Transportation, open data and artificial intelligence - Challenges and opportunities

Univ.-Prof. Dr. Constantinos Antoniou

c.antoniou@tum.de

https://www.mos.ed.tum.de/en/vvs/

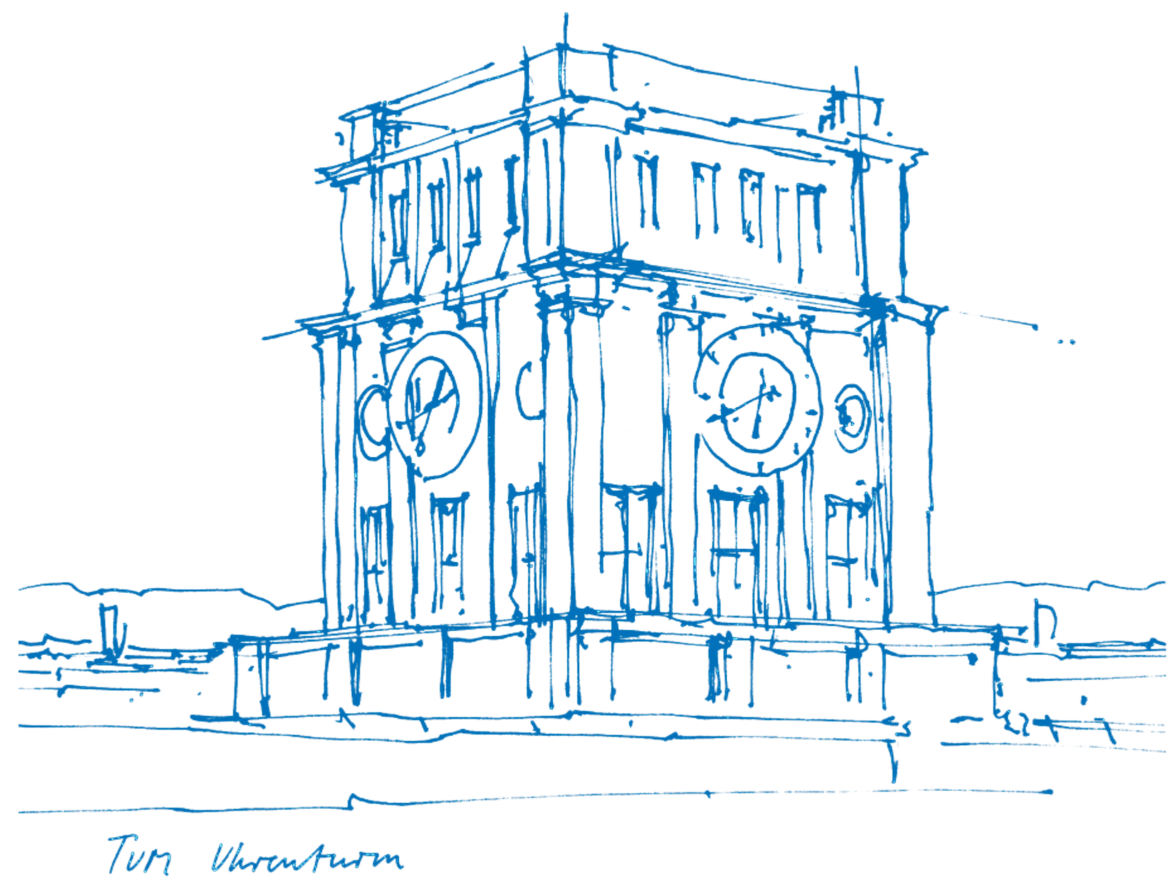
May 22., 2023

Technical University of Munich

TUM School of Engineering and Design

Department of Mobility Systems Engineering

Chair of Transportation Systems Engineering



Background - Constantinos (Costas) Antoniou

Diploma of Civil Engineering (1995)

National Technical University of Athens, Greece

M. S. in Transportation (1997)

Massachusetts Institute of Technology

Ph. D. in Transportation Systems (2004)

Massachusetts Institute of Technology

Associate Professor (until 2014 Assistant Professor) (2009-2015)

National Technical University of Athens, Greece

Chair of Transportation Systems Engineering (TSE)

TSE Thematic Research Profile

Modeling, optimization and simulation of transportation systems

- Demand and supply modeling (multimodal, incl. freight)
- Emerging transport modes (autonomous, urban air mobility, etc.)
- Optimization, calibration, and validation

Data science and data analytics

- Big data acquisition, e.g. via social media
- Data-driven models and machine learning

Human factors analysis

- Road safety modeling
- Behavioral economics applications
- Modeling of factors affecting transportation systems user engagement (adoption, satisfaction, etc.)

Outline

Thanks to many funding agencies, including:

DFG Deutsche Forschungsgemeinschaft

Transport data openness

Opportunistic applications

Scalable processing of drone videography data

Indirect traffic flow estimation from link speeds

TUM TSE at the NeurIPS 2022 Traffic4cast challenge

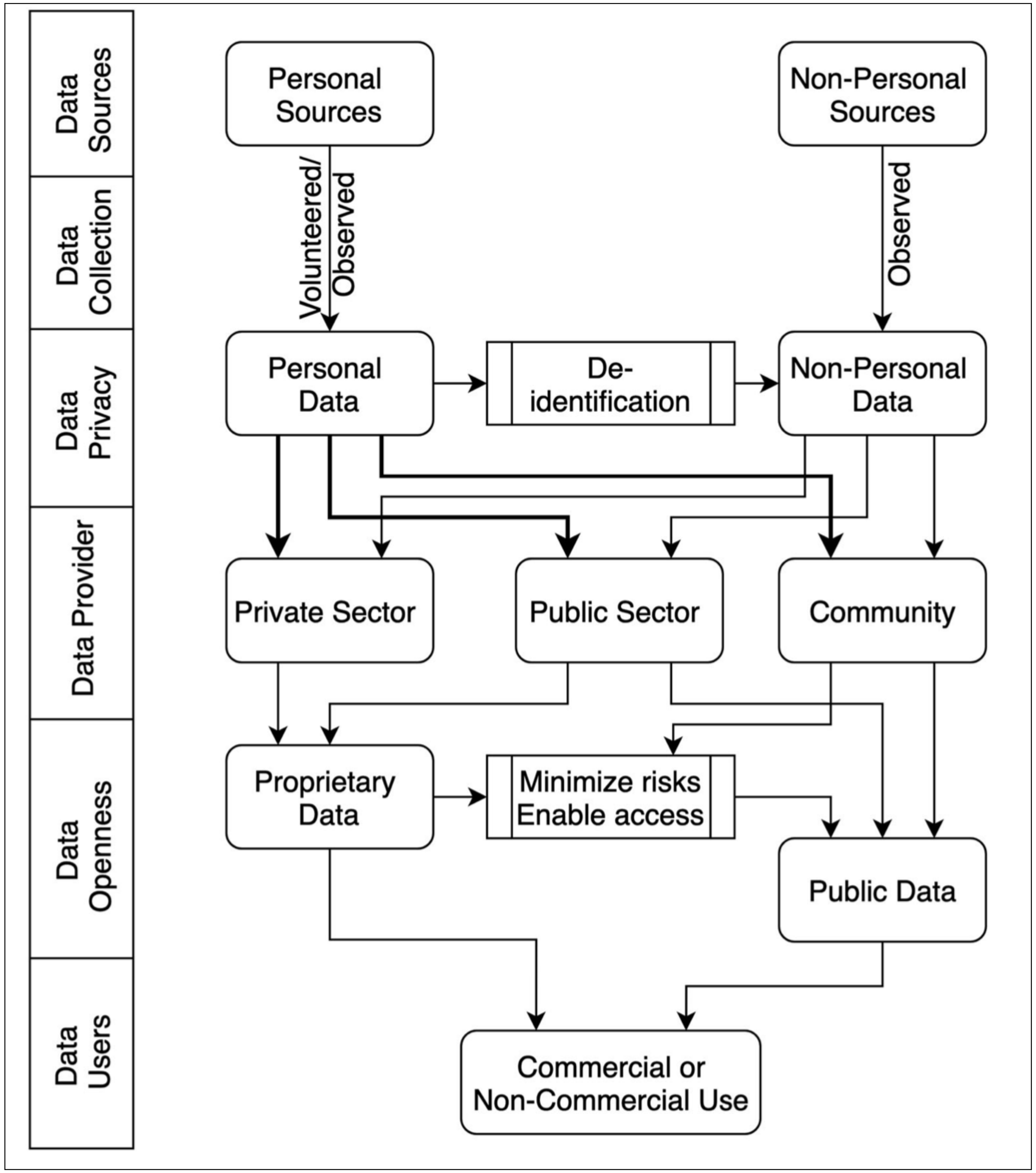
Some ongoing projects

Outlook and conclusion

Transport data openness

Transportation data

Production and operational flow of the data



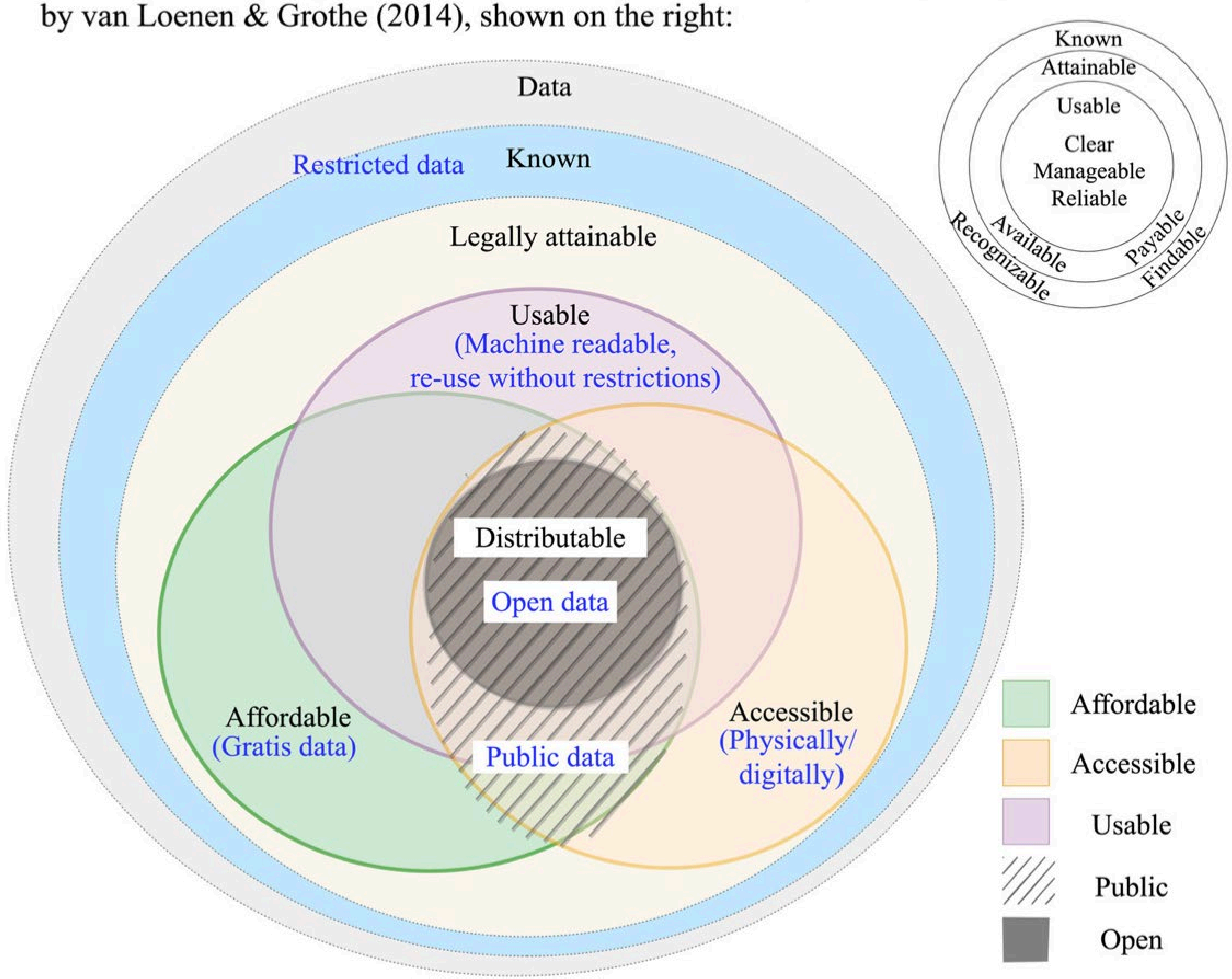
Source: Mahajan, V., N. Kuehnel, A. Intzevidou, G. Cantelmo, R. Moeckel & C. Antoniou (2021)
Data to the people: a review of public and proprietary data for transport models, Transport Reviews, 42 (4), 415-440. DOI: 10.1080/01441647.2021.1977414

Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

Data openness

Are data publicly available?

Visualisation was inspired by the concentric shell model by Backx (2003), further used by van Loenen & Grothe (2014), shown on the right:



Source: Mahajan, V., N. Kuehnel, A. Intzevidou, G. Cantelmo, R. Moeckel & C. Antoniou (2021)
Data to the people: a review of public and proprietary data for transport models, Transport Reviews, 42 (4), 415-440. DOI: 10.1080/01441647.2021.1977414

Data insufficiency

Commendable open data initiative

Needs vs. availability

Data availability varies

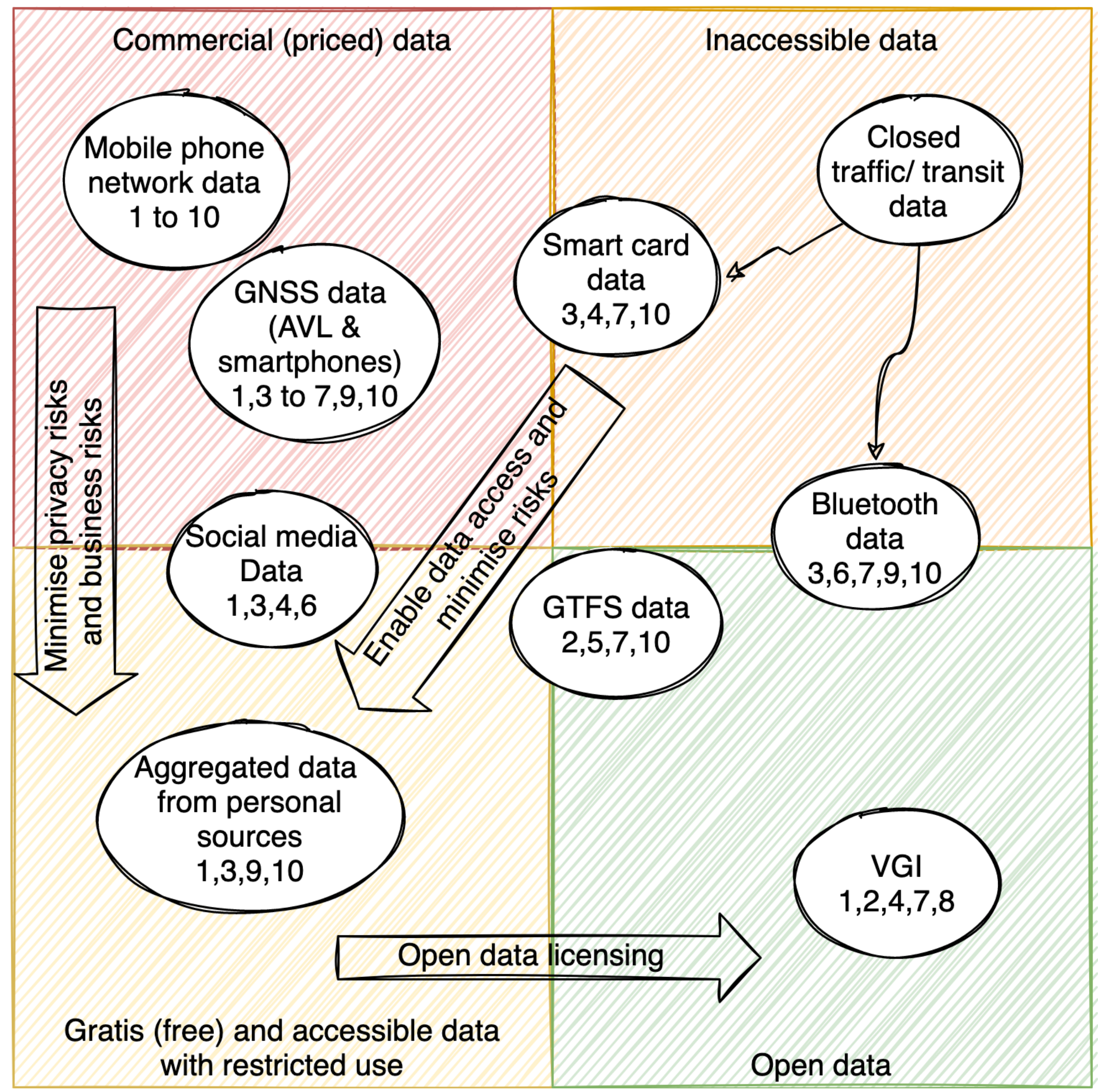
Key (Numbers in the circles represent modelling task):
1: Trip Generation, **2**: Accessibility, **3**: Trip Distribution/ OD flows,
4: Destination Choice/ Activity Spaces/ Trip Purpose,
5: Departure Time Choice, **6**: Mode Detection, **7**: Route Choice,
8: Network Modelling, **9**: Traffic Speed, **10**: Travel-time,

Crowd-sensed GNSS (POI/ Map data by Google, Foursquare, etc.)

AVL data (Traffic data by TomTom, INRIX, etc)

MPND (OD Matrix by telecom operators)

Social media data (Premium Twitter API)



Closed Transit data (if transit operators do not adopt GTFS and share the data publicly)

Closed Traffic data (if traffic operators do not share the detector data)

Aggregated AVL data (UBER Movement aggregated Travel-time and Speed data)

Social media data (Standard/ Free Twitter API)

Bluetooth travel-time (Open Data Toronto)

VGI (OpenStreetMaps)

Static GTFS data for public transport in Germany (www.gtfs.de)

Traffic counts (Open Data Paris)

Source: Mahajan, V., N. Kuehnel, A. Intzevidou, G. Cantelmo, R. Moeckel & C. Antoniou (2021)
Data to the people: a review of public and proprietary data for transport models, Transport Reviews, 42 (4), 415-440. DOI: 10.1080/01441647.2021.1977414

Opportunistic applications

Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

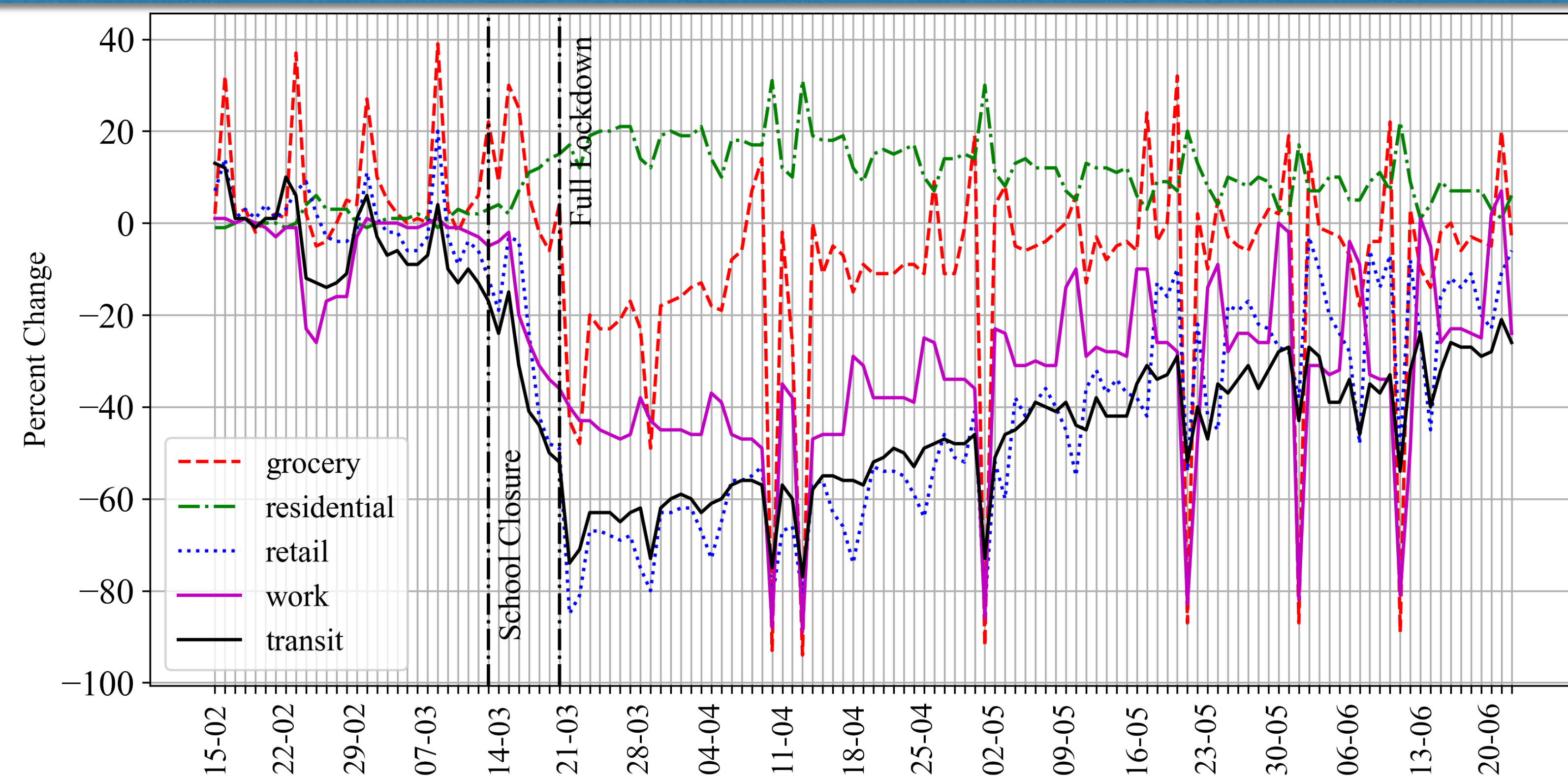
COVID-19

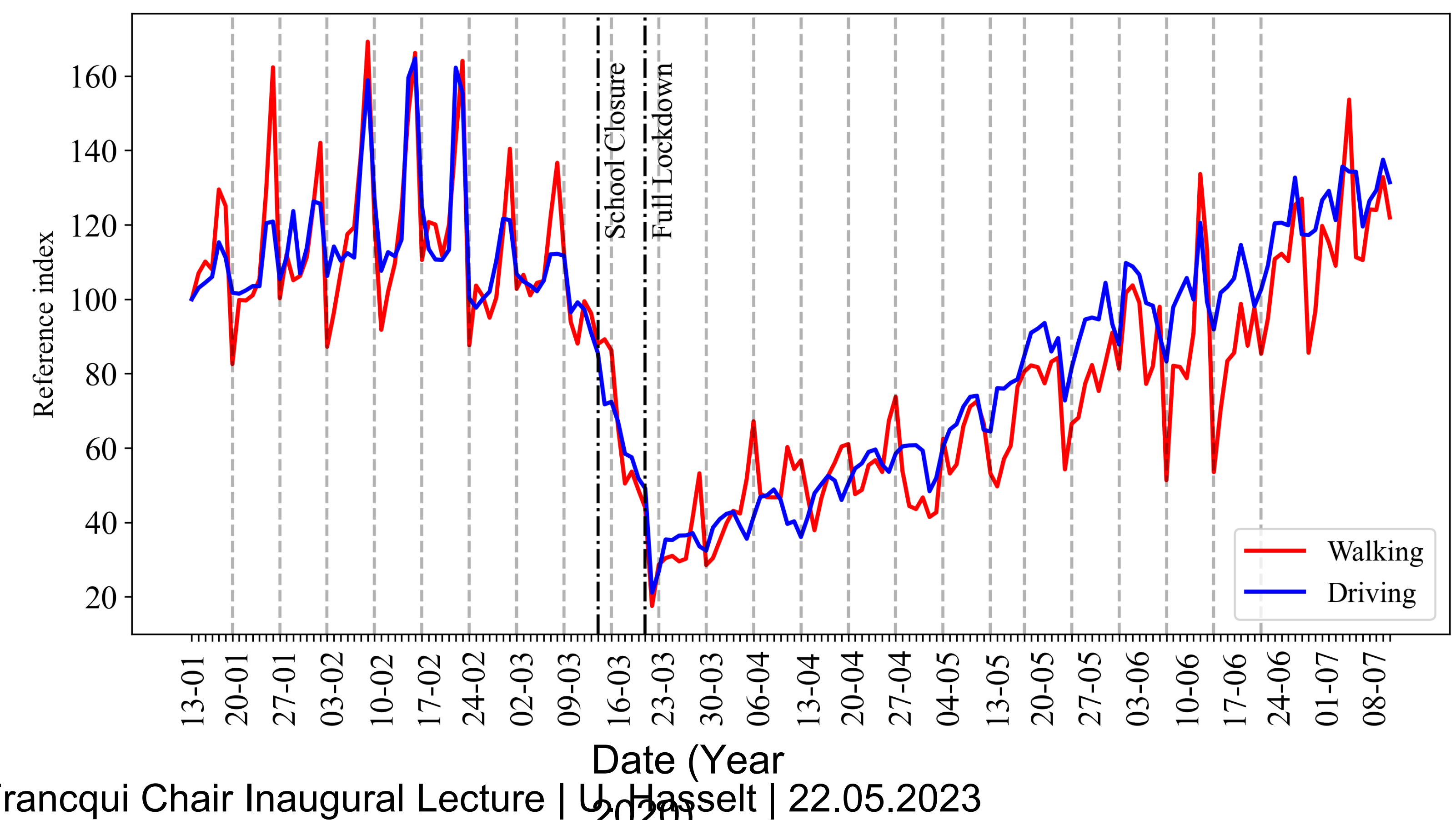
Triggers:
- Lockdown/ mobility restrictions
- Social distancing
- Work from home
- ….

Observations:
- Panic buying
- Drop in transit ridership
- Dip in non-essential retail
- ….

Data source:
Google Community Mobility Reports,
Apple Mobility Trends

Movement patterns in Bavaria (top) and Munich (bottom) in 2020





Google's popular time graph

Google's popular time graph for POIs (Place of Interest)

POI's popularity graph

- Past trend (average over past few weeks)
- **Live popularity (real-time)**
- Relative busyness on [0,100] scale
- Indication of busyness/ visitation

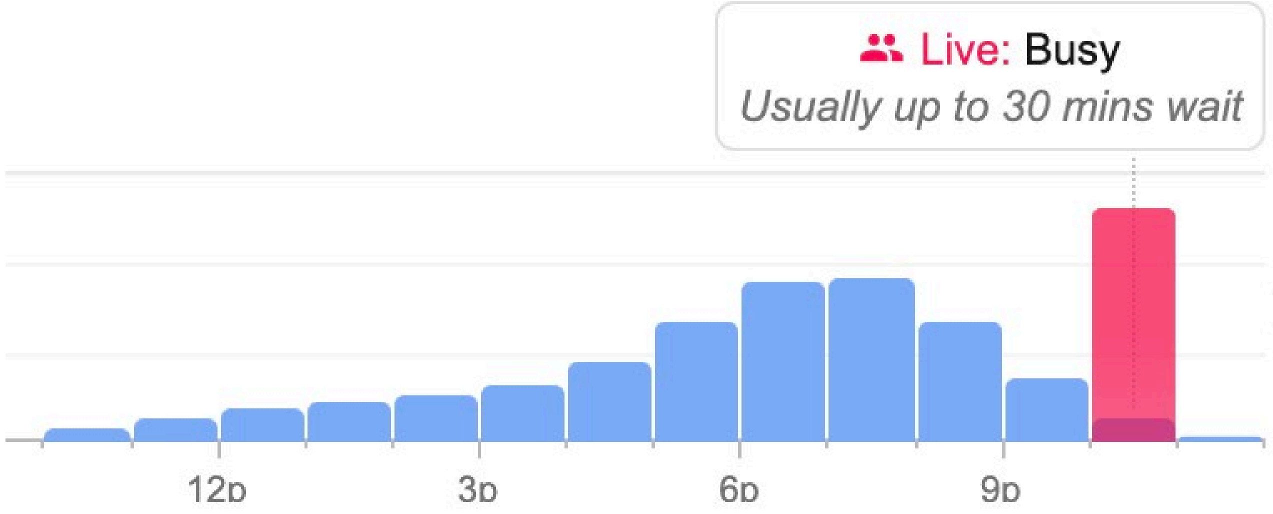


Image Source: Google Maps

https://github.com/m-wrzr/populartimes

Previous studies

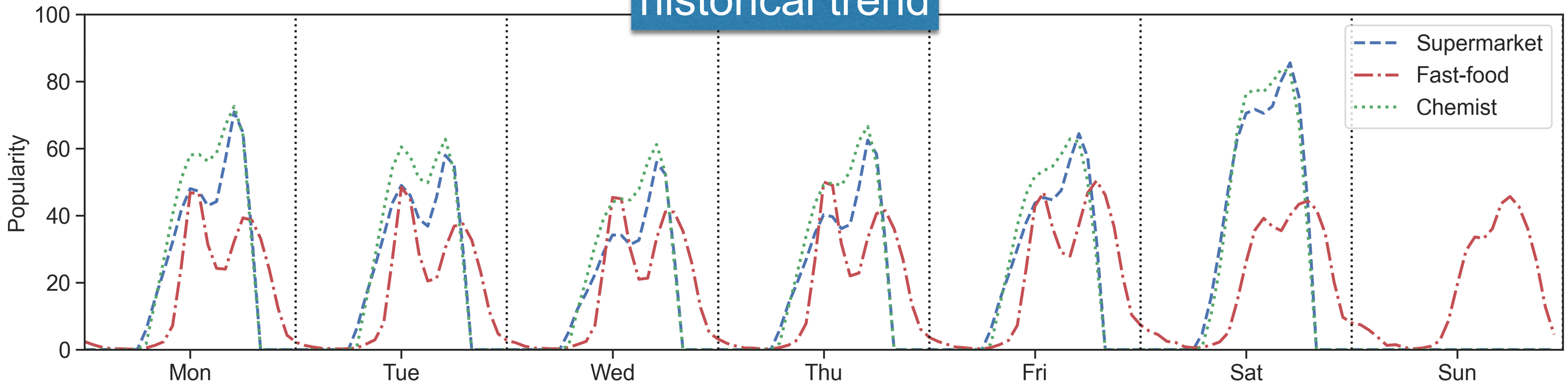
- Venue popularity[1]
- Local business attractiveness[2]
- Demand expansion factors[3]
- Consumer behavior[4]

[1]Timokhin et al., 2020, [2]Capponi et al., 2019, [3]MacKenzie & Cho, 2020, [4]Möhring et al., 2020

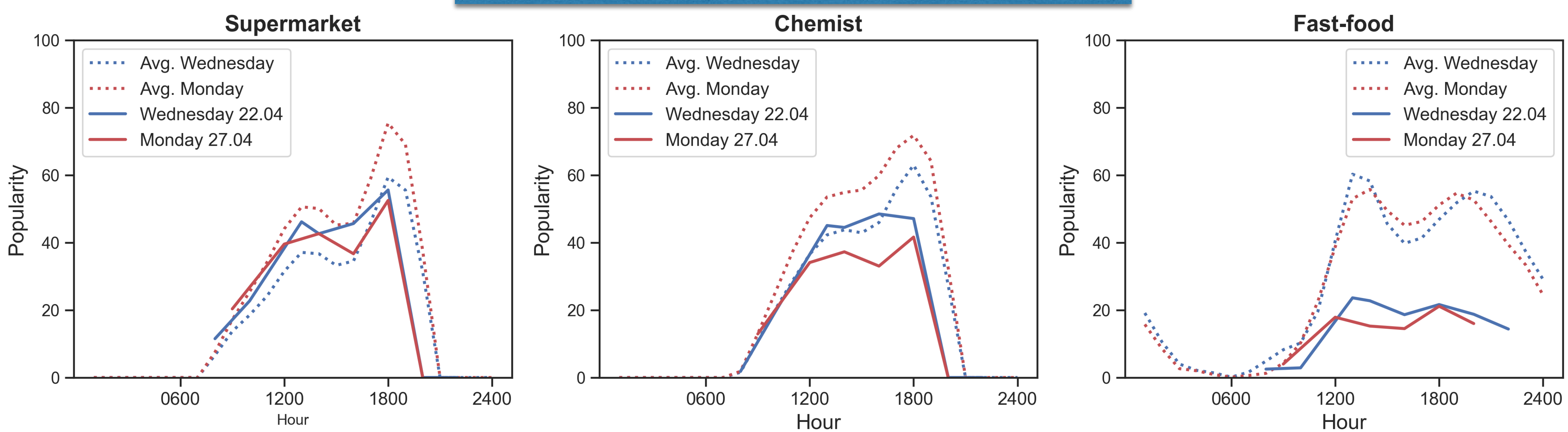
Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

Popularity trend

historical trend



historical vs. live trend (April 2020)



Source: Mahajan, V., Cantelmo, G. & Antoniou, C. (2021) Explaining demand patterns during COVID-19 using opportunistic data: a case study of the city of Munich, European Transport Research Review 13(1), 1-14

Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

Results

- Variability in the effects of lockdown
- Significant correlation of POI popularity with:
  - POI type
  - Proximity to a transit stop
  - Day of the week
- Focus on popularity dip for transit-close POIs
  - Work from home
  - Change of travel modes

Linear regression

	Dependent variable: P_{i-d}	
	1:OLS	2: RLM
Intercept	48.74***	53.37***
	(6.93)	(7.04)
fast-food	−2.56	−3.57
	(3.75)	(3.81)
lockdown	3.66	−3.50
	(9.82)	(9.98)
lockdown:fast-food	−19.47***	−16.73***
	(5.31)	(5.39)
lockdown:monday	−16.45***	−16.24***
	(1.52)	(1.55)
lockdown:average stop distance/1000	19.49**	20.19**
	(8.93)	(9.07)
lockdown:supermarket	4.53**	4.84**
	(2.02)	(2.05)
monday	11.20***	11.29***
	(1.08)	(1.09)
average stop distance/1000	−10.52*	−11.40*
	(6.32)	(6.42)
Observations	718	718
R^2	0.34	
Adjusted R^2	0.32	

Source: Mahajan, V., Cantelmo, G. & Antoniou, C. (2021) Explaining demand patterns during COVID-19 using opportunistic data: a case study of the city of Munich, European Transport Research Review 13(1), 1-14

From relative to absolute numbers

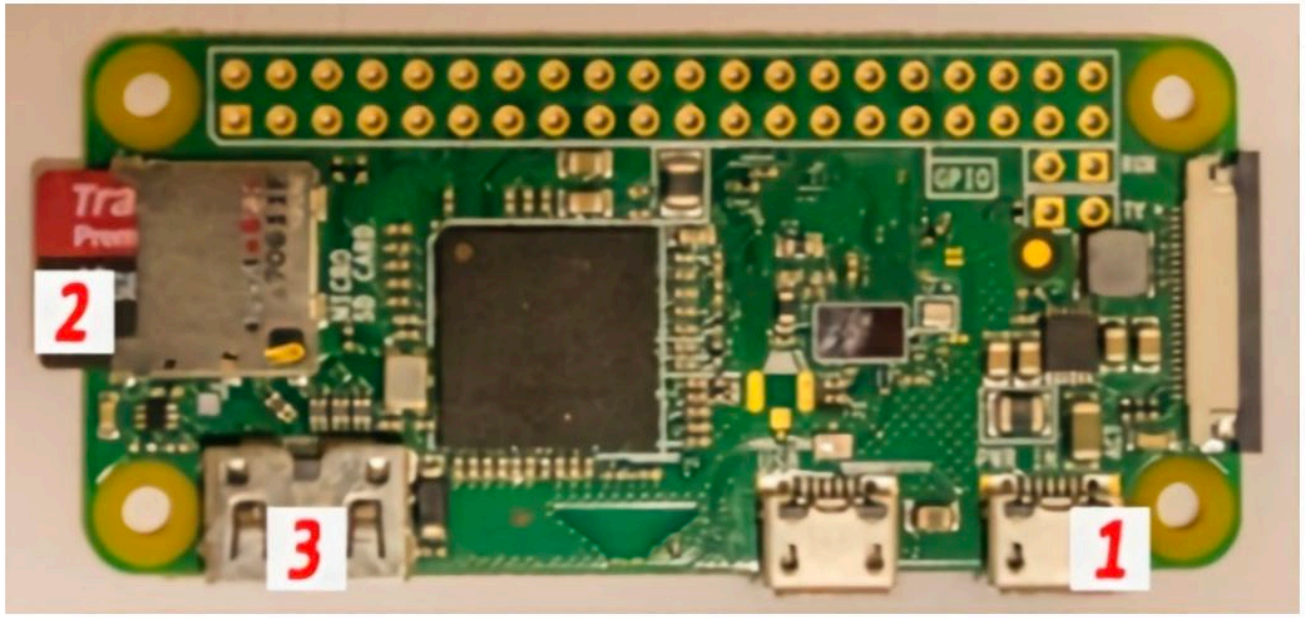


Figure 12. Photo of the microcontroller setup.

Table 6. The price of each item used in the microcontroller setup (2018, amazon.de).

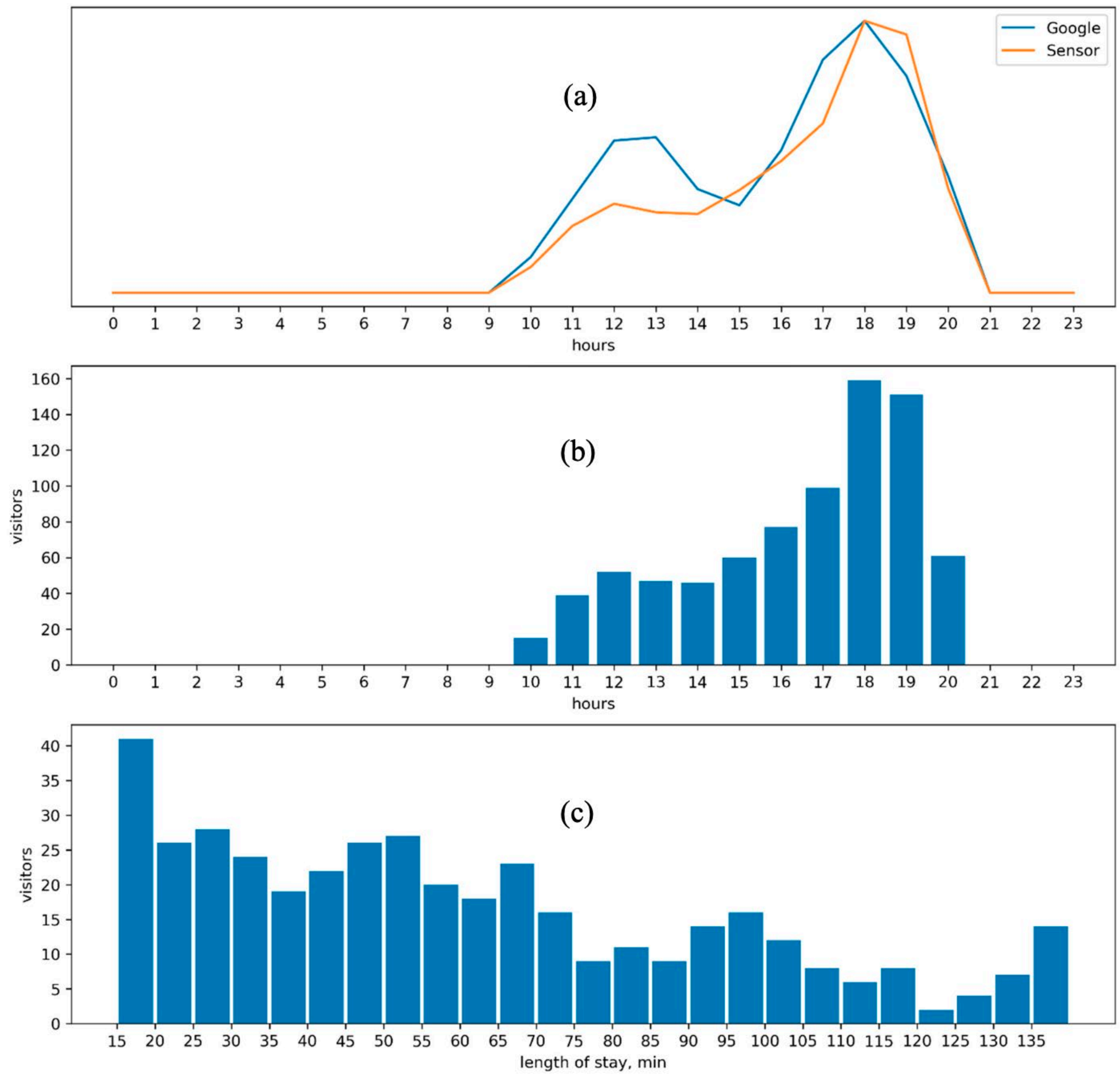
Item	Cost, EUR
Raspberry Pi Zero W	10
Micro SD card (16 GB)	6.49
Power Bank (5000 mAh)	8.99

Timokhin, S.; Sadrani, M.; Antoniou, C. (2020). Predicting Venue Popularity Using Crowd-Sourced and Passive Sensor Data. Smart Cities, 3(3), 818-841.

Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

Takumi

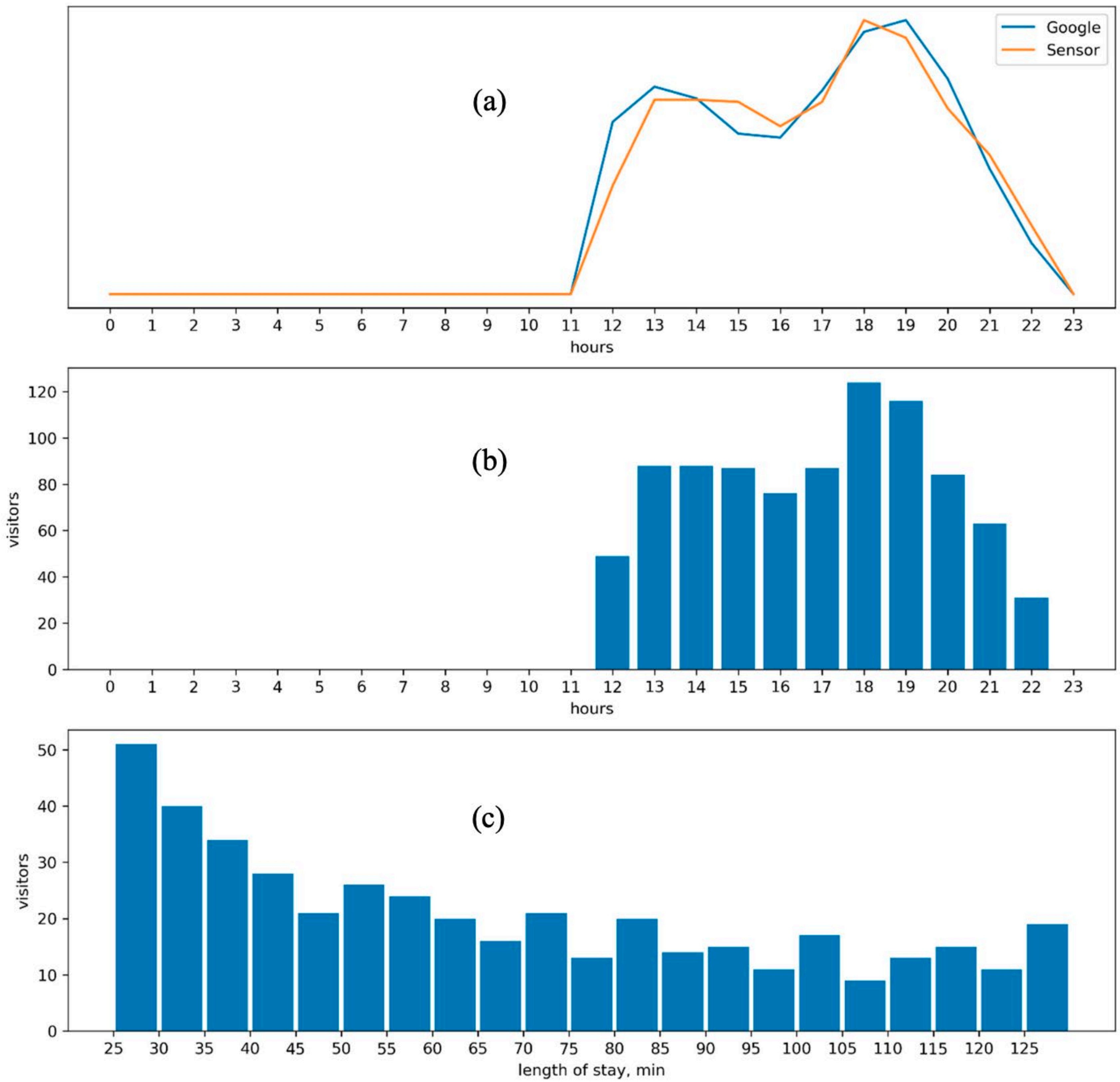
Min threshold: 15 min
Max threshold: 140 min
Correlation with Google: 0.93



Timokhin, S.; Sadrani, M.; Antoniou, C. (2020). Predicting Venue Popularity Using Crowd-Sourced and Passive Sensor Data. Smart Cities, 3(3), 818-841.

Lo Studente

Min threshold: 25 min
Max threshold: 130 min
Correlation with Google: 0.98

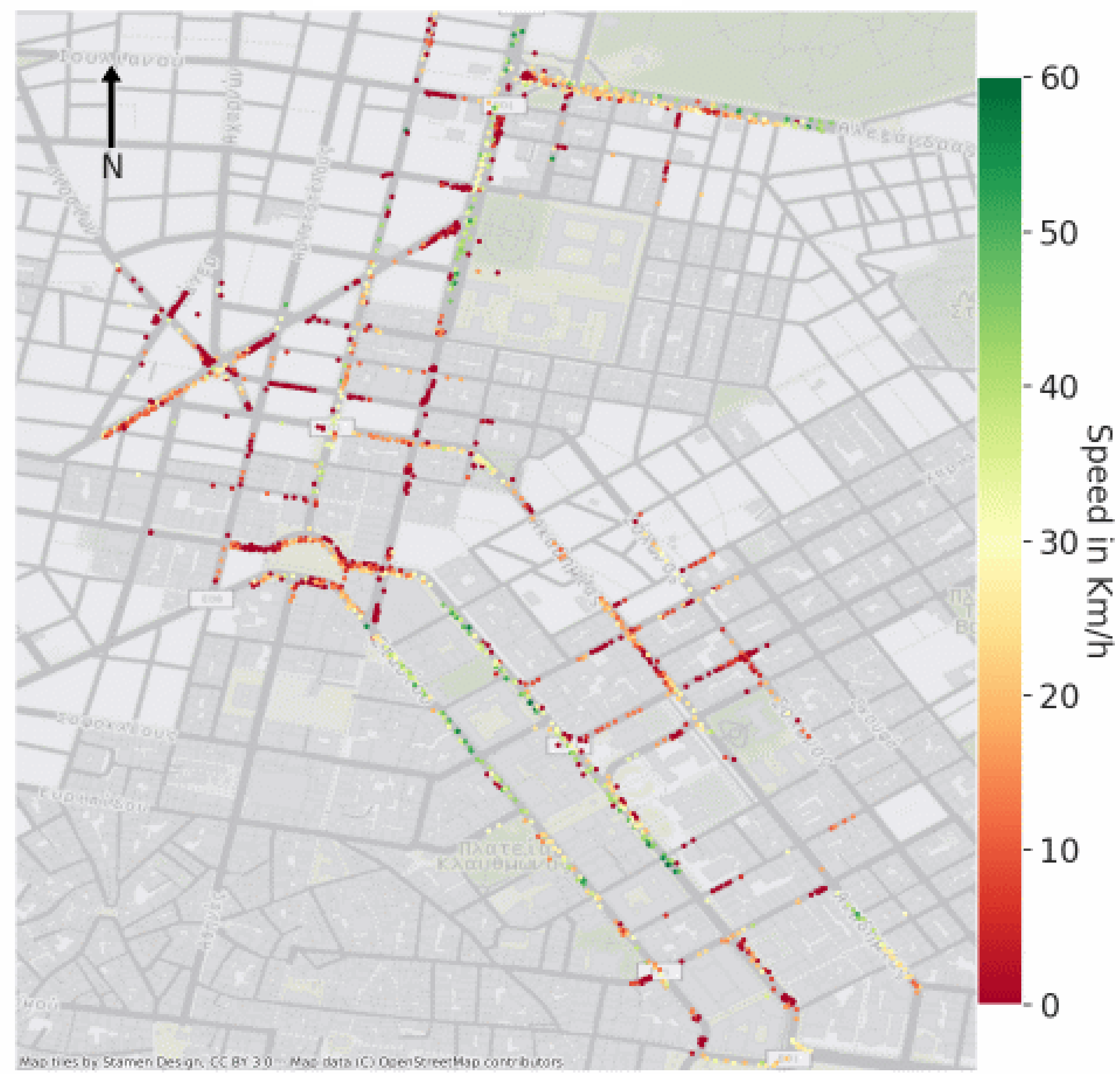


Timokhin, S.; Sadrani, M.; Antoniou, C. (2020). [Predicting Venue Popularity Using Crowd-Sourced and Passive Sensor Data](). Smart Cities, 3(3), 818-841.

Addressing quality of drone videography data

Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

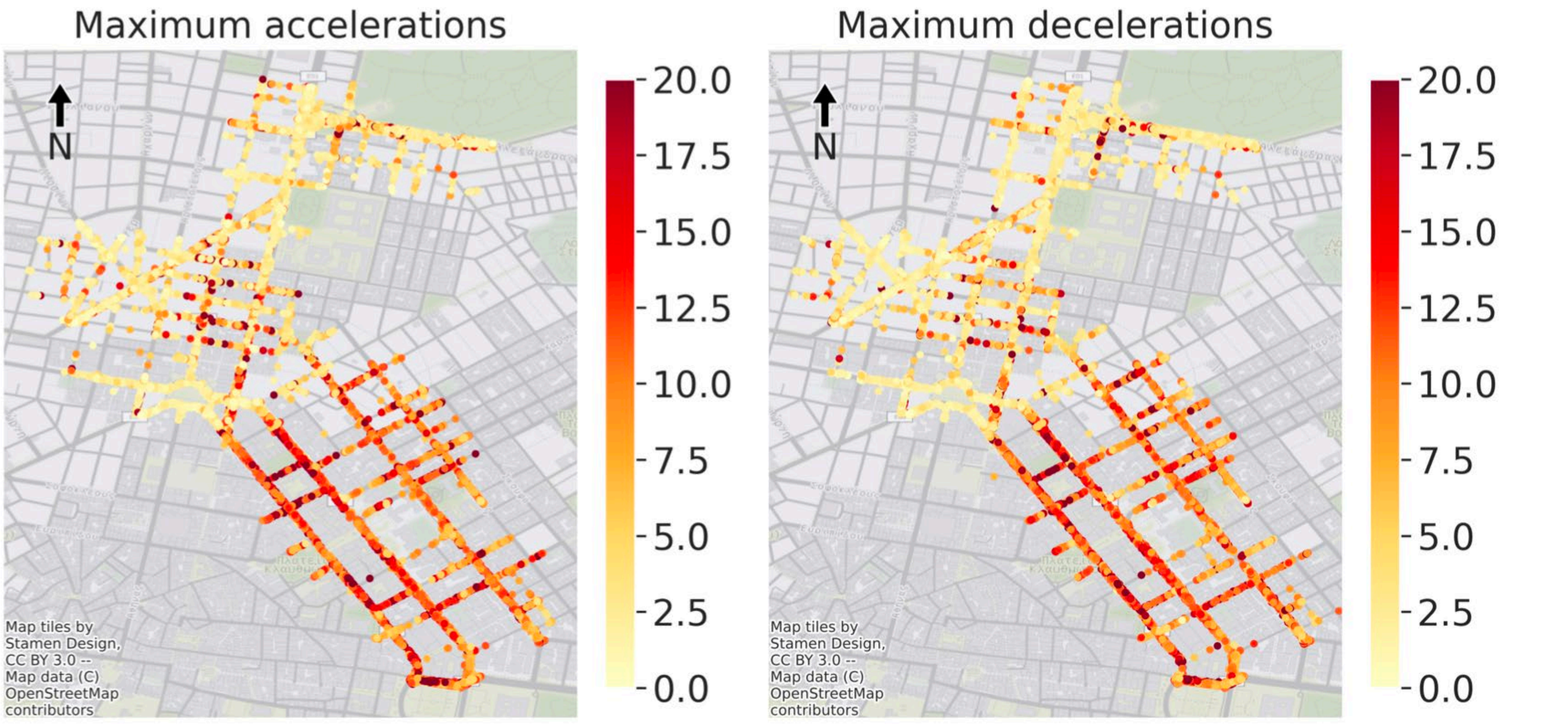
Drones for data collection



Data source: Barmpounakis, E. and Geroliminis, N., On the new era of urban traffic monitoring with massive drone data: The **pNEUMA** large-scale field experiment, Transportation Research Part C: Emerging Technologies, Volume 111, 2020.

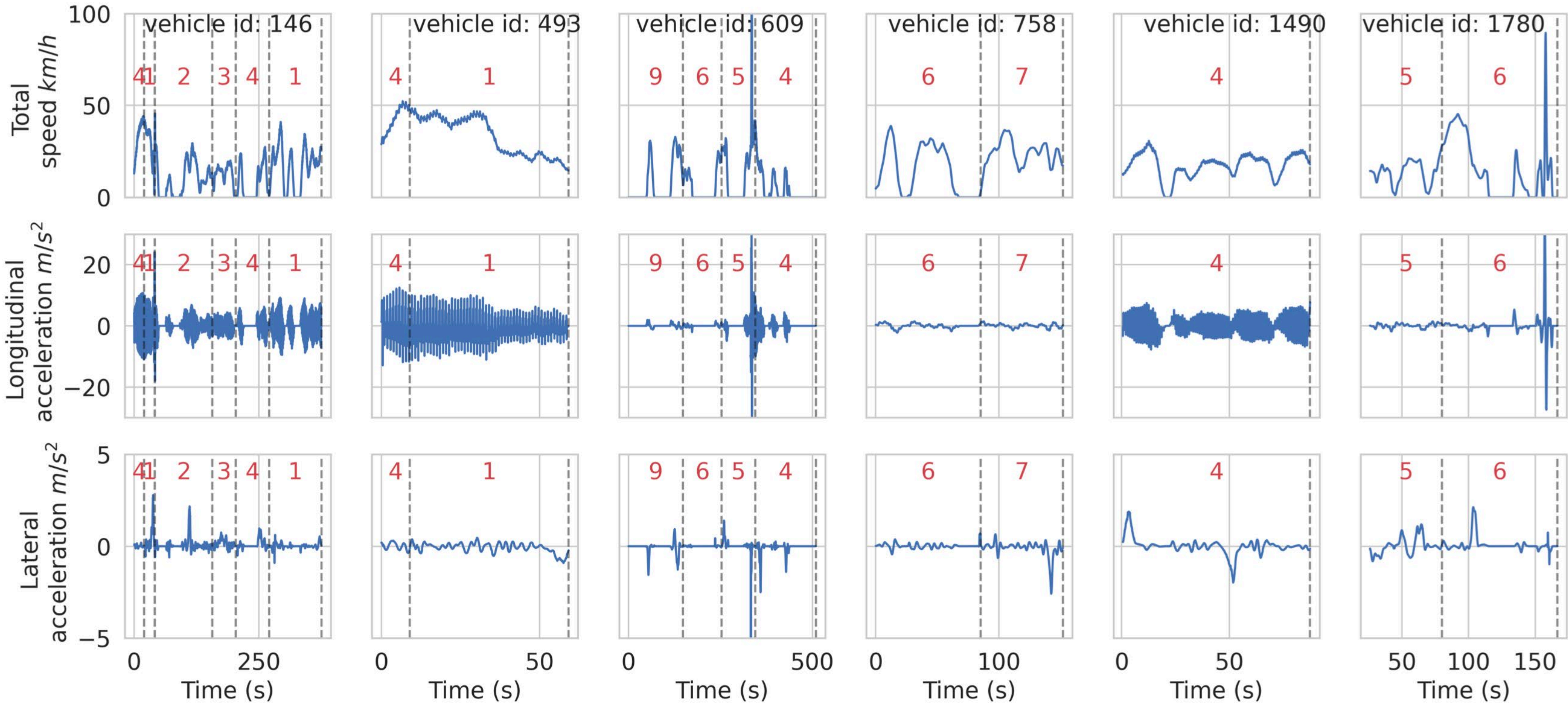
https://open-traffic.epfl.ch/

Errors: noise and anomalies



Source: V. Mahajan, E. Barmpounakis, M. R. Alam, N. Geroliminis and C. Antoniou, "Treating Noise and Anomalies in Vehicle Trajectories From an Experiment With a Swarm of Drones," in IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2023.3268712.

Errors: noise and anomalies

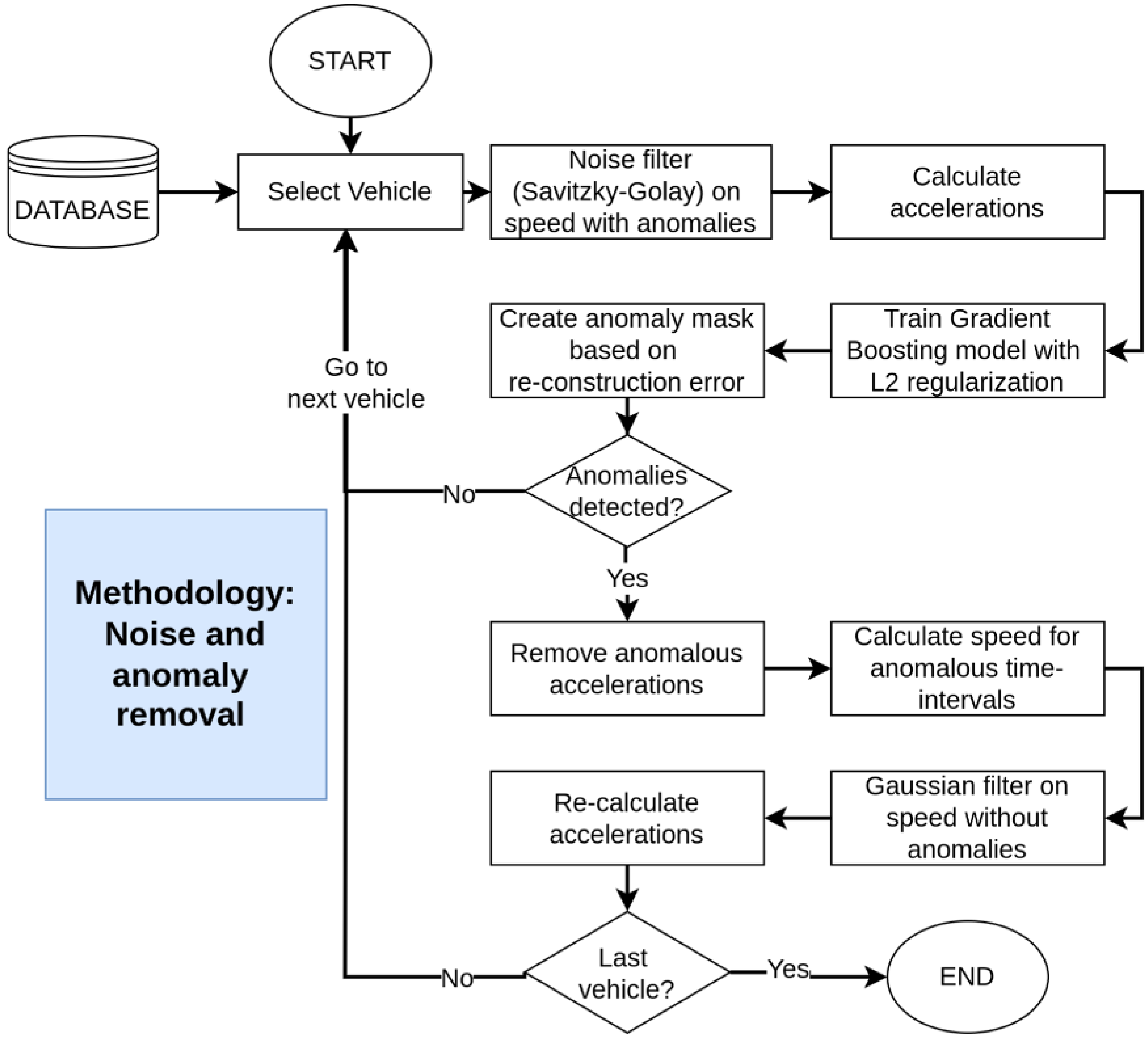


Source: V. Mahajan, E. Barmpounakis, M. R. Alam, N. Geroliminis and C. Antoniou, "Treating Noise and Anomalies in Vehicle Trajectories From an Experiment With a Swarm of Drones," in IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2023.3268712.

Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

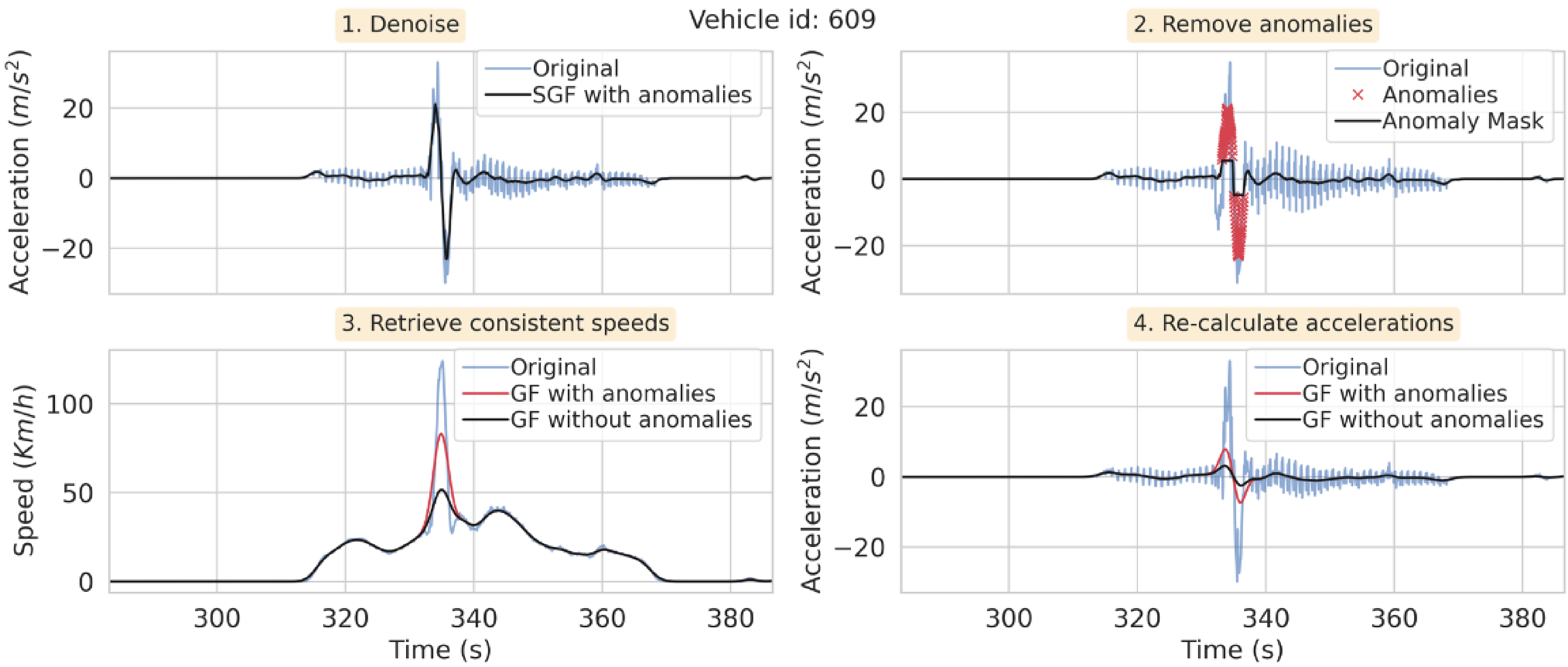
Methodology

Methodology: Noise and anomaly removal



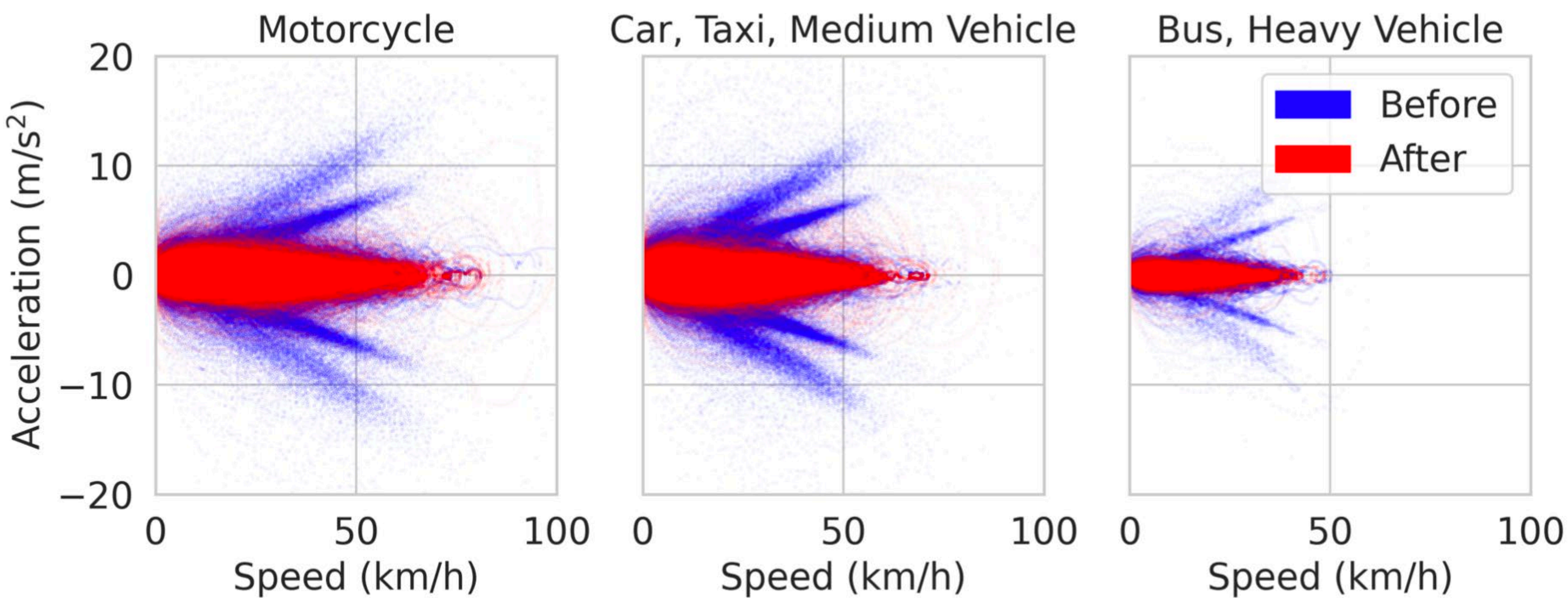
Source: V. Mahajan, E. Barmpounakis, M. R. Alam, N. Geroliminis and C. Antoniou, "Treating Noise and Anomalies in Vehicle Trajectories From an Experiment With a Swarm of Drones," in IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2023.3268712.

Noise and anomaly removal



Source: V. Mahajan, E. Barmpounakis, M. R. Alam, N. Geroliminis and C. Antoniou, "Treating Noise and Anomalies in Vehicle Trajectories From an Experiment With a Swarm of Drones," in IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2023.3268712.

Results



Source: V. Mahajan, E. Barmpounakis, M. R. Alam, N. Geroliminis and C. Antoniou, "Treating Noise and Anomalies in Vehicle Trajectories From an Experiment With a Swarm of Drones," in IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2023.3268712.

Indirect traffic flow estimation from link speeds

Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

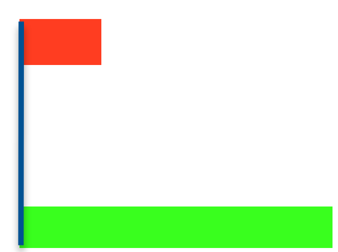
Some data are easier to collect

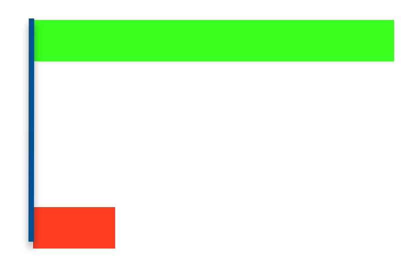
Ease of collection/ availability

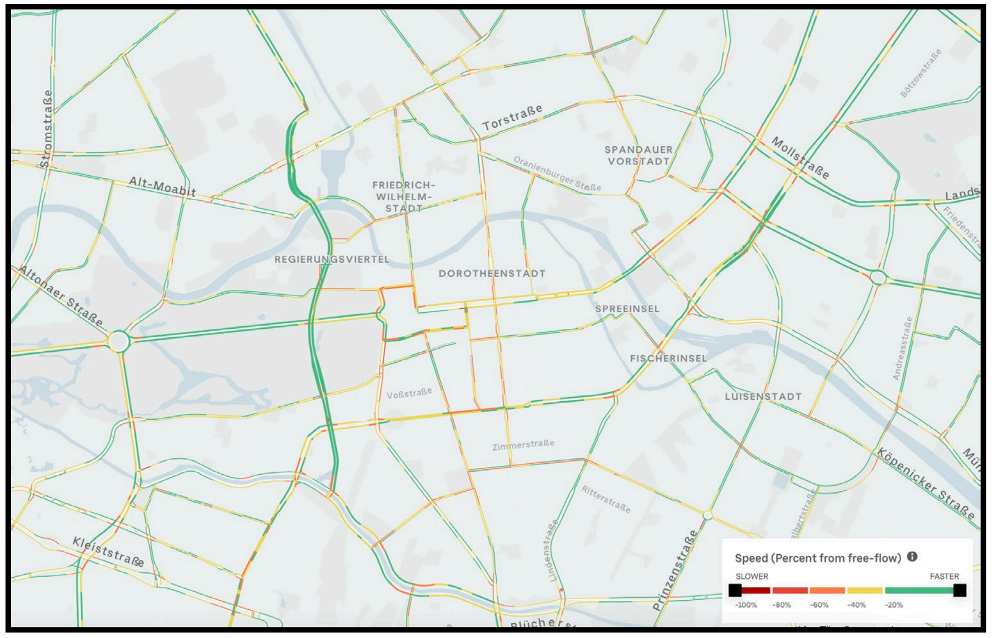
Ease of demand calibration

Link count data

Link speed data

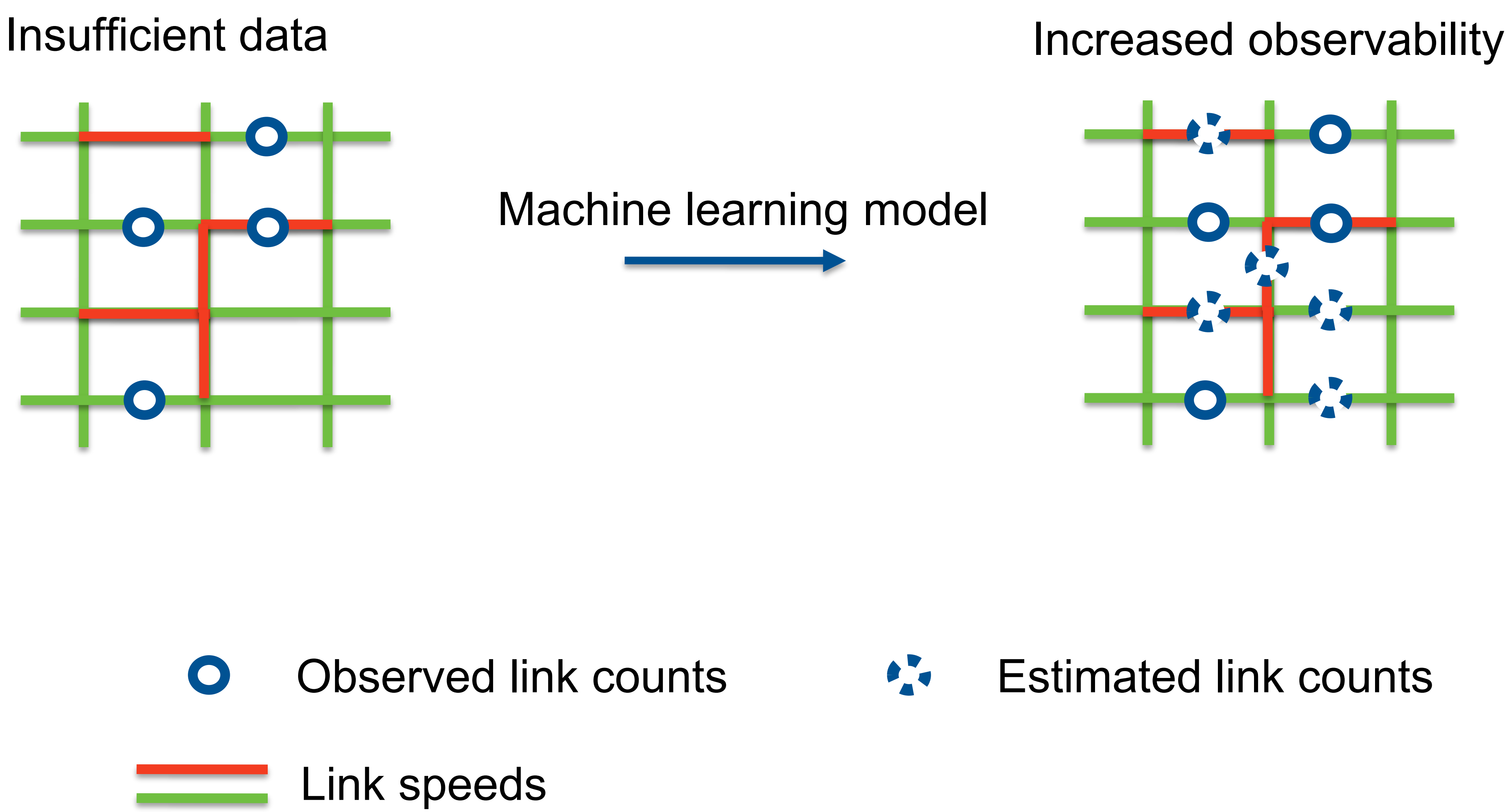






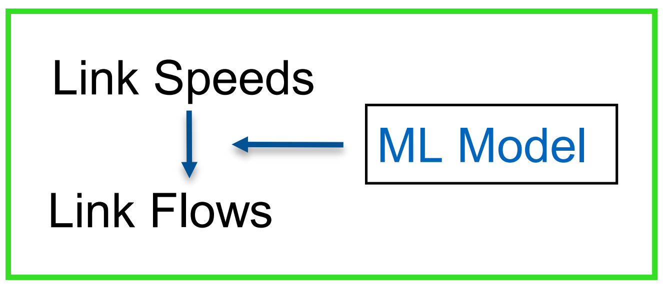
Source: UBER Movement

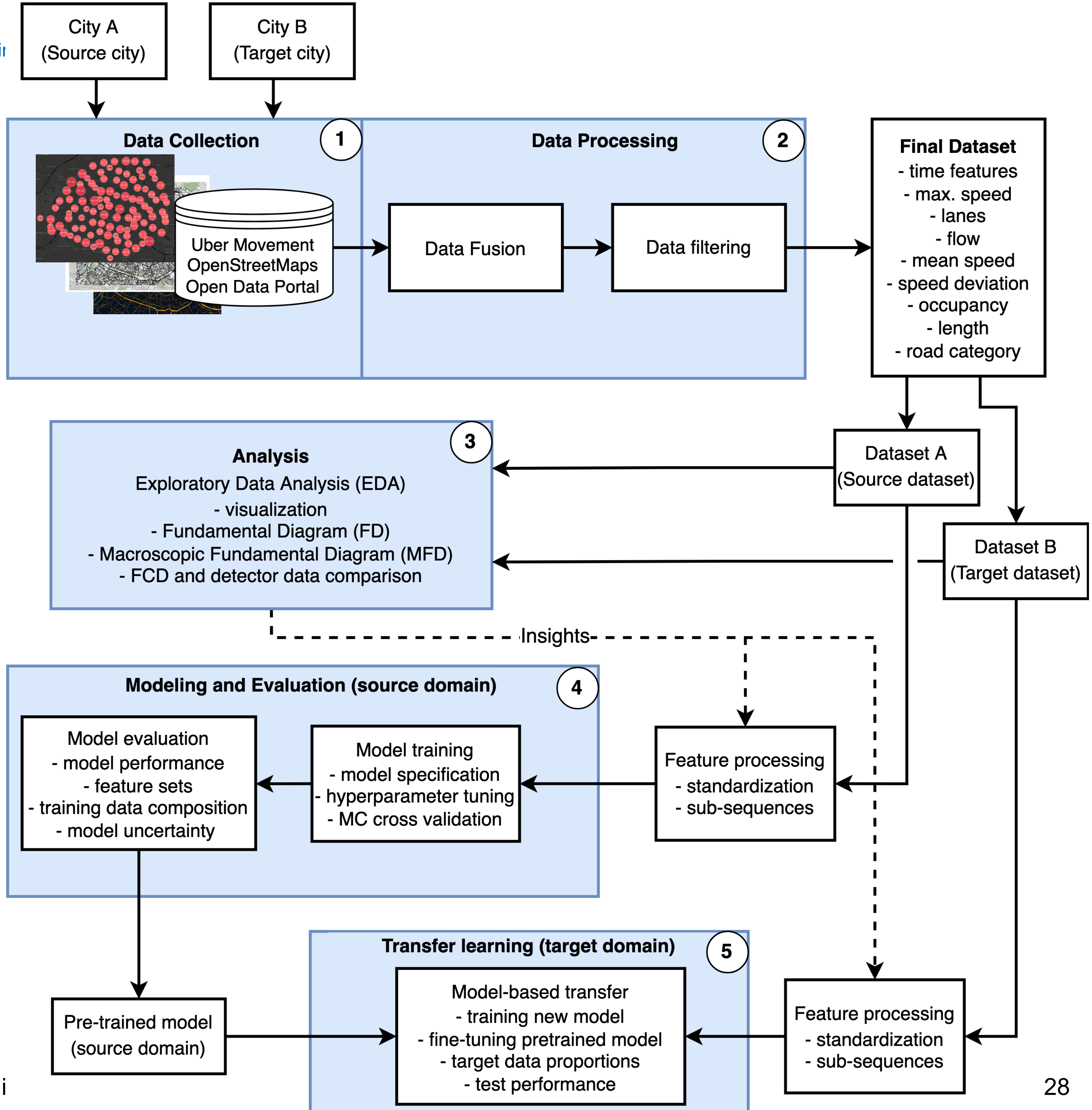
Approach


Insufficient data
Increased observability
Machine learning model
Observed link counts
Estimated link counts
Link speeds

Chair of Transportation Systems Engineeri
TUM School of Engineering and Design
Technical University Munich

Methodology





Mahajan, V., Cantelmo, G., Rothfeld, R., Antoniou, C.: Predicting network flows from speeds using open data and transfer learning. *IET Intell. Transp. Syst., 17 (4), 804-824* (2023). https://doi.org/10.1049/itr2.12305

Data Fusion

Speed from FCD (Uber Movement)

Detector: Flow (Open Data portal)

OpenStreetMaps (OSM)

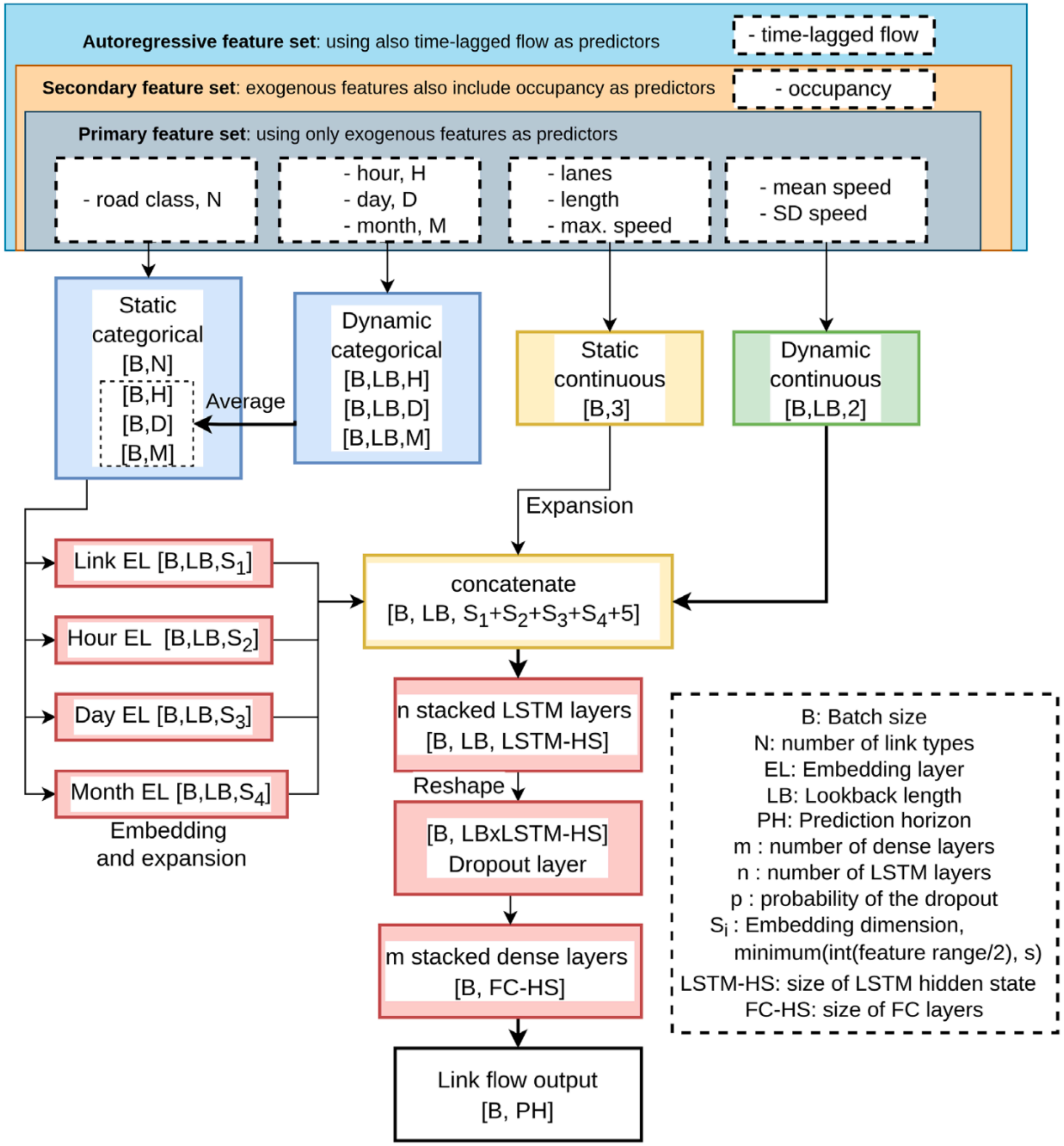
shst_id	osm_id	hour	day	month	maxspeed	highway	oneway	length	lanes	q	k	speed_kph_mean	speed_kph_stddev
de935d5b90bca31cecf7963	369834796	0	1	1	50.0	secondary	True	65.409	3.0	394.0	9.31722	36.718	5.040
f1ed9a923d64d0484d79590	[454300572, 879803157]	0	1	1	50.0	secondary	True	113.265	2.0	394.0	9.31722	35.557	14.453
f1ed9a923d64d0484d79590	84593975	0	1	1	50.0	secondary	True	30.823	2.0	394.0	9.31722	33.332	8.752

Hourly aggregation

Case studies: Paris and Madrid

Model Architecture

Architecture of the deep learning model using embedding and LSTM layers



Mahajan, V., Cantelmo, G., Rothfeld, R., Antoniou, C.: Predicting network flows from speeds using open data and transfer learning. *IET Intell. Transp. Syst., 17 (4), 804-824* (2023). https://doi.org/10.1049/itr2.12305

Results: Deep learning vs. XGBoost

Model	Link types	Loss criteria	Performance Metric	Training data	Test data
XGBoost	all	SMAPE (%)	SMAPE (%)	45.15 $\pm$2.02	51.76 $\pm$5.28
	trunk		SMAPE (%) RMSE	14.04 $\pm$1.39 725 $\pm$88	21.65 $\pm$2.93 862 $\pm$157
LSTM	all	SMAPE (%)	SMAPE (%)	40.75 $\pm$0.51	**40.17 $\pm$0.90**
	trunk		SMAPE (%) RMSE	14.05 $\pm$0.47 634 $\pm$19	**16.89 $\pm$0.31** **743 $\pm$14**

Note: RMSE is not reported for link types *"all"*, since the scale of target variable largely varies across the primary, secondary, and trunk link types.

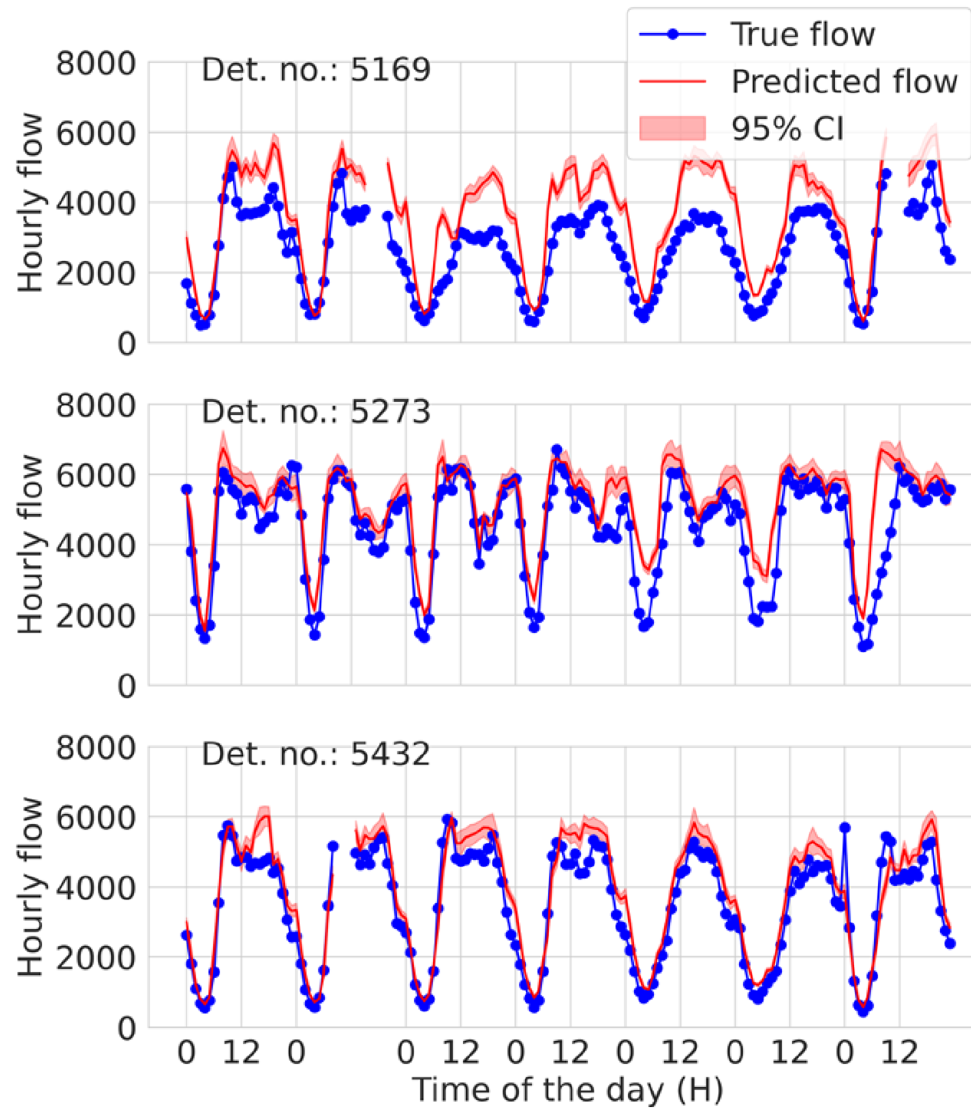
Mahajan, V., Cantelmo, G., Rothfeld, R., Antoniou, C.: Predicting network flows from speeds using open data and transfer learning. *IET Intell. Transp. Syst., 17 (4), 804-824* (2023). https://doi.org/10.1049/itr2.12305

Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

Results

Model - Input features	Lookback length (hour)	Prediction horizon (hour)	SMAPE (%)	RMSE (vehicles/hour)
LSTM - (exogenous)	**6**	**1**	**16.89 ±0.31**	**743 ±14**
	6	3	20.20 ±1.06	919 ±42
	6	6	27.18 ±0.60	1247 ±20
	6	9	77.80 ±16.45	2602 ±317
	3	1	17.30 ±0.47	751 ±24
	6	**1**	**16.89 ±0.31**	**743 ±14**
	9	1	18.00 ±1.38	815 ±73
	12	1	17.11 ±1.08	777 ±56
LSTM - (exogenous, o)	6	1	9.95 ±2.01	466 ±92
LSTM - (exogenous, o, q)	6	1	6.47 ± 0.97	321 ±37

Note: o: occupancy, q: time-lagged flow.

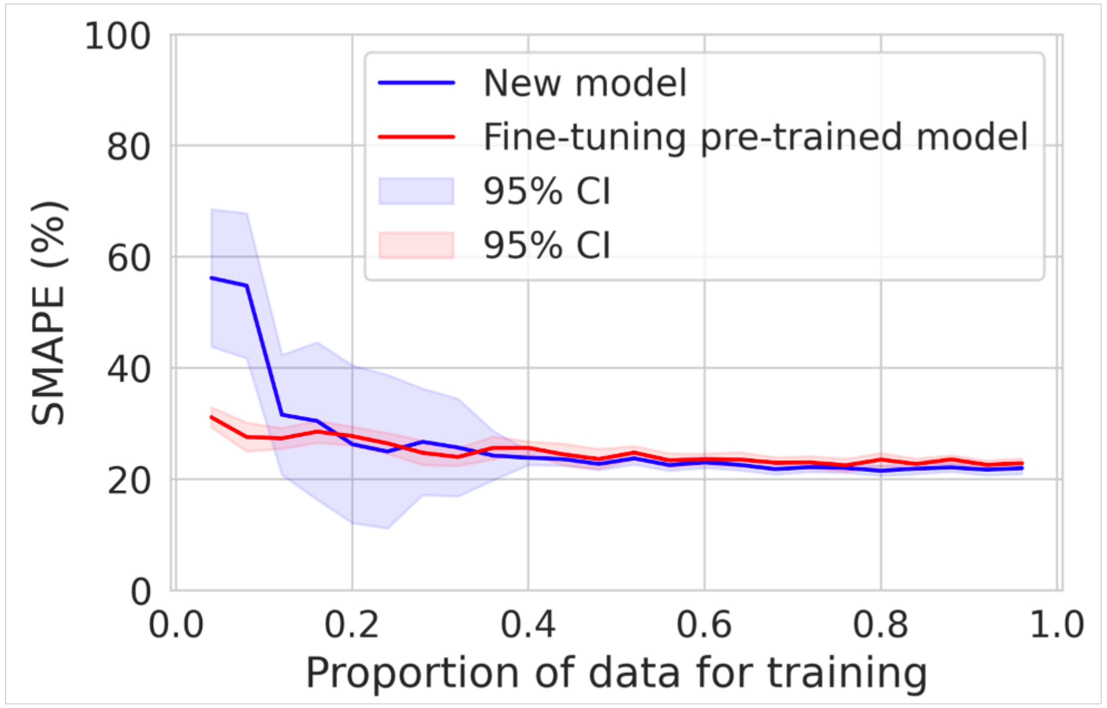


Mahajan, V., Cantelmo, G., Rothfeld, R., Antoniou, C.: Predicting network flows from speeds using open data and transfer learning. *IET Intell. Transp. Syst., 17 (4), 804-824* (2023). https://doi.org/10.1049/itr2.12305

Model transfer (from Paris to Madrid)

LSTