger counting (APC)
- New: Smart card data

Mobility monitoring

- GSM/Mobile phone data

Door to door monitoring

- Traveller app data

1) AVL data

AVL data: The Dutch approach: NDOV

NDOV is a nationwide initiative to make transit data available to authorities and the public.

Focus on dynamic traveler information

Timetable and AVL data available from the majority of the transit vehicles.

(source: GOVI)

NDOV

CROW

Cooperation of 15 transit authorities in NL

Two platforms to share all data (real time and offline)

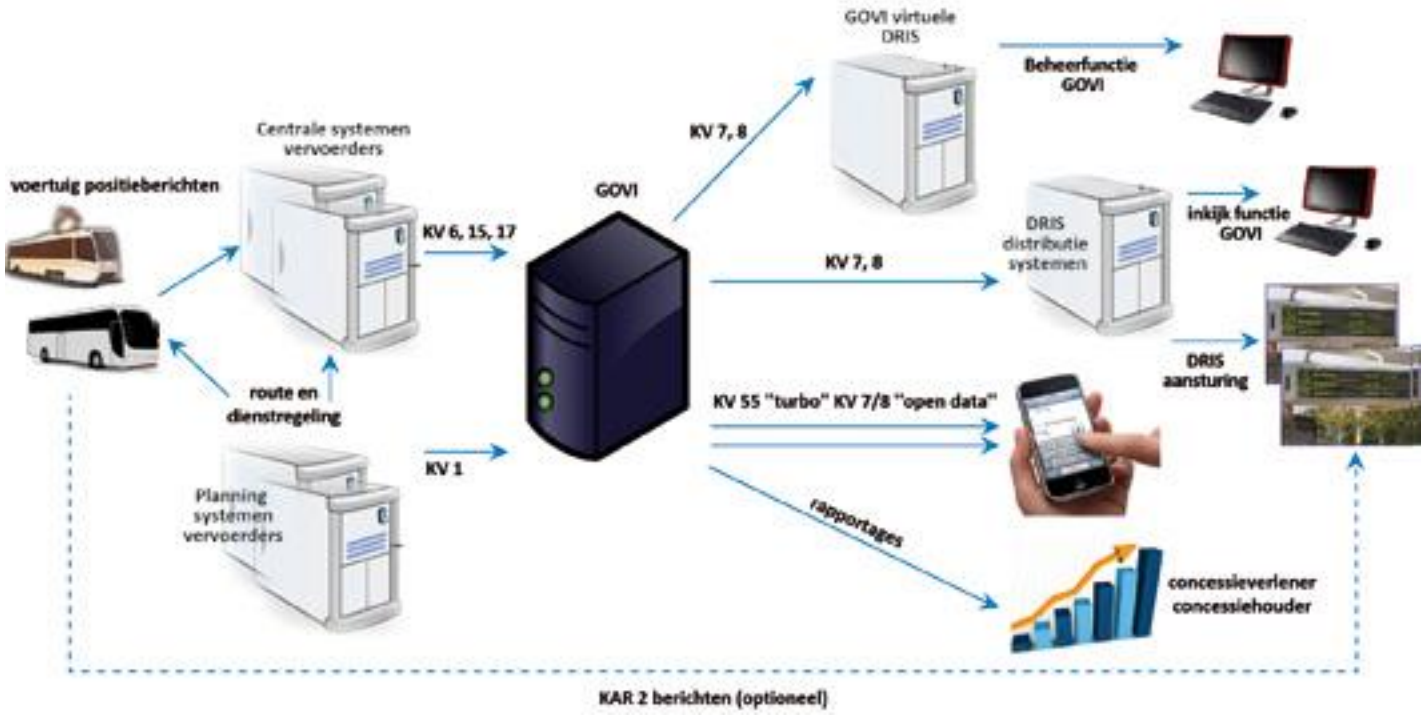
Displays

Standard data format: http://bison.connekt.nl/

The NDOV data architecture

Data interfaces defined in the BISON standard

e.g. KV1 – timetable, KV6 – AVL, KV15 – free text msgs



(source: GOVI)

Real time positions: www.OVradar.nl

The NDOV data architecture

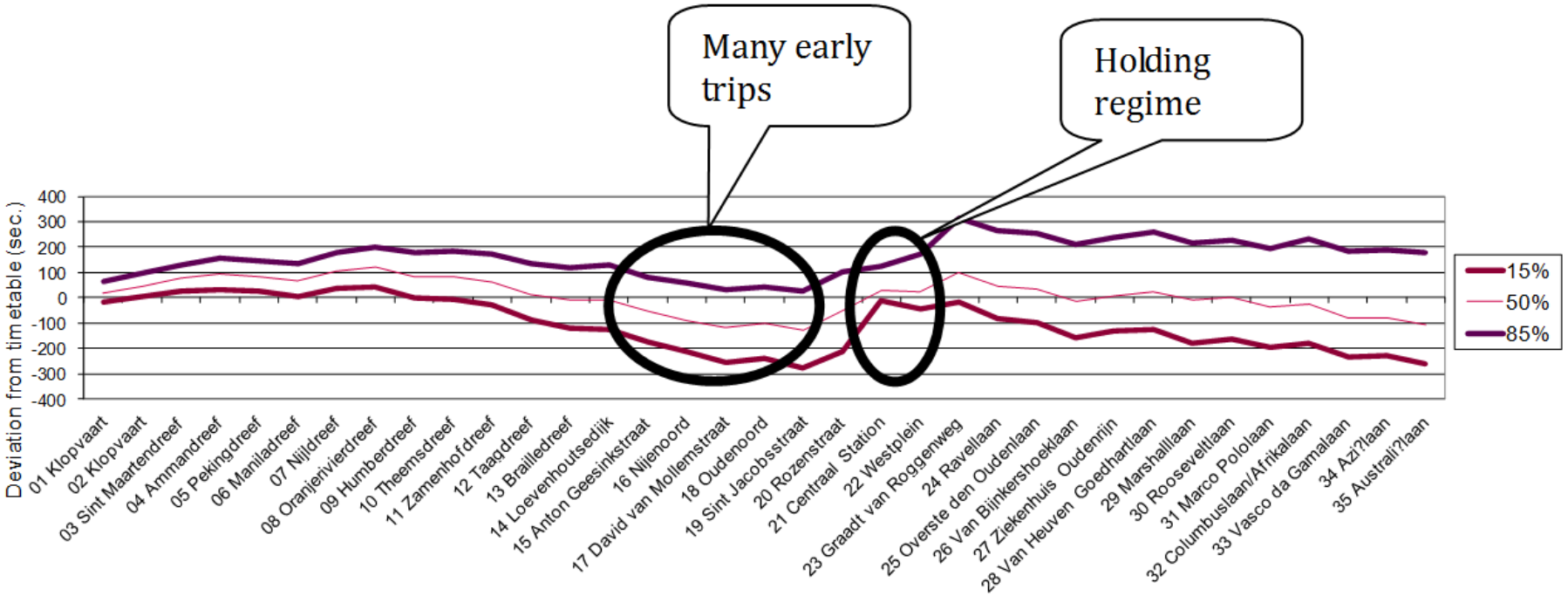
All AVL data is publicly available without restrictions.

Example:

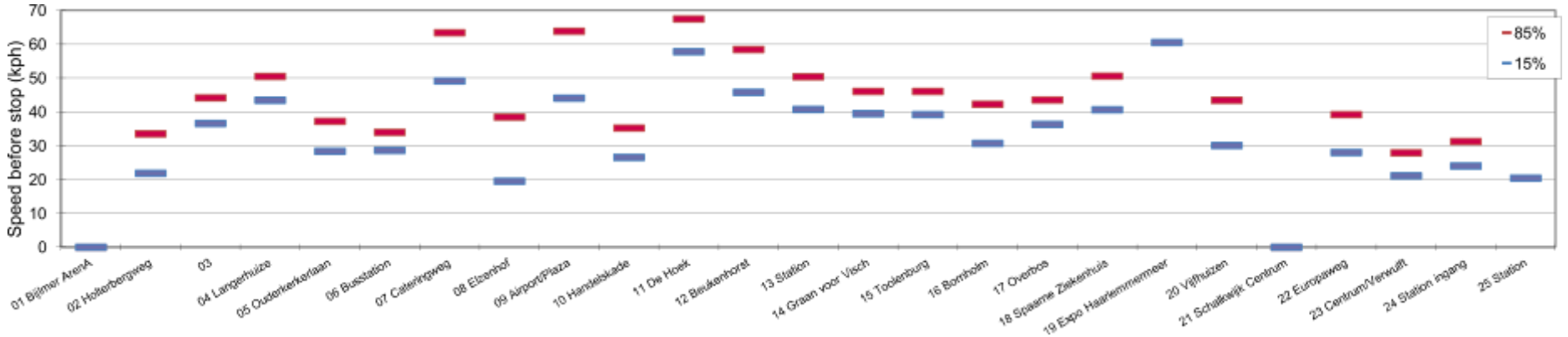
Time	Message type	Operator	Line	Journey	Stop	Punctuality
08:29:00	INIT	...	B120	7001	99990140	
08:29:00	ONSTOP	...	B120	7001	99990140	60
08:29:22	DEPARTURE	...	B120	7001	99990140	82
08:31:28	DEPARTURE	...	B120	7001	99990290	88
...						
08:51:04	ONROUTE	...	B120	7001		
08:52:37	ARRIVAL	...	B120	7001	99990500	-202
08:52:37	END	...	B120	7001	99990500	

Van Oort, N., D. Sparing, T. Brands, R.M.P. Goverde (2015), Data driven improvements in public transport: the Dutch example, Public Transport, Vol 7(3), pp.369-389.

