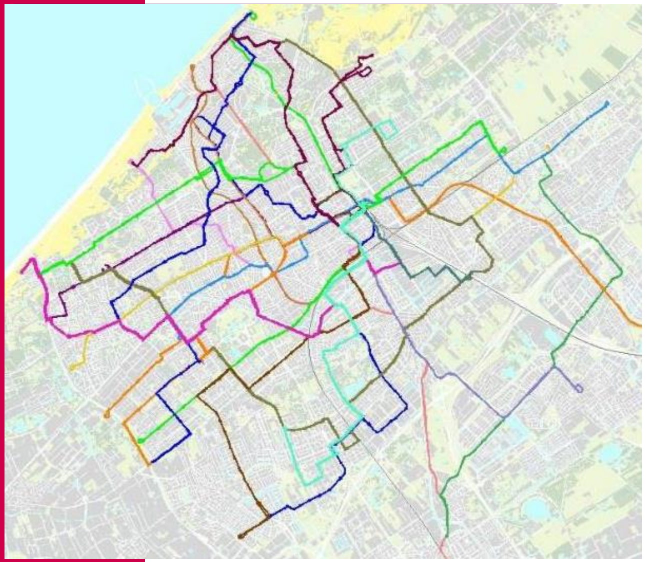
Data-driven PT ridership prediction approach

including comfort aspects



dr. N. van Oort
ir. T. Brands
ir. M.P. Drost
ir. M. Yap

Delft University of Technology

HTM

adviseurs mobiliteit
Goudappel Coffeng

Policy questions

- Impact of construction works (rerouting, ridership decrease)

- Simple efficiency improvements (schedule, fares)

- Dealing with budget savings (least damage)

Supporting decision making taking into account:

- Passenger impacts
- Costs (service) and revenues (tickets)
- Societal costs/benefits (value of time)

Available tools

	Multimodal model	Quick-Scan model
Modes	Car, public transport, bike	Public transport
Scale	National, regional, urban	Urban
Time horizon	10-20 years	< 5 years
Project type	Strategic, policies, infrastructure changes	Tactical, changing lines, frequencies
Usage	Modal split, cost-benefit analysis	Route choice effects

New generation of models

Traditional (4-step) model	Simple calculation
Multimodal (~PT)	PT only
Network	Line
Complex	Transparent
Long calculation time	Short calculation time
Visualisation	Only numbers
Much data	Little data
Detailed results	Assessments

Short term predictions

Elasticity method based on smartcard data

Smartcard data

Our research focus:

Connecting to transport model
- **Evaluating history**
- **Predicting the future**

- **Whatif scenario's**
  - Stops: removing
  - Shorter travel times and higher frequencies
  - Route changes

- **Quick insights into**
  - Expected Ridership
  - Expected costs (coverage)

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 accessibility via operators**

Data

- **19 million smartcards; 42 million transactions every week**

Connecting data to transport model

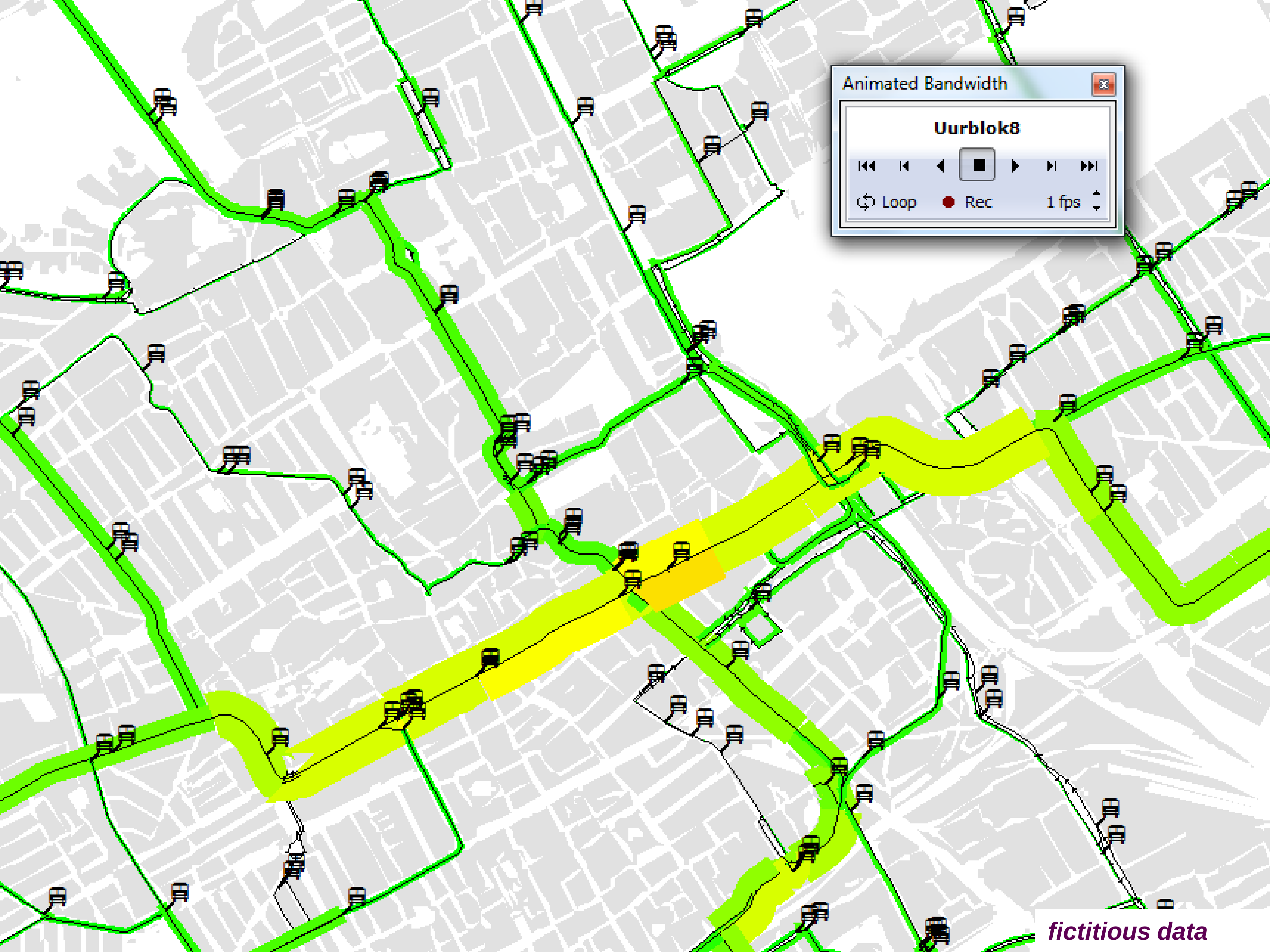
1) Importing PT networks (GTFS) (Open data)

2) Importing smartcard data (Closed data)

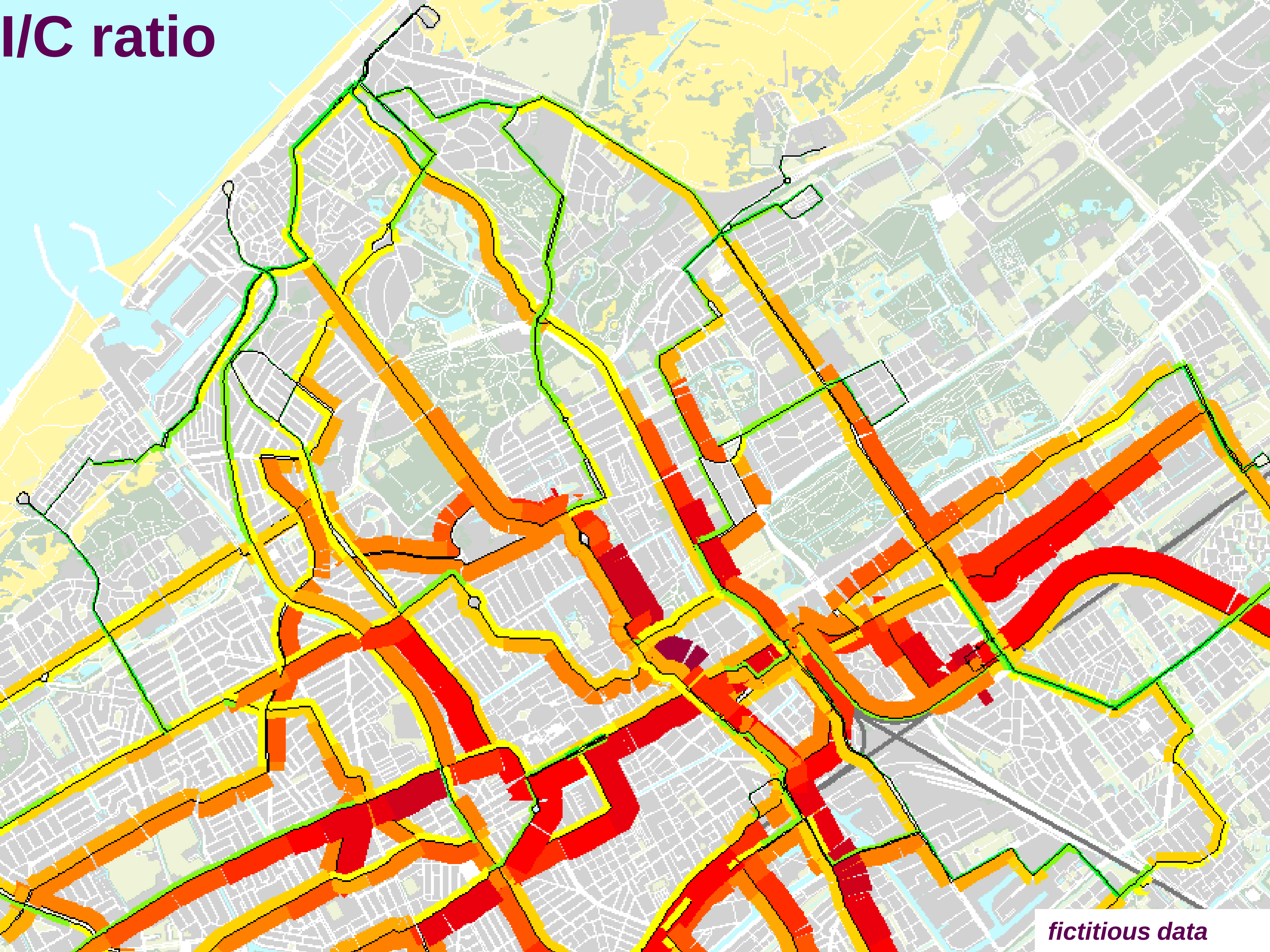
3) Processing, cleaning and matching

4) Route choice and visualization options of transport model

What if?


Animated Bandwidth
Uurblok8
Loop
Rec
1 fps
fictitious data

I/C ratio



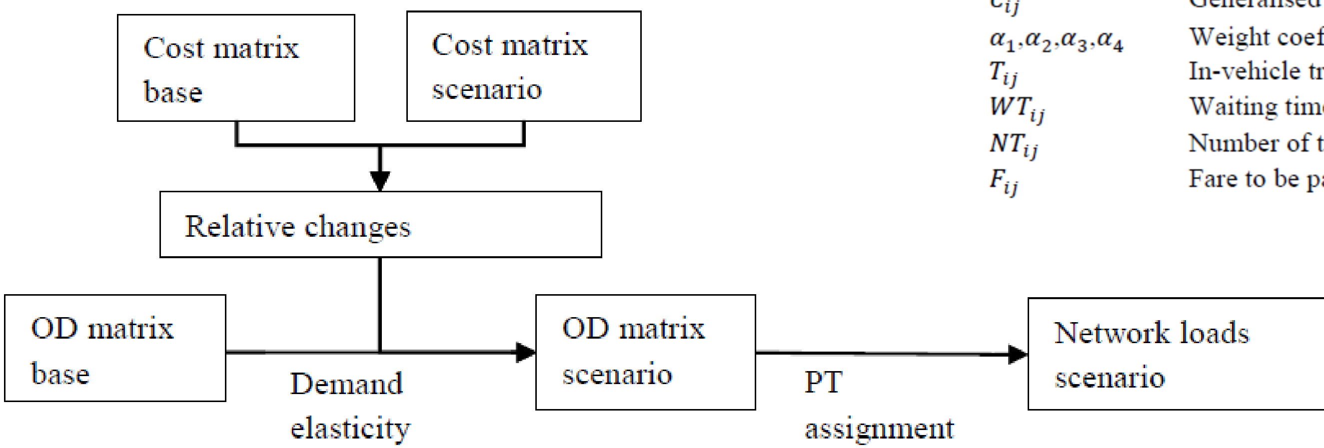
fictitious data

What if: elasticity approach

$$C_{ij} = \alpha_1 T_{ij} + \alpha_2 WT_{ij} + \alpha_3 NT_{ij} + \alpha_4 F_{ij}$$

With:

C_{ij}	Generalised costs on OD pair *i,j*
$\alpha_1, \alpha_2, \alpha_3, \alpha_4$	Weight coefficients
T_{ij}	In-vehicle travel time on OD pair *i,j*
WT_{ij}	Waiting time on OD pair *i,j*
NT_{ij}	Number of transfers on OD pair *i,j*
F_{ij}	Fare to be paid by the traveler on OD pair *i,j*



NOTE:

Simple changes

Short term

Primarly LOS changes

Accuracy

Elasticities

Literature (e.g. Balcombe)

"Proven " rules of thumb

Crowding in PT

- **Perception of in-vehicle time of travellers: a crowded vehicle is less attractive**
  - Travel time may remain the same
- **Dwell time may increase in a crowded vehicle**
  - Boarding and alighting of passengers takes more time
- **Very crowded vehicles result in denied boarding**
  - Additional waiting time of one (or more) entire headway

Crowding model

- **(perceived) in-vehicle time depends on crowding level**
- **Iterative assignment is needed**
- **Two values indicate capacity:**
  - Number of seats
  - Crush capacity: maximum capacity of vehicle: sitting and standing passengers together

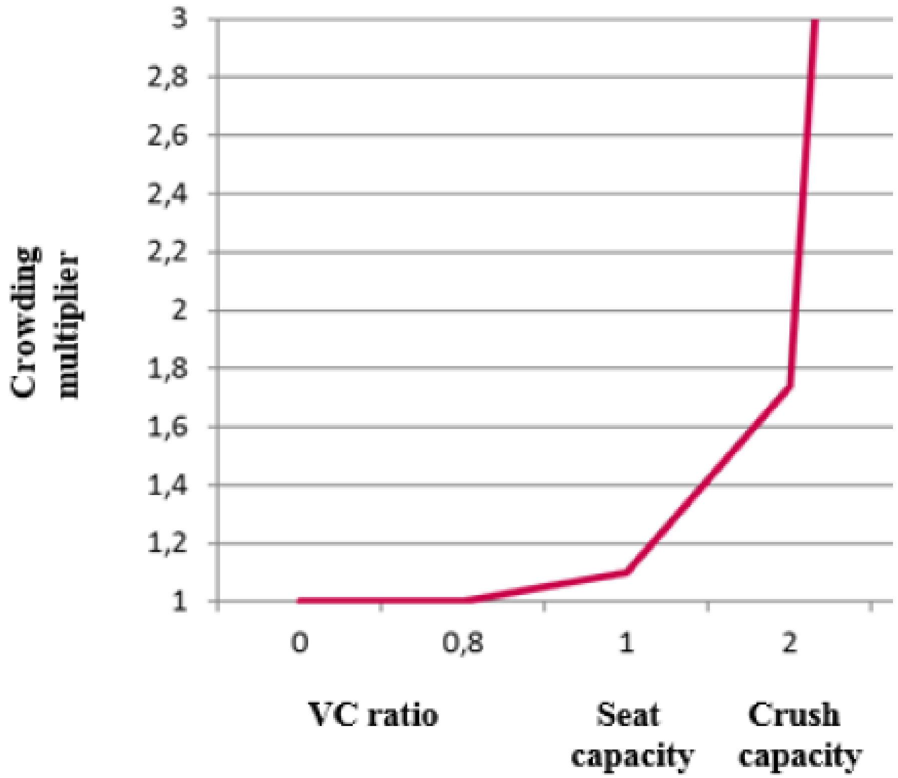
$$VC = \left\{ \frac{\frac{L}{C_{seats}}}{1+\frac{L-C_{seats}}{C_{crush}-C_{seats}}} \right.$$

- **Distinguish between vehicles with relatively large / small number of seats**

Crowding model

$$T_{ij}^{per} = T_{ij} * F$$

$$C_{ij} = \alpha_1 T_{ij}^{per} + \alpha_2 WT_{ij} + \alpha_3 NT_{ij} + \alpha_4 F_{ij}$$

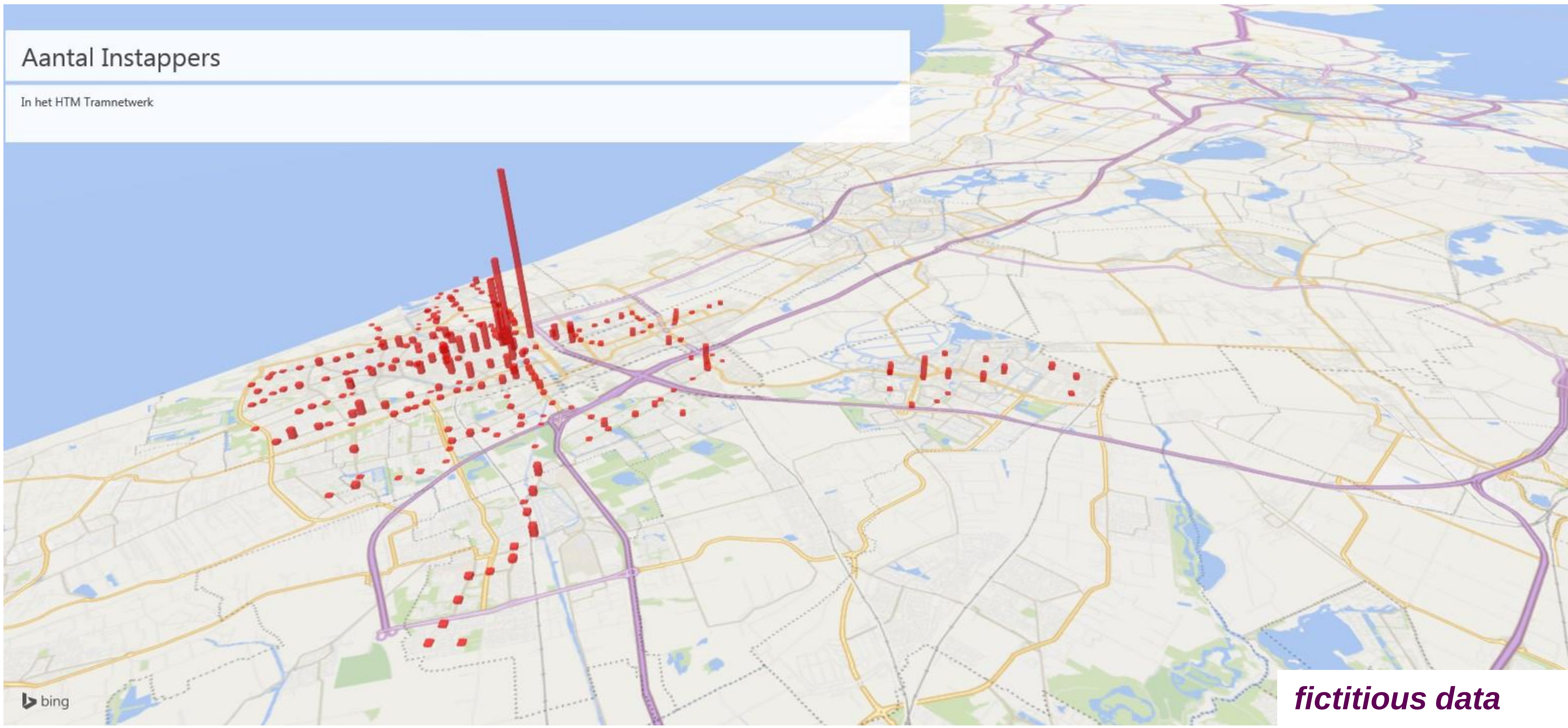


Douglas Economics (2006)
MVA Consultancy (2008)
Wardman and Whelan (2011)

Crowding model

- **Modelled time period**
  - Usually an entire peak period of 2 hours is modelled
  - Some vehicles may be busier than other vehicles
  - Evenness of the load distribution over this period
  - → a correction factor may be applied that is lower than 2, to incorporate this effect

Case study: The Hague tram network



fictitious data

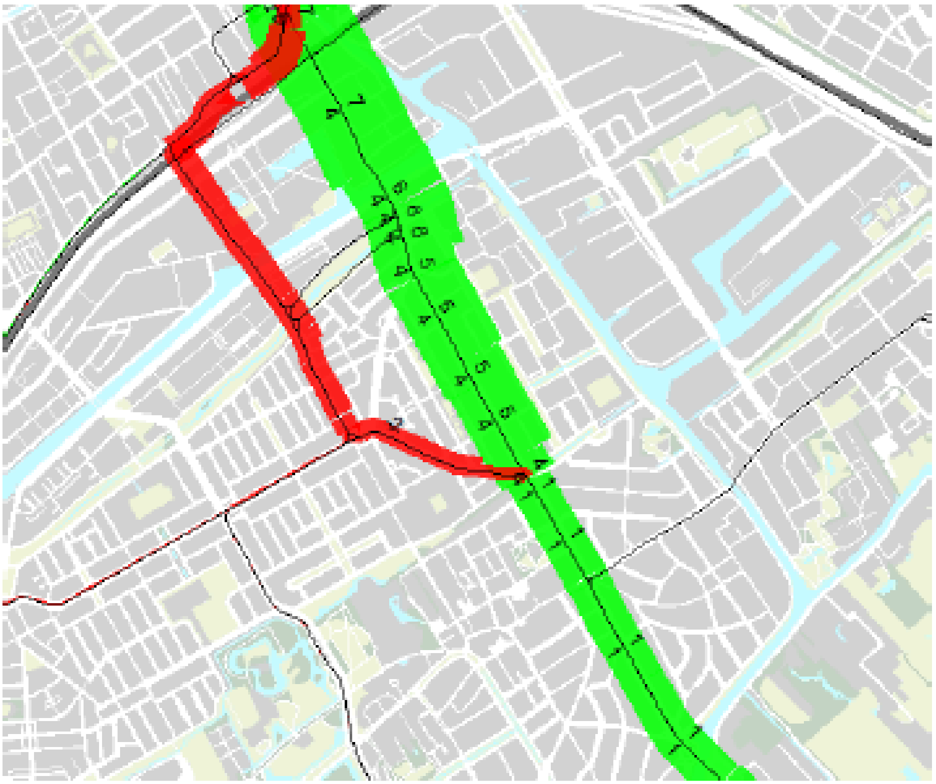
Network changes

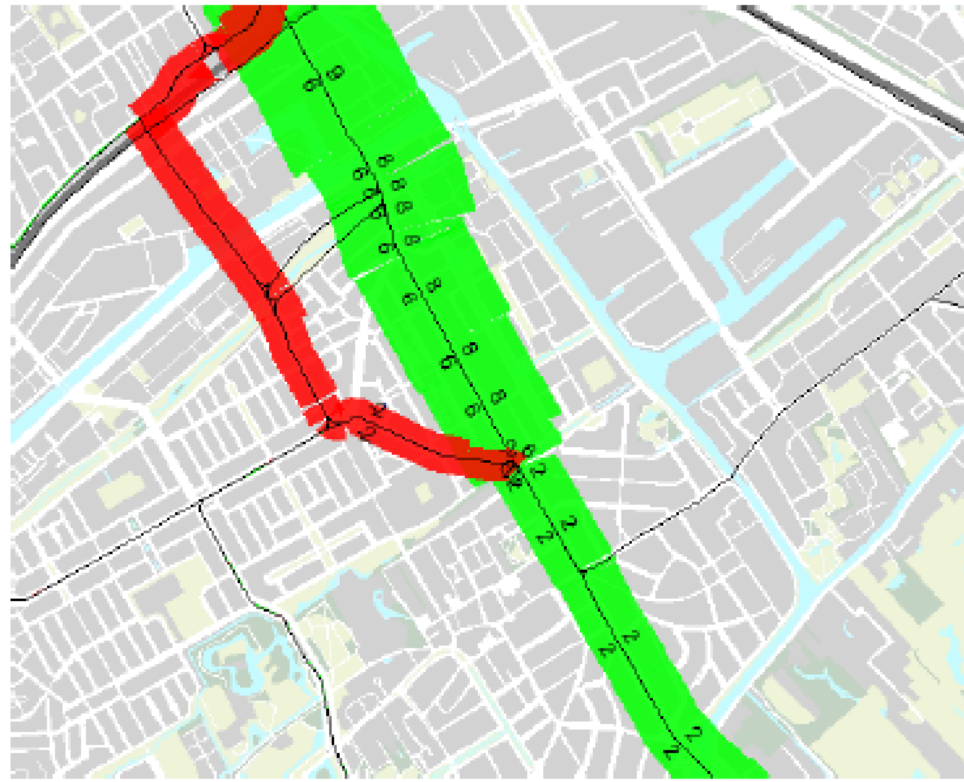
- **Increase frequency of tram line 15 from 6 to 8 times per hour during moring peak and evening peak**

- **Results:**

	Model without comfort	Model including comfort
Average work day	+8%	+10%

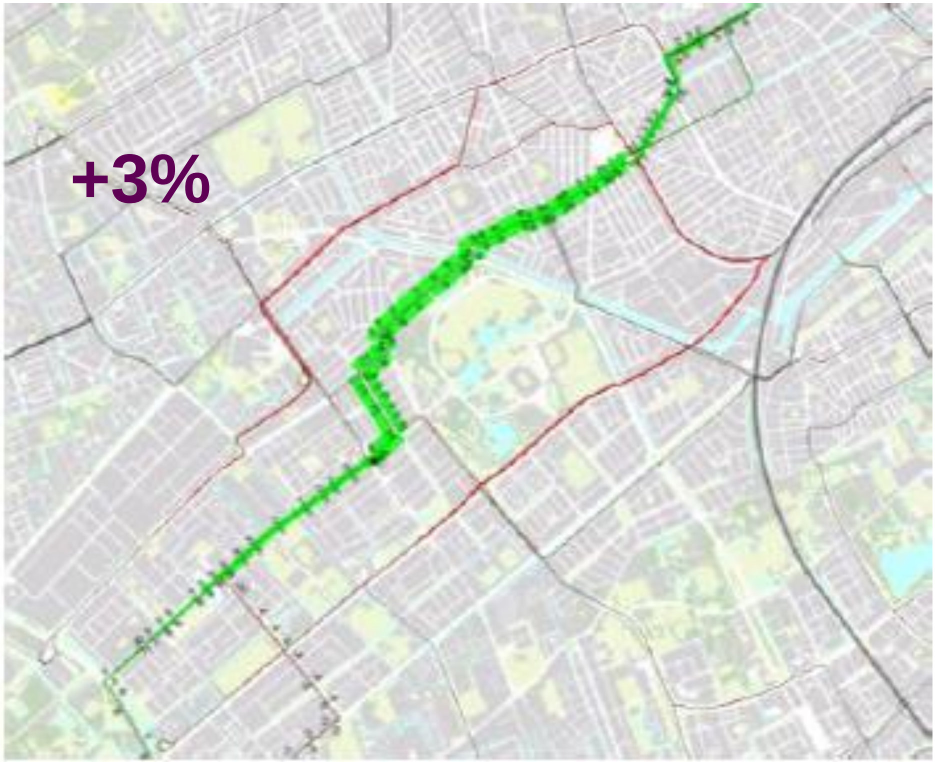
Result frequency increase line 15





Network changes

- **Transformation of line 25 from bus line to tram line**
- **Due to larger vehicles, the frequency decreases**



Conclusions

- Smartcard data supports ridership predictions
- Combining strengths of both worlds
- Comfort is explicitely taken into account
- Limited computation time needed for real size networks

- Benefits (revenues and societal) of certain measures become larger when comfort is taken into account
- Up to 30% underestimation of the effects when comfort is neglected
- Reduce crowding may compensate frequency reduction

Work in progress

- Validating the model using revealed data (smart card data), including behaviour during disturbances
- Incorporating denied boarding and extended dwell times
- Adding service unreliability costs
- Applying the quantified comfort effect in cost-benefit analysis

Questions

Related papers:

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

- Niels van Oort
- NvOort@goudappel.nl / N.vanOort@TUDelft.nl

- Ties Brands: TBrands@goudappel.nl

http://www.utwente.nl/ctw/vvr/People/brands/

- Marc Drost: M.Drost@htm.net

- Menno Yap: MYap@goudappel.nl