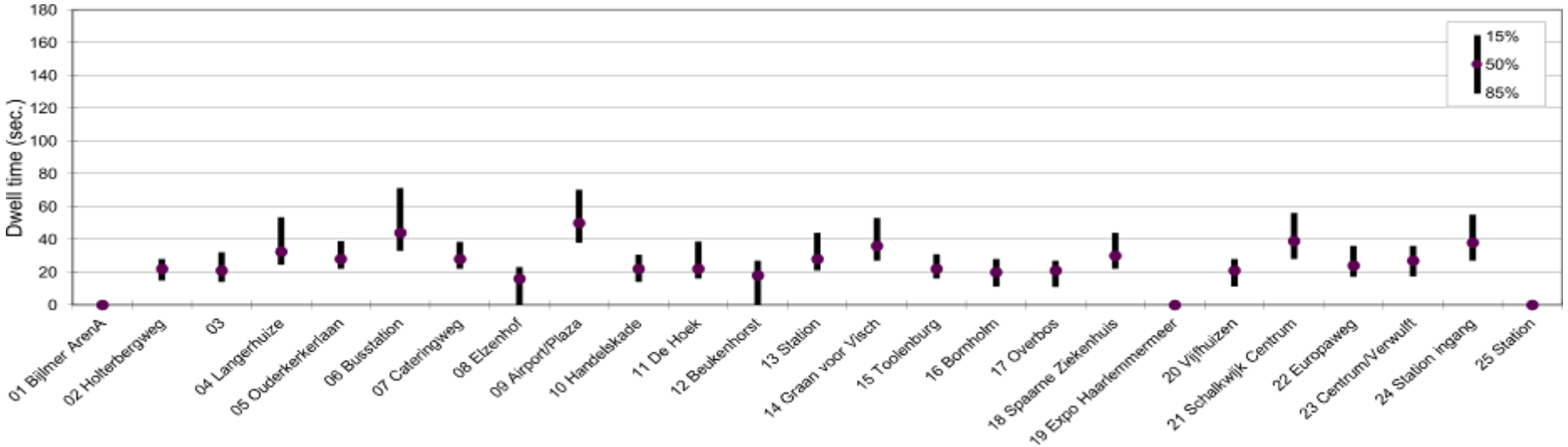
AVL(1)


Many early trips
Holding regime
Deviation from timetable (sec.)
400
300
200
100
0
-100
-200
-300
-400
15%
50%
85%
01 Klopvaart
02 Klopvaart
03 Sint Maartendreef
04 Ammandreef
05 Pekingdreef
06 Maniladreef
07 Nijldreef
08 Oranjerivierdreef
09 Humberdreef
10 Theemsdreef
11 Zamenhofdreef
12 Taagdreef
13 Brailledreef
14 Loevenhoutsedijk
15 Anton Geesinkstraat
16 Nijenoord
17 David van Mollemstraat
18 Oudenoord
19 Sint Jacobsstraat
20 Rozenstraat
21 Centraal Station
22 Westplein
23 Graadt van Roggenweg
24 Ravellaan
25 Overste den Oudenlaan
26 Van Biinkershoeklaan
27 Ziekenhuis Oudenrijn
28 Van Heuven Goedhartlaan
29 Marshalllaan
30 Rooseveltlaan
31 Marco Pololaan
32 Columbuslaan/Afrikalaan
33 Vasco da Gamalaan
34 Azi?laan
35 Australi?laan

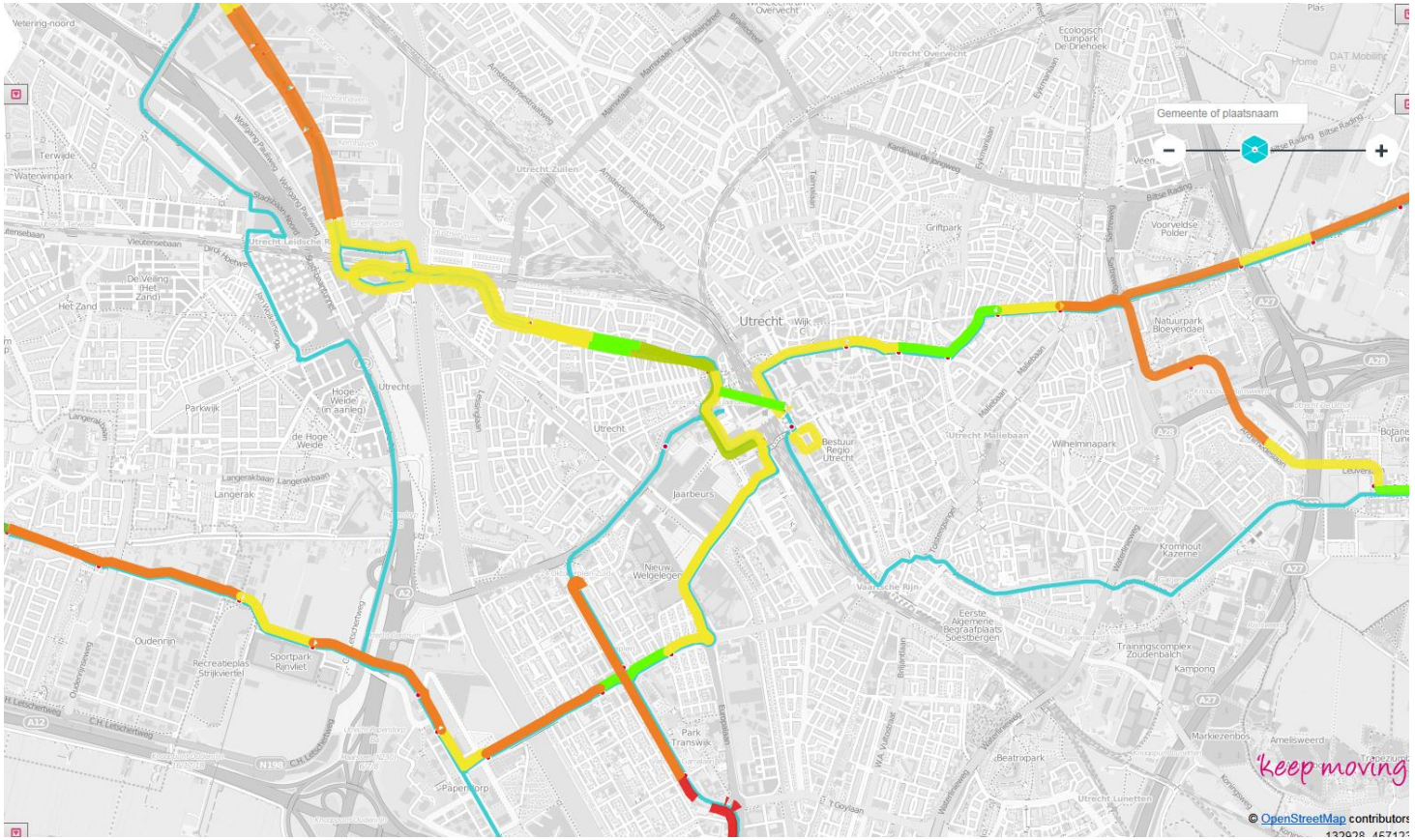
AVL (2)


Speed before stop (kph)
70
60
50
40
30
20
10
0
85%
15%
01 Bijlmer ArenA
02 Holterbergweg
03
04 Langerhuize
05 Ouderkerkerlaan
06 Busstation
07 Cateringweg
08 Elzenhof
09 Airport/Plaza
10 Handelskade
11 De Hoek
12 Beukenhorst
13 Station
14 Graan voor Visch
15 Toolenburg
16 Bornholm
17 Overbos
18 Spaarne Ziekenhuis
19 Expo Haarlemmermeer
20 Vijfhuizen
21 Schalkwijk Centrum
22 Europaweg
23 Centrum/Verkuijlt
24 Station ingang
25 Station

AVL(3)


Dwell time (sec.)
180
160
140
120
100
80
60
40
20
0
15%
50%
85%
01 Bijlmer ArenA
02 Holterbergweg
03
04 Langerhuize
05 Ouderkerkerlaan
06 Busstation
07 Cateringweg
08 Elzenhof
09 Airport/Plaza
10 Handelskade
11 De Hoek
12 Beukenhorst
13 Station
14 Graan voor Visch
15 Toolenburg
16 Bornholm
17 Overbos
18 Spaarne Ziekenhuis
19 Expo Haarlemmermeer
20 Vijfhuizen
21 Schalkwijk Centrum
22 Europaweg
23 Centrum/Verwulft
24 Station ingang
25 Station

AVL(4)


Gemeente of plaatsnaam
Utrecht
keep moving
© OpenStreetMap contributors

Bottleneck detection

Brands, T., N. van Oort, M. Yap (2018),

Automatic bottleneck detection using AVL data: a case study in Amsterdam,

Conference on Advanced Systems in Public Transport and TransitData (CASPT), Brisbane, Australia.

Bottleneck definition

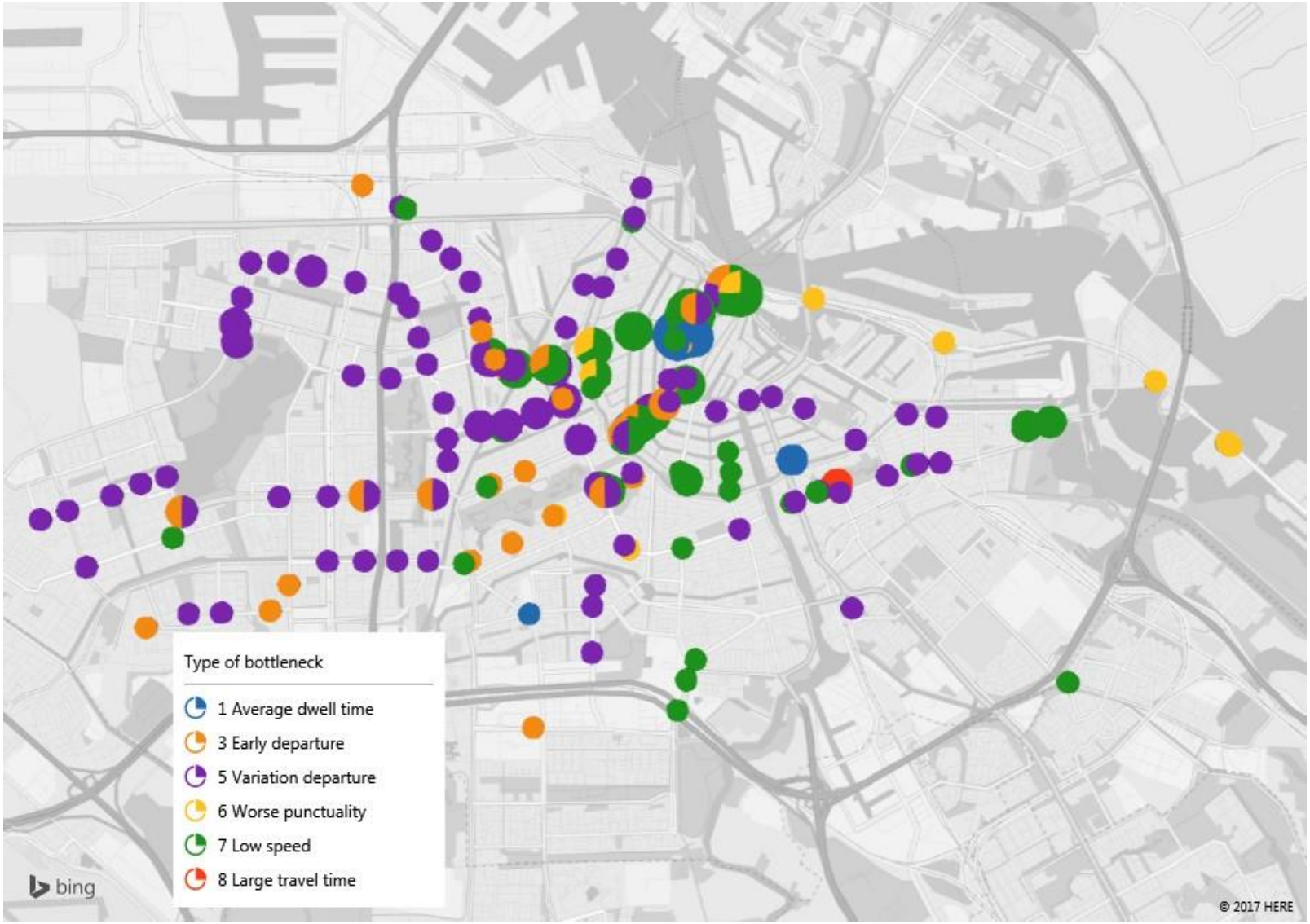
- Contract requirements, literature and expert judgement

• Dwell time	> 60s.
• Variation dwell time (85 and 15 percentile)	> 120s.
• Early departure	< -60s.
• Late departure	>180s.
• Variation departure time: (85 and 15 percentile)	300s.
• Punctuality change compared to previous stop	> 60s.
• Speed	< 15 km/h.
• Trip time compared to free flow (15th percentile of the travel time on Sundays)	> 60s.

Van Oort, N. (2014),
Incorporating service reliability in public transport design and performance requirements: International survey results and recommendations, Research in Transportation Economics, Volume 48, pp 92-100.

Results: bottlenecks

Line	Direction	Period	Stop	Stop number	Average dwell time	Dwell time variation	Punctuality 50%	Variation departure time	Punctuality change wrt previous stop	Average speed	Difference with free flow travel time
10	Back	Evening	14Elandsgracht	6028	20	13	-28	217	65	30	18
10	Back	Saturday	13Leidseplein	6061	31	25	66	182	18	14	17
10	Back	Sunday	13Leidseplein	6061	29	22	32	163	14	15	14
10	Back	Sunday	14Elandsgracht	6028	20	17	-42	177	73	32	12
2	Forth	AM peak	09Hoofddorpplein	7049	25	18	73	199	45	13	39
2	Forth	AM peak	10Amstelveenseweg	7034	22	15	95	223	22	15	28
2	Forth	AM peak	13Van Baerlestraat	7322	23	11	50	303	2	18	37


Type of bottleneck
1 Average dwell time
3 Early departure
5 Variation departure
6 Worse punctuality
7 Low speed
8 Large travel time
bing
© 2017 HERE

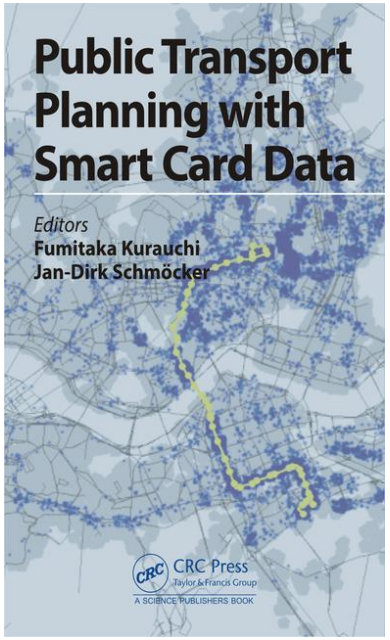
2) Smartcard data

Smartcard data

The Netherlands
- OV Chipkaart
- Nationwide (since 2012)
- All modes: train, metro, tram, bus
- Tap in and tap out
- Bus and tram: devices are in the vehicle

Issues
- Privacy
- Data accaccessibility via operators

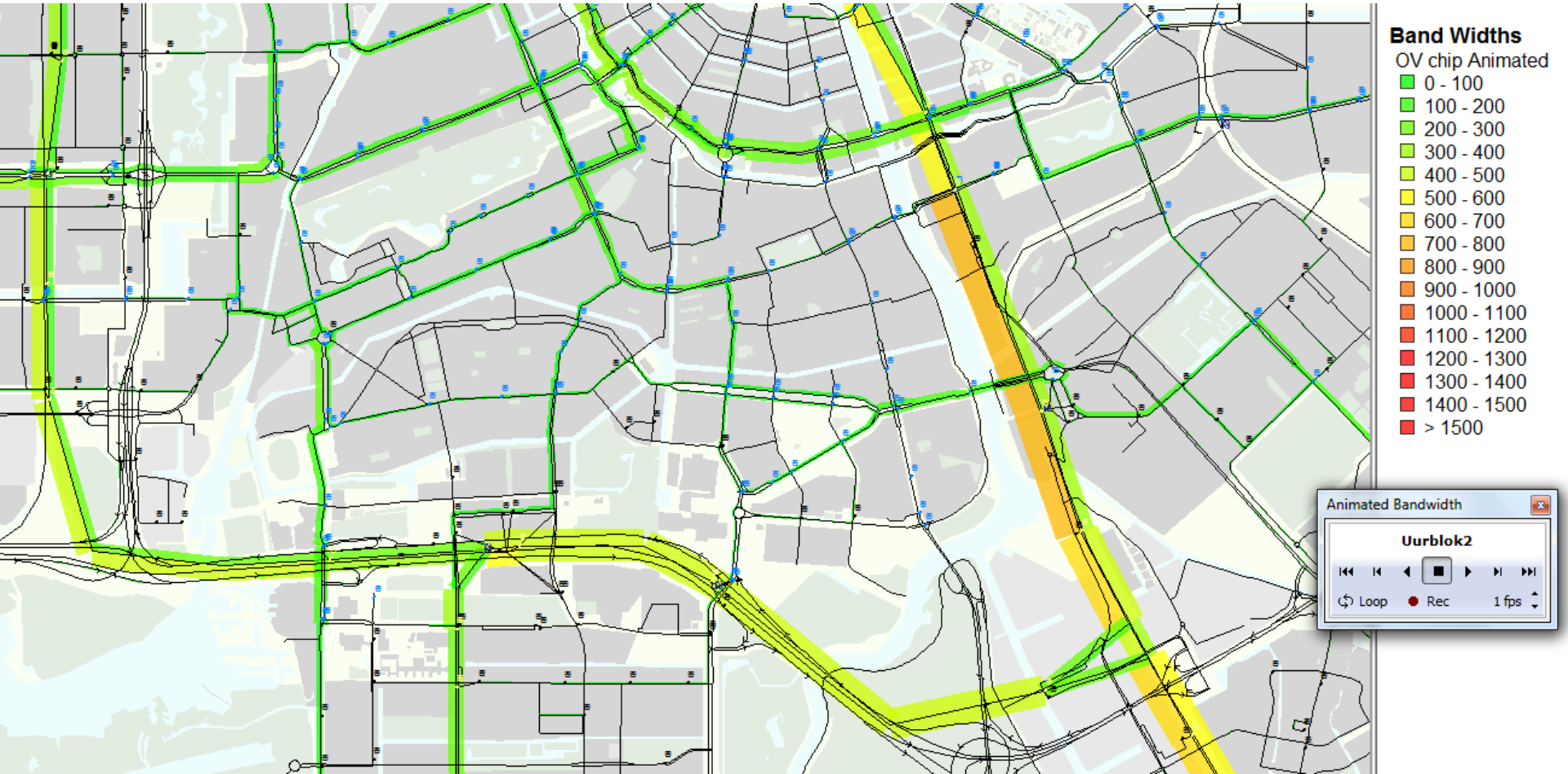
- Several applications of smartcard data:
Pelletier et. al (2011). Transportation Research Part C



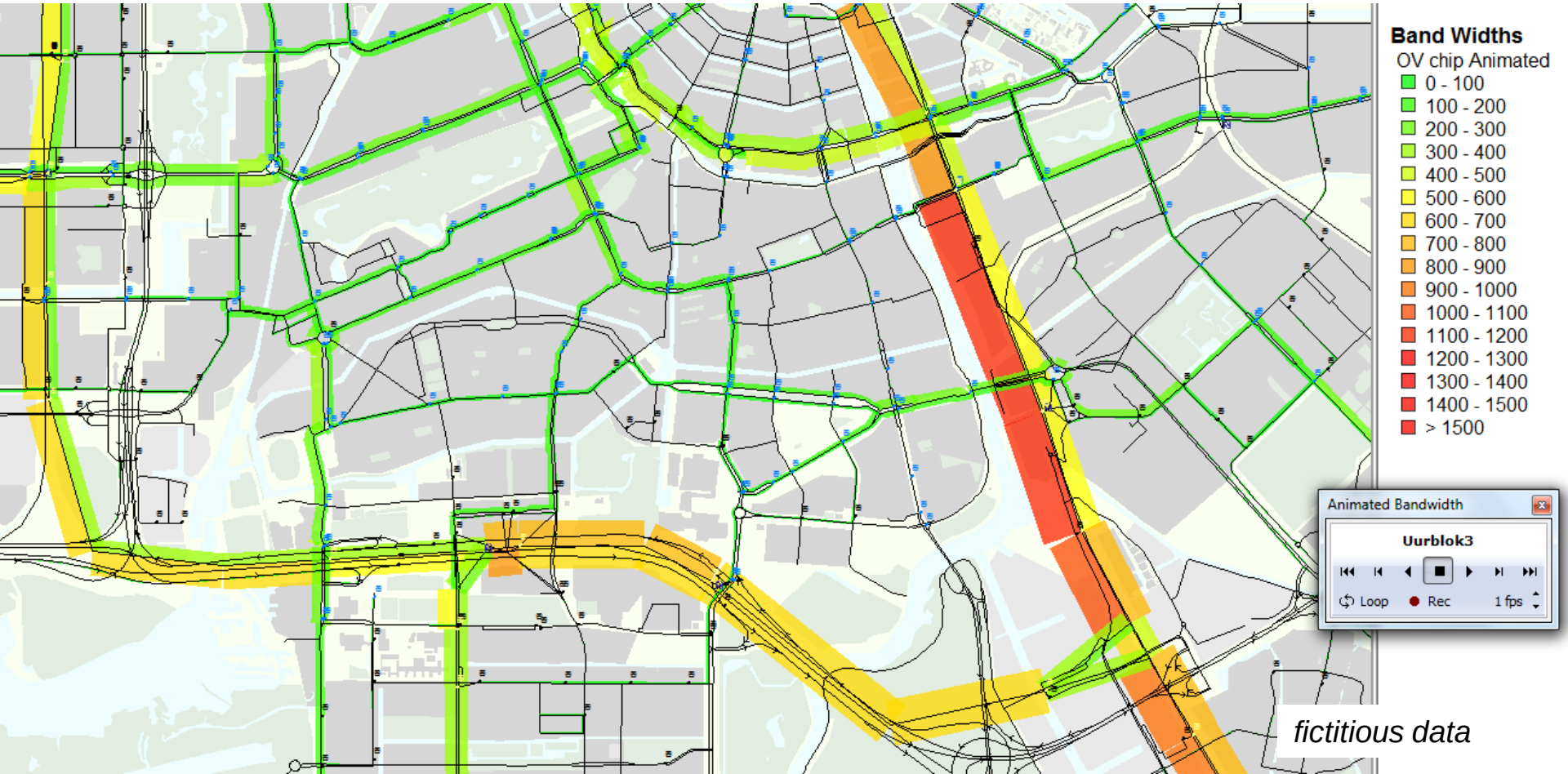
translink

- Task: introduce and maintain smartcard system in NL
- Owned by all PT operators in NL
- >25 million smart card issued
- > 2 billion transactions annually

Examples: Monitoring and predicting passenger numbers

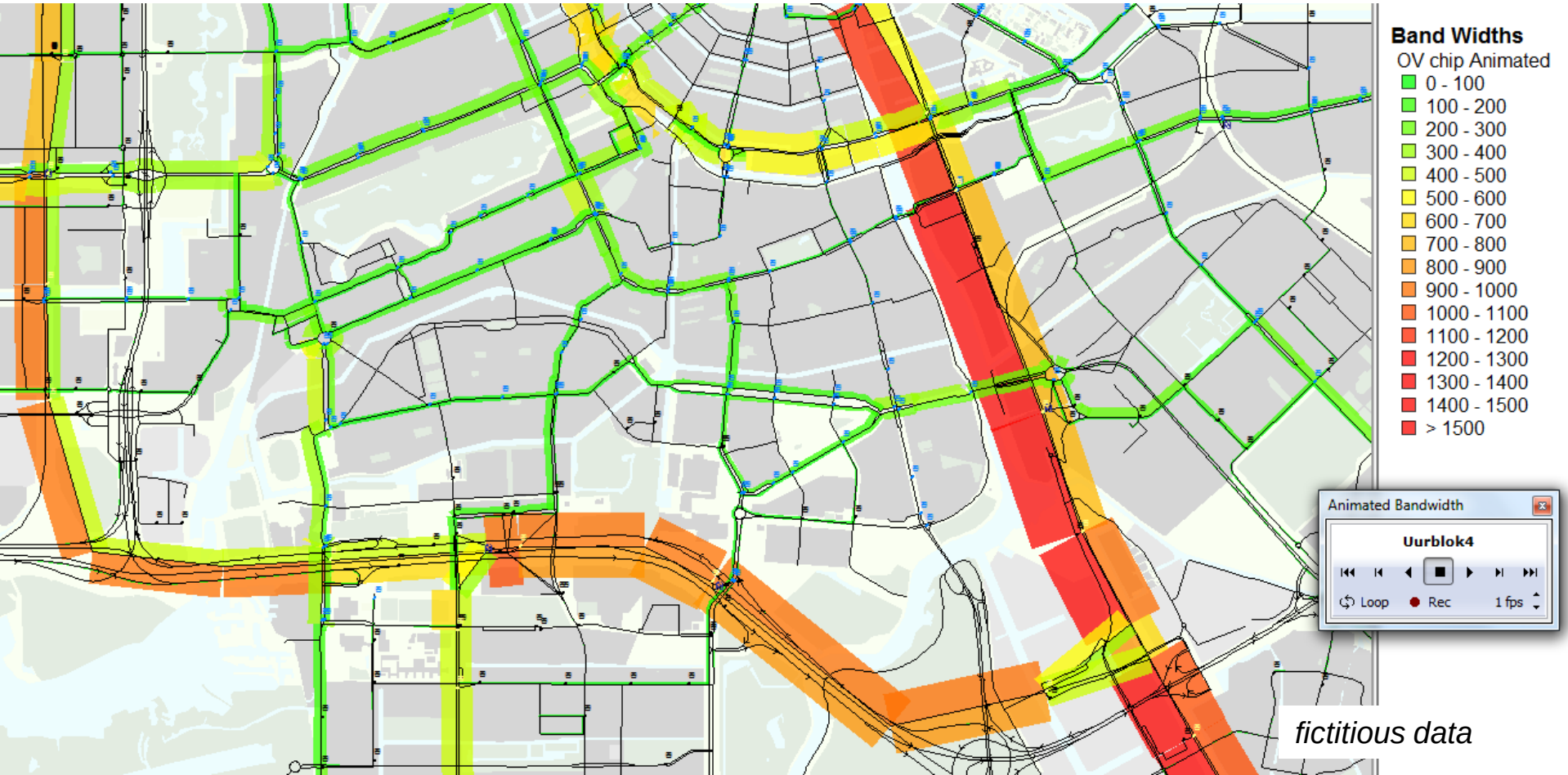


Examples: Monitoring and predicting passenger numbers



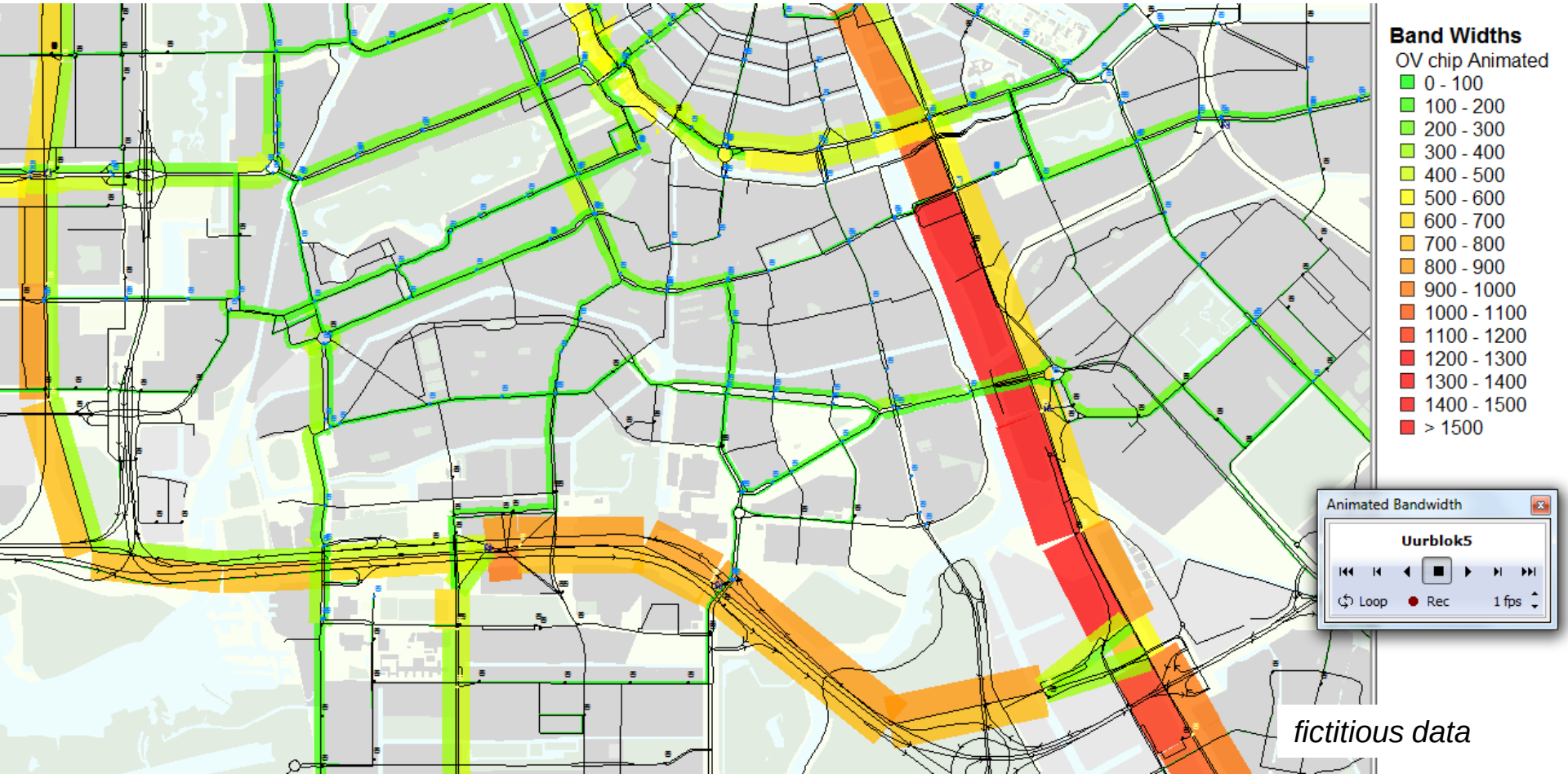
fictitious data

Examples: Monitoring and predicting passenger numbers



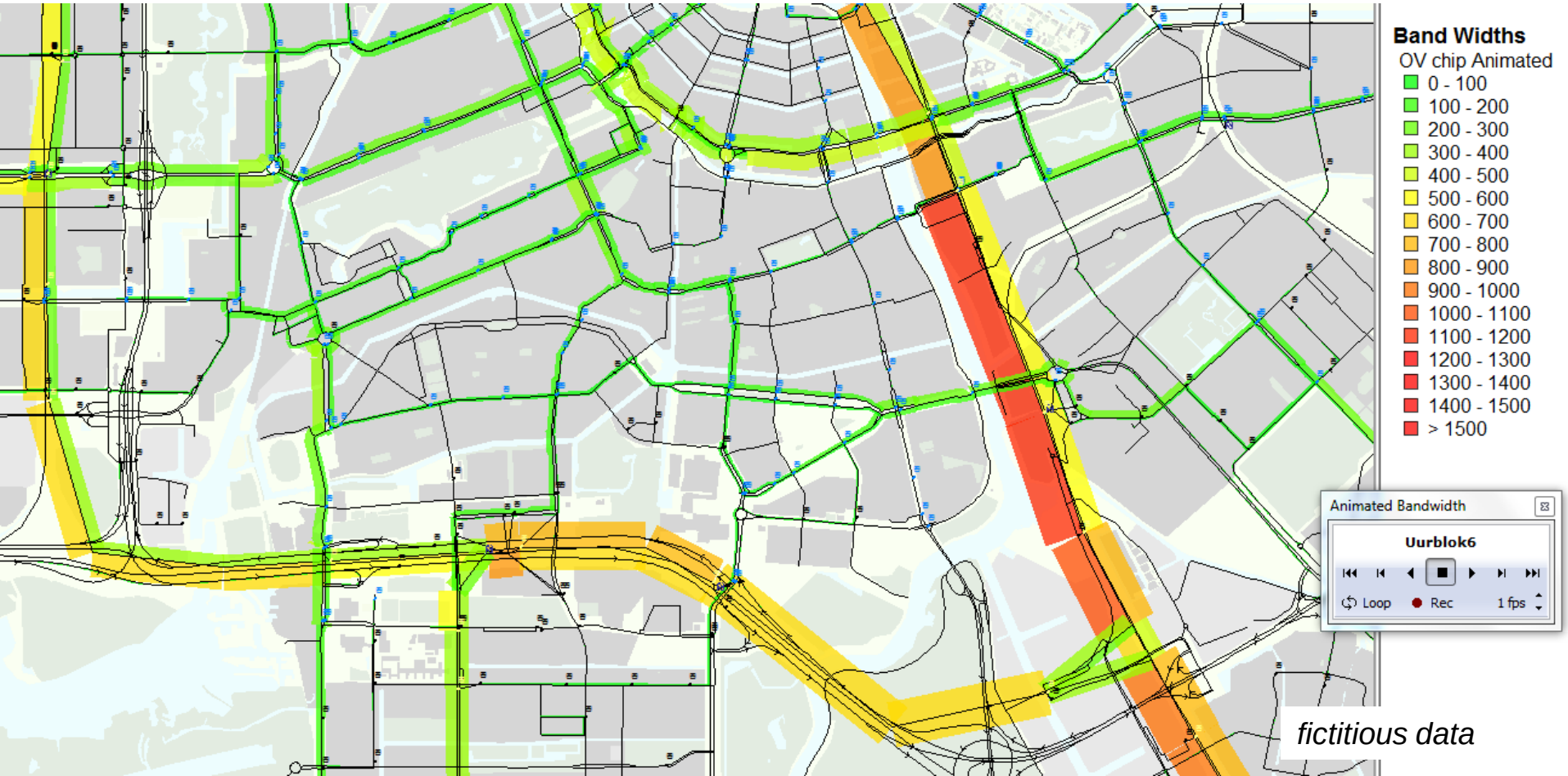
fictitious data

Examples: Monitoring and predicting passenger numbers



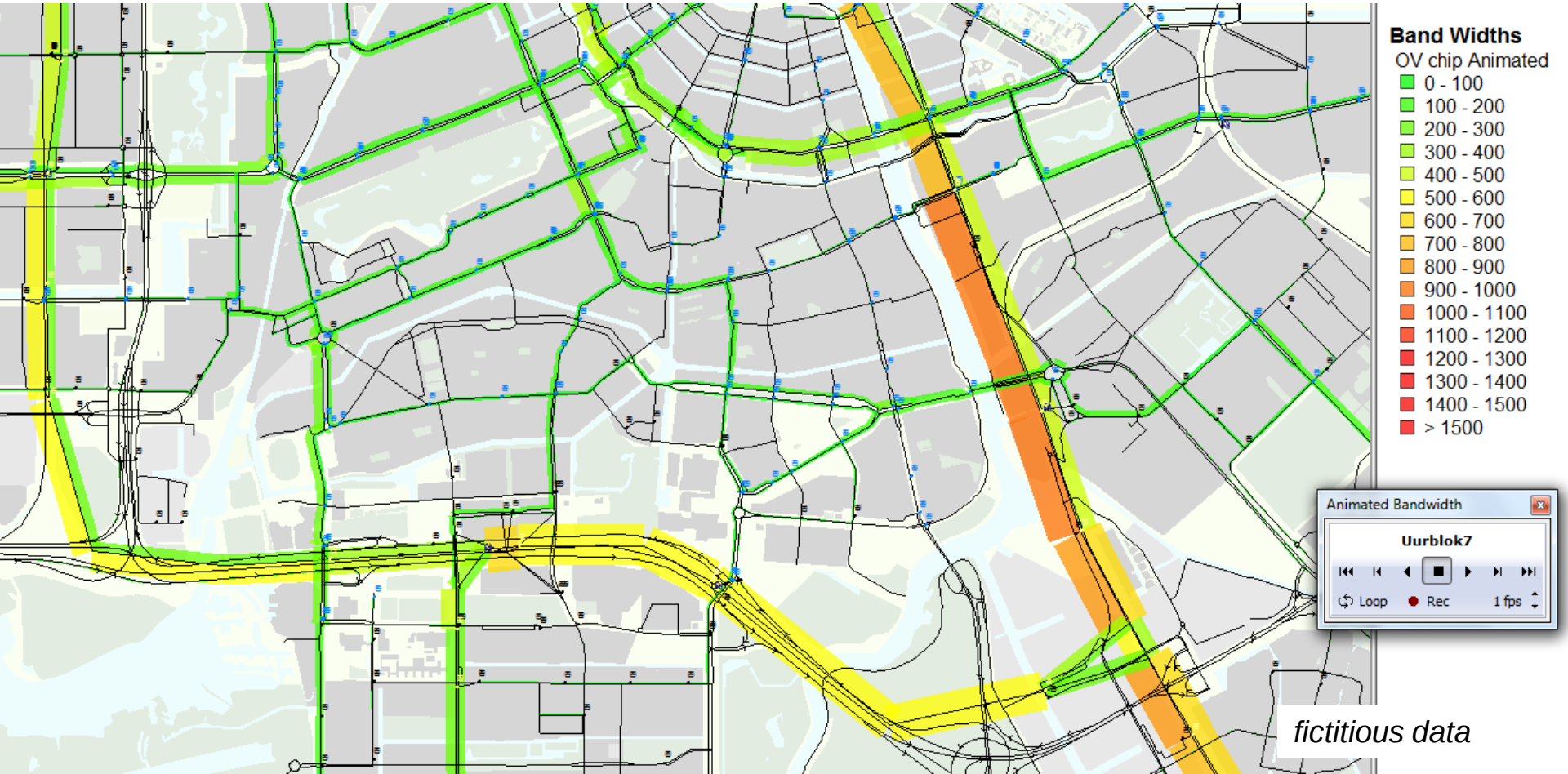
fictitious data

Examples: Monitoring and predicting passenger numbers



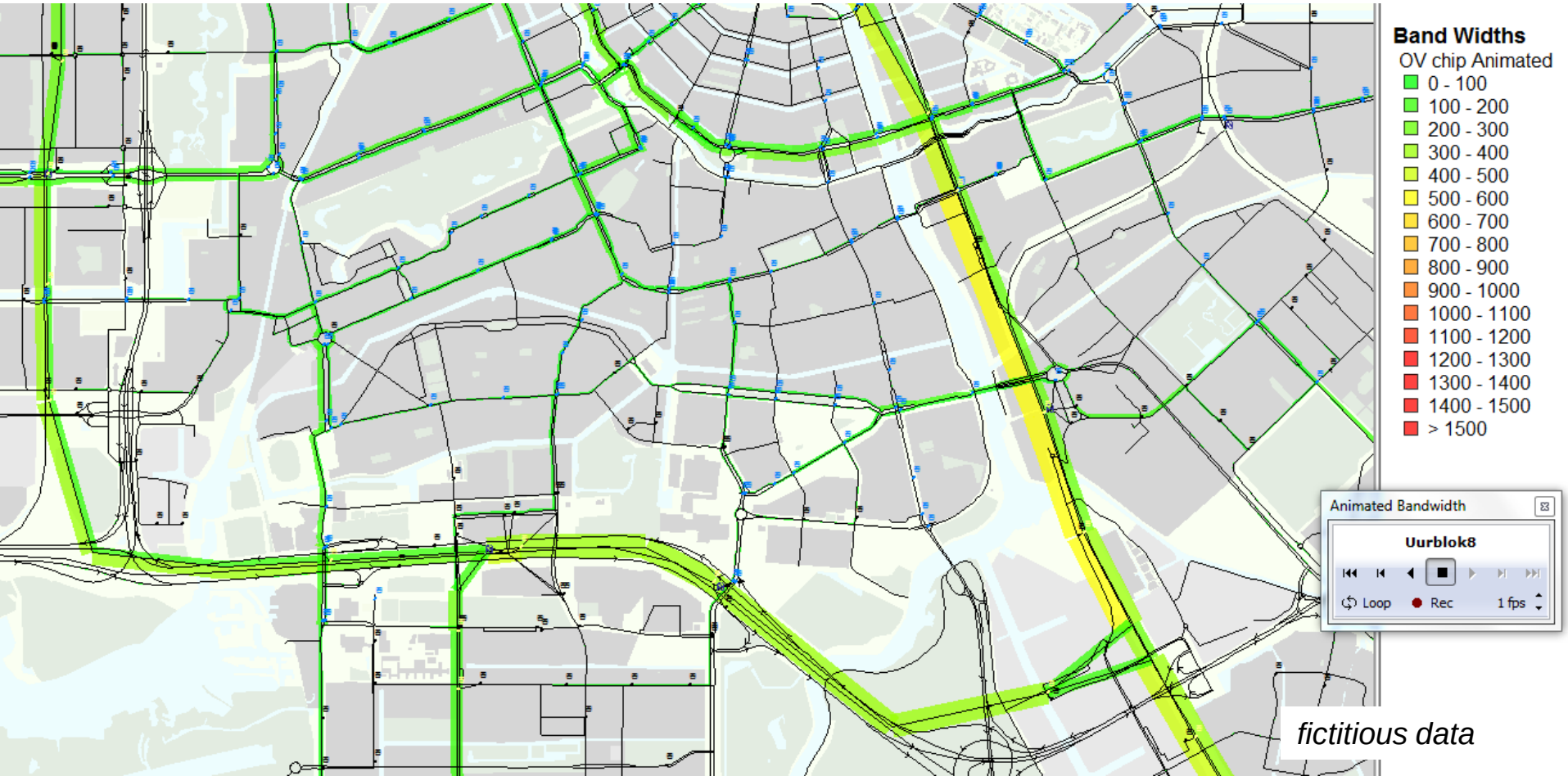
fictitious data

Examples: Monitoring and predicting passenger numbers



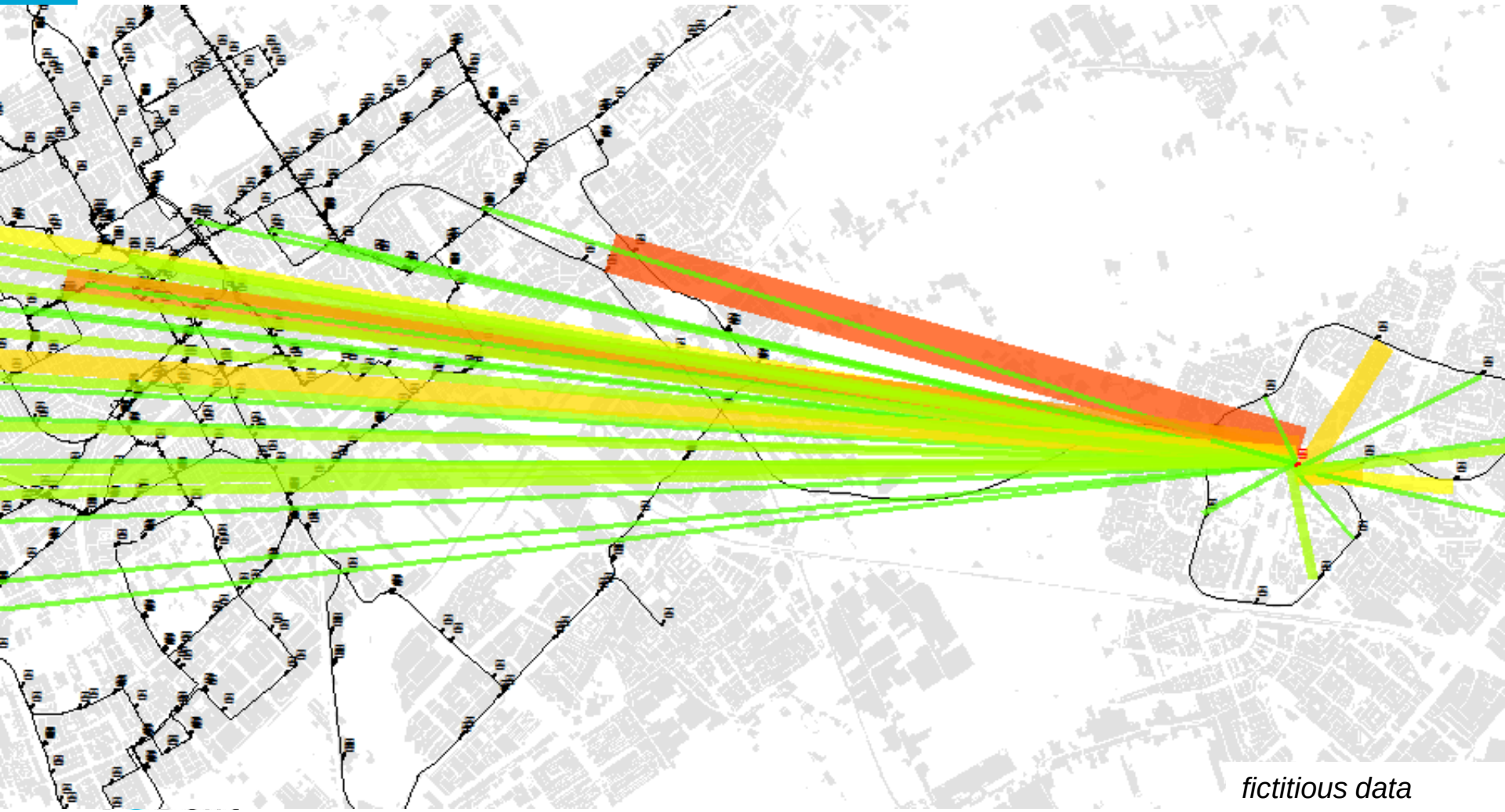
fictitious data

Examples: Monitoring and predicting passenger numbers



fictitious data

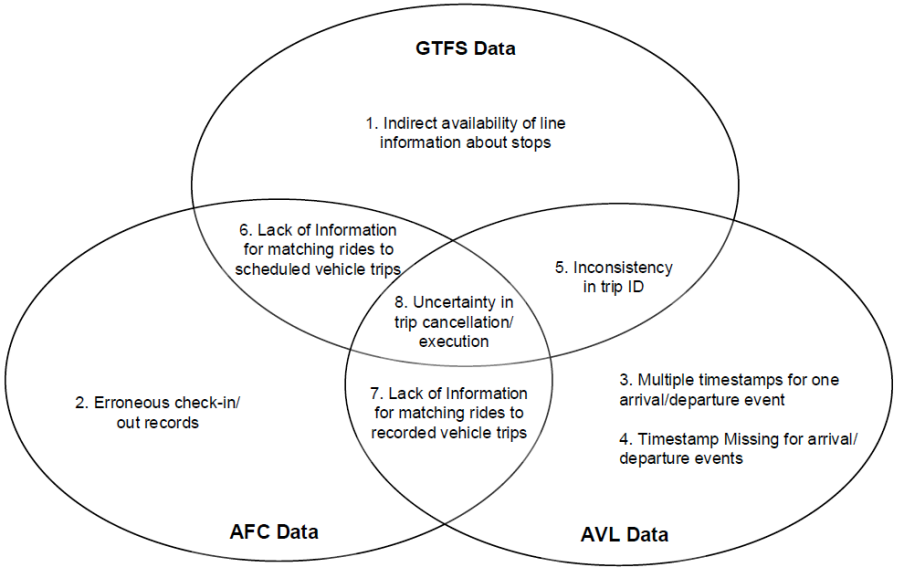
OD-patterns



fictitious data

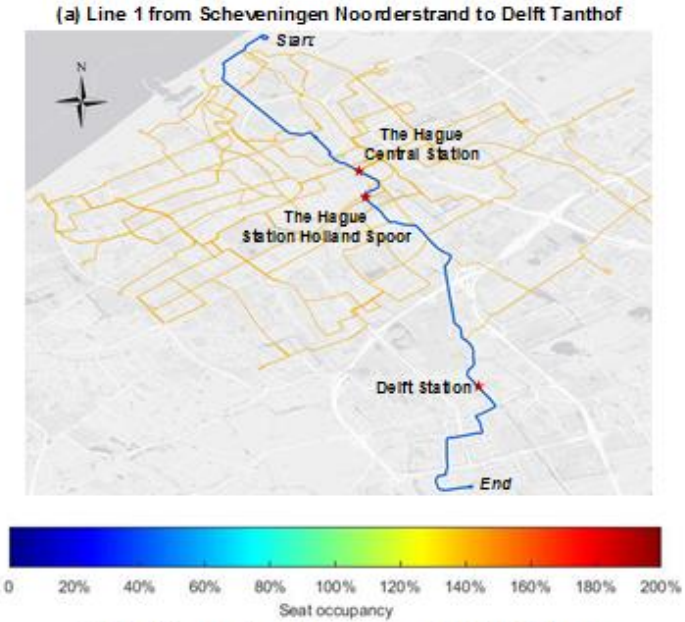
Spatiotemporal Load Profiles

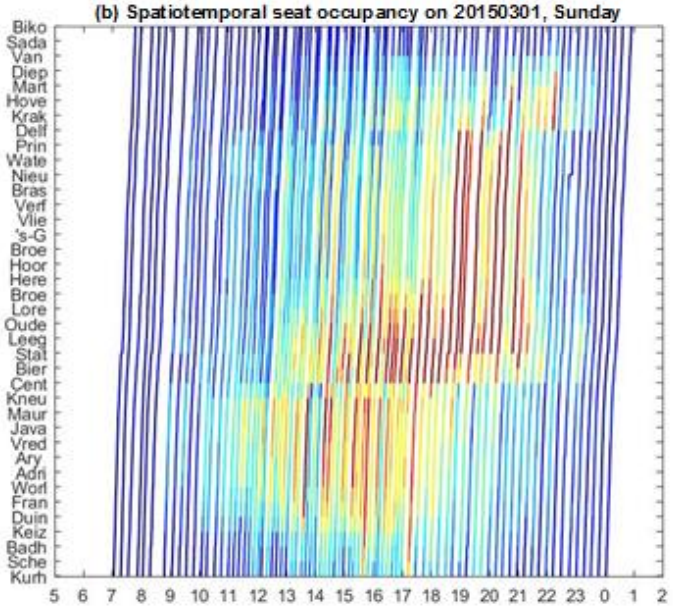
- Spatiotemporal load profiles of transit vehicles can be constructed using multiple data sources.
- Combination of data sets: AFC/smartcard, AVL and GTFS.

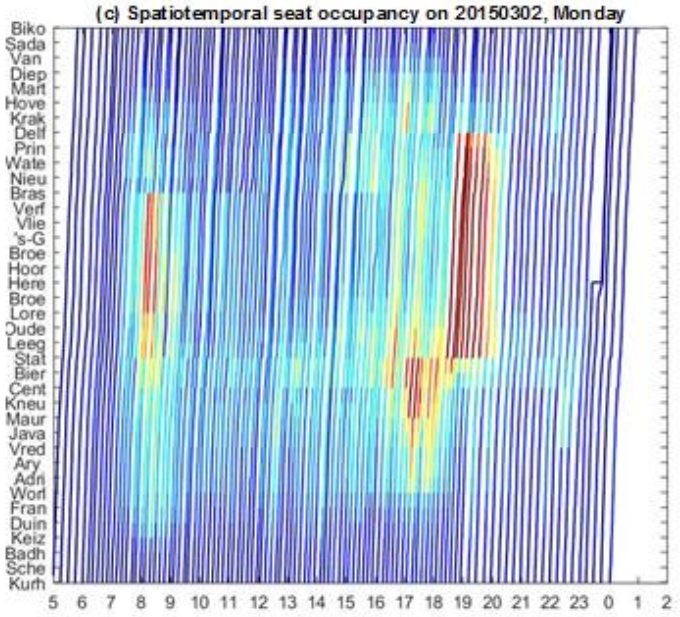


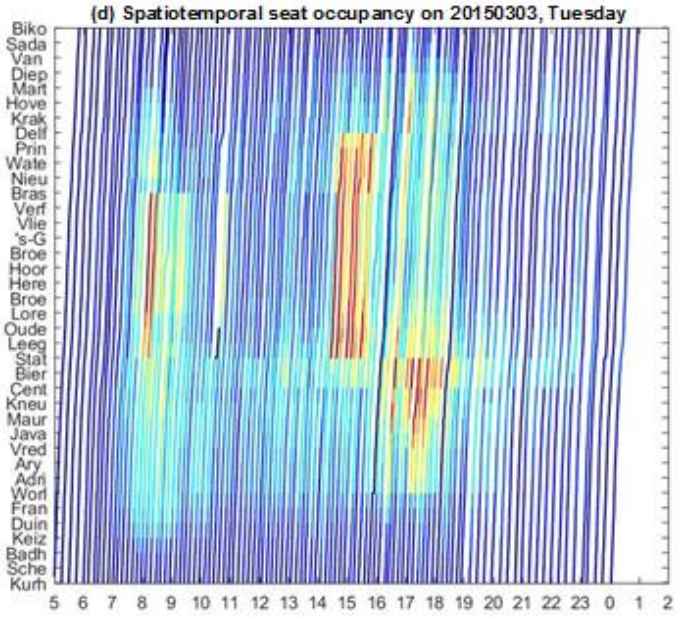
Luo, D., Bonnetain, L., Cats, O. & van Lint, H. (2018) Constructing spatiotemporal load profiles of transit vehicles with multiple data sources. Transportation Research Record: Journal of the Transportation Research Board.

Acknowledgement: HTM and Stichting OpenGeo provided the AFC and AVL datasets, resp.


(a) Line 1 from Scheveningen Noorderstrand to Delft Tanthof
Start
N
The Hague
Central Station
The Hague
Station Holland Spoor
Delft Station
End
0
20%
40%
60%
80%
100%
120%
140%
160%
180%
200%
Seat occupancy

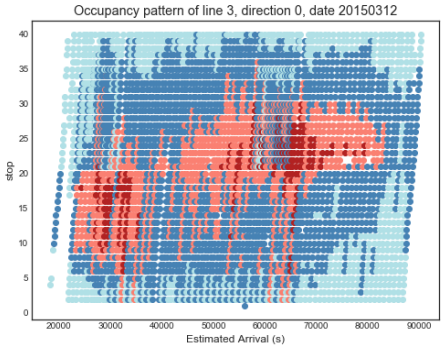

(b) Spatiotemporal seat occupancy on 20150301, Sunday
Biko
Sada
Van
Diep
Mart
Hove
Krak
Delf
Prin
Wate
Nieu
Bras
Verf
Vlie
's-G
Broe
Hoor
Here
Broe
Lore
Oude
Leeg
Stat
Bier
Cent
Kneu
Maur
Java
Vred
Ary
Adri
Worl
Fran
Duin
Keiz
Badh
Sche
Kurh
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 0 1 2

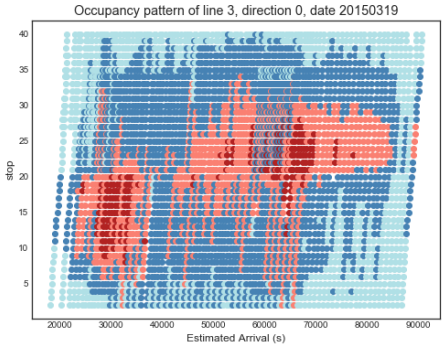

(c) Spatiotemporal seat occupancy on 20150302, Monday
Biko
Sada
Van
Diep
Mart
Hove
Krak
Delf
Prin
Wate
Nieu
Bras
Verf
Vlie
's-G
Broe
Hoor
Here
Broe
Lore
Oude
Leeg
Stat
Bier
Cent
Kneu
Maur
Java
Vred
Ary
Adri
Worl
Fran
Duin
Keiz
Badh
Sche
Kurh
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 0 1 2

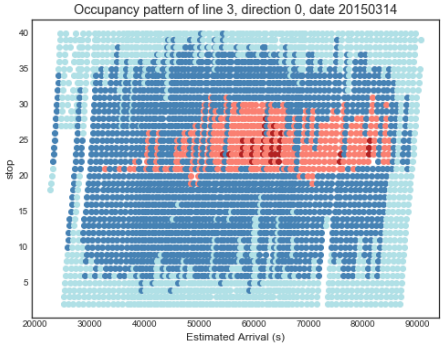

(d) Spatiotemporal seat occupancy on 20150303, Tuesday
Biko
Sada
Van
Diep
Mart
Hove
Krak
Delf
Prin
Wate
Nieu
Bras
Verf
Vlie
's-G
Broe
Hoor
Here
Broe
Lore
Oude
Leeg
Stat
Bier
Cent
Kneu
Maur
Java
Vred
Ary
Adri
Worl
Fran
Duin
Keiz
Badh
Sche
Kurh
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 0 1 2

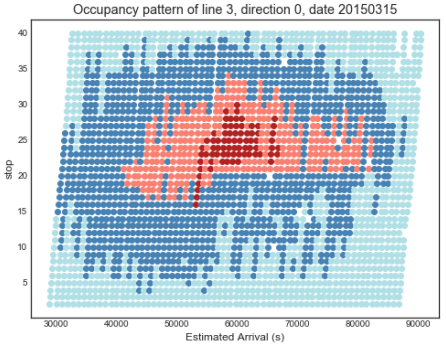
Occupancy pattern clustering and recognition (Artificial intelligence)

Legend:
- Almost empty
- Sit alone
- Sit next to someone
- All seats occupied



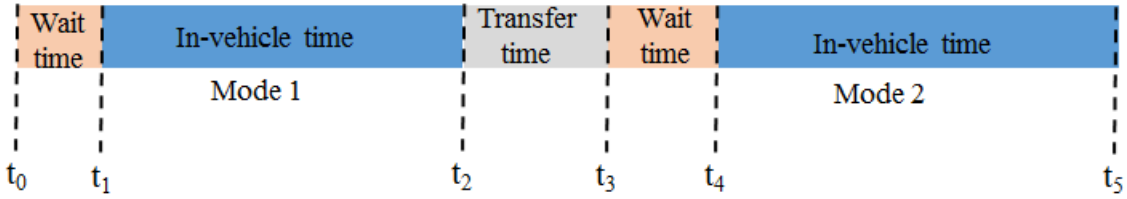


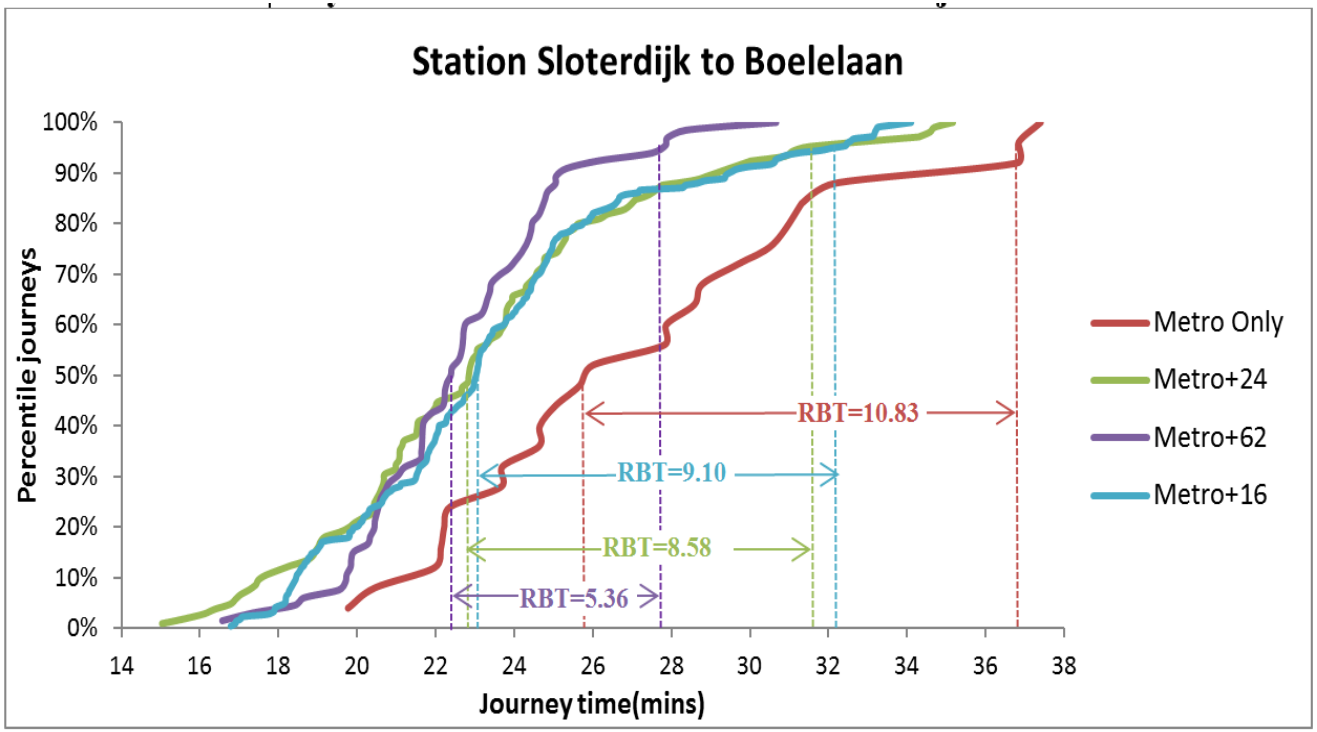




Heydenrijk-Ottens et al. (2018), Supervised learning: Predicting passenger load in public transport, CASPT conference, Brisbane

Passenger service reliability





AVL+Smartcard

Dixit et al. (2019). TRB

What-if analysis with smart card data

PT modelling

Traditional (4-step) model
Multimodal (~PT)
Network
Complex
Long calculation time
Visualisation
Much data
Detailed results

Simple calculation
PT only
Line
Transparent
Short calculation time
Only numbers
Little data
Assessments

Short term predictions

Elasticity method based on smartcard data

Combining models and smartcard data

Connecting to transport model

- Evaluating history
- Predicting the future

- Whatif scenario's
  - Stops: removing or adding
  - Faster and higher frequencies
  - Route changes

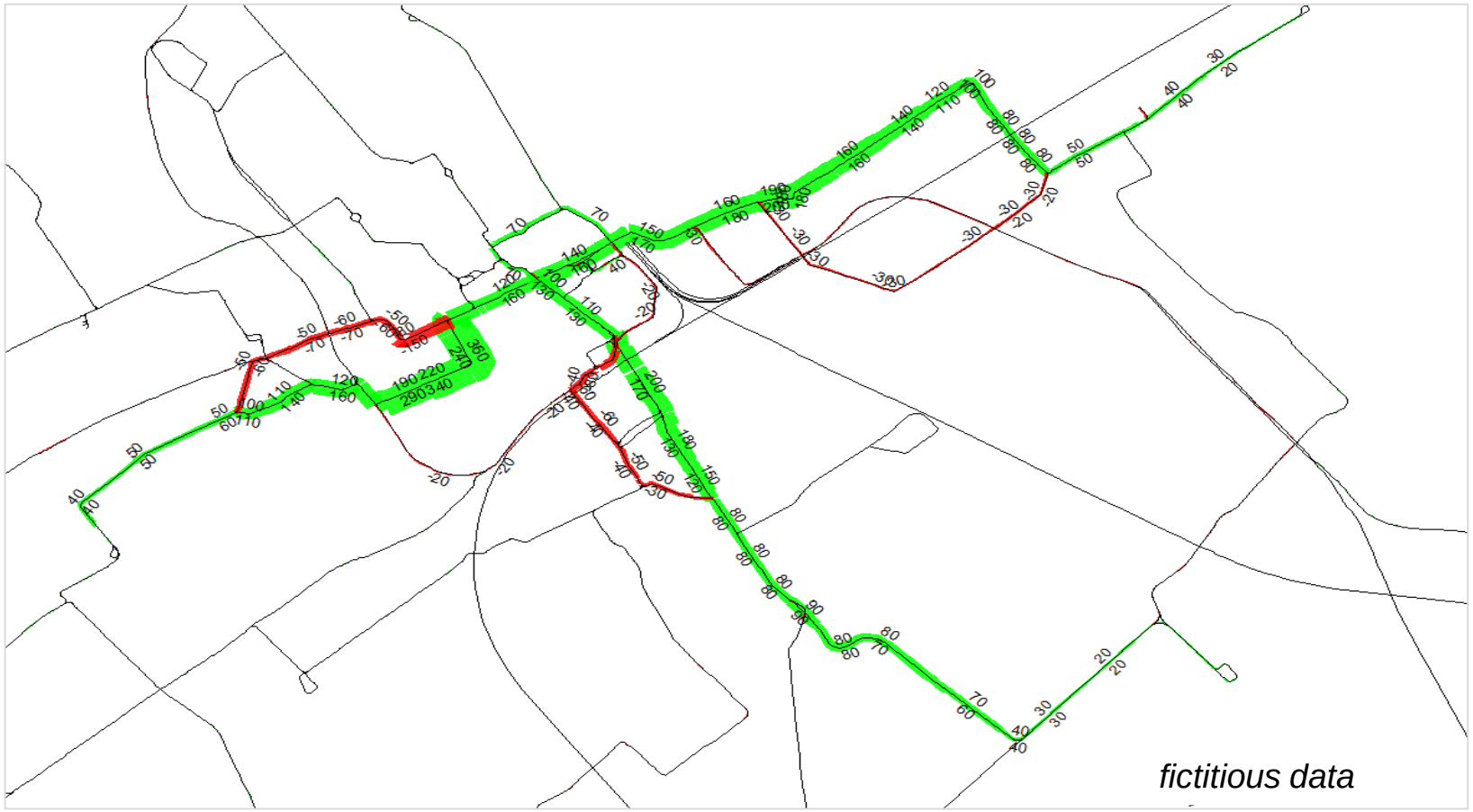
- Quick insights into
  - Expected cost coverage
  - Expected occupancy

OmniTRANS
Transport Planning Software

New generation of transport models: data driven

Van Oort, N., T. Brands, E. de Romph (2015), Short-Term Prediction of Ridership on Public Transport with Smart Card Data, Transportation Research Record, No. 2535, pp. 105-111.

Whatif: increased frequencies



Whatif: increased speed

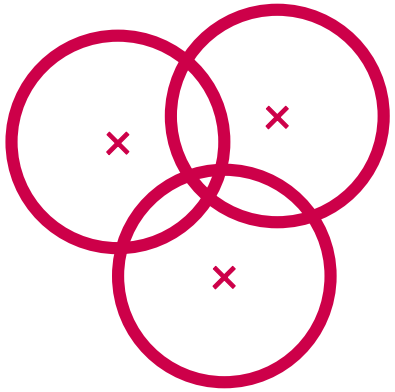
fictitious data

3) GSM/cell phone data

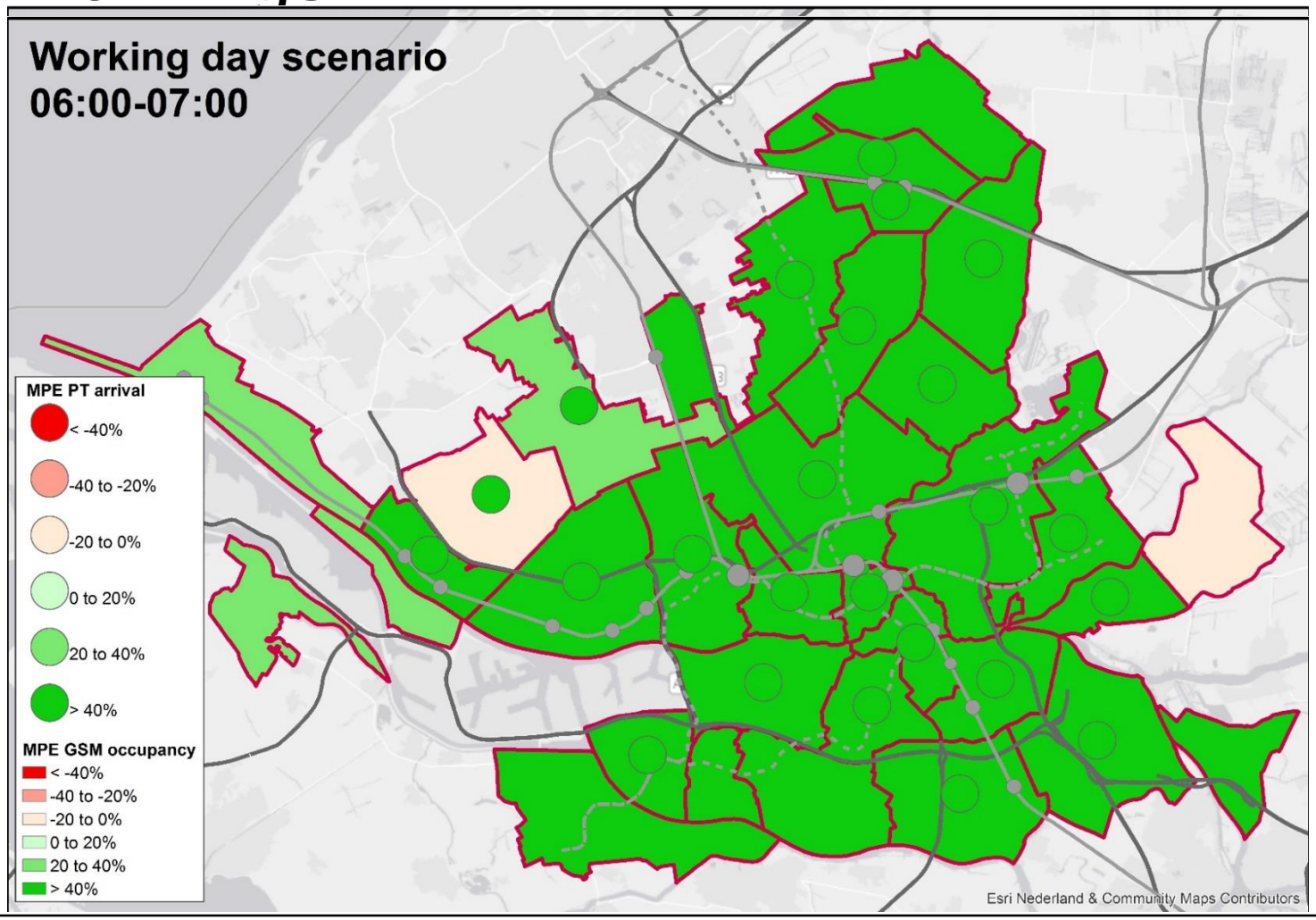
Data pre-processing
GSM data

Value: Amount of visitors detected in a zone, on a specific day, per hour

- Data from one network operator
  - Algorithm to increase sample data to total population
- Distinction inhabitants or visitors
  - Place of residence based on overnight stays per month
- Spatial level of detail: zone level
  - Antennas have overlapping reach
- Difference occupancy between subsequent hours is a net change



Rotterdam late evenings & early mornings



More applied data examples
http://nielsvanoort.weblog.tudelft.nl/

Cost benefit analysis
Van Oort, N. (2016). Incorporating enhanced service reliability of public transport in cost-benefit analyses. Public Transport, Volume 8 (1), pp 143-160.

Optimizing planning and real time control
Van Oort, N. and R. van Nes (2009), Control of public transport operations to improve reliability: theory and practice, Transportation research record, No. 2112, pp. 70-76.

Durand, A., N. van Oort and S. Hoogendoorn (2018). Assessing and Improving Operational Strategies for the Benefit of Passengers in Rail-Bound Urban Public Transport Systems. Transportation Research Record.

Optimizing synchronization multimodal transfers
Lee, A. N. van Oort, R. van Nes (2014), Service reliability in a network context: impacts of synchronizing schedules in long headway services, TRB

Yap, M.D., O. Cats, N. van Oort, S.P. Hoogendoorn (2017), A robust transfer inference algorithm for public transport journeys during disruptions, Transportation Research Procedia 27 .

Improved scheduling
Van Oort, N. et al. (2012). The impact of scheduling on service reliability: trip time determination and holding points in long-headway services. Public Transport, 4(1), 39-56.

Van Oort, N., N.H.M. Wilson, R. van Nes (2010), Reliability improvement in short headway transit services: schedule-based and headway-based holding strategies, Transportation Research Record, No. 2143, pp.67-76.

Passenger behavior and modelling
Yap, M.D., Nijenstein, S., & van Oort, N. (2018). Improving predictions of public transport usage during disturbances based on smart card data. Transport Policy, Vol. 61, pp. 84-95.

Literature

http://nielsvanoort.weblog.tudelft.nl/

Brands, T., N. van Oort, M. Yap (2018),
Automatic bottleneck detection using AVL data: a case study in Amsterdam, Conference on Advanced Systems in Public Transport and TransitData (CASPT), Brisbane, Australia.

De Regt K., Cats O., van Oort N. and van Lint H. (2017). **Investigating Potential Transit Ridership by Fusing Smartcard Data and GSM Data**. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2652

Dixit, M., T. Brands, N. van Oort, O. Cats, S.P. Hoogendoorn (2019), **Passenger Travel Time Reliability For Multi-Modal Public Transport Journeys**, 98th Transportation Research Board Annual Meeting, Washington DC.

Heydenrijk-Ottens, L. et al. (2018), Supervised learning: **Predicting passenger load in public transport**, CASPT conference, Brisbane

Luo, D., Bonnetain, L., Cats, O. & van Lint, H. (2018) **Constructing spatiotemporal load profiles of transit vehicles with multiple data sources**. *Transportation Research Record*.

Van Oort, N., T. Brands, E. de Romph (2015), **Short-Term Prediction of Ridership on Public Transport with Smart Card Data**, *Transportation Research Record,* No. 2535, pp. 105-111.

Van Oort, N., D. Sparing, T. Brands, R.M.P. Goverde (2015), **Data driven improvements in public transport: the Dutch example**, *Public Transport*, Vol 7(3), pp.369-389

Questions / Contact

N.vanOort@TUDelft.nl

@Niels_van_Oort

Publications:
http://nielsvanoort.weblog.tudelft.nl/

www.smartPTlab.tudelft.nl

TUDelft
PPTS COURSE
Planning Public Transport Services
TU Delft, The Netherlands

Amsterdam **May 2019**

http://ppts-course.org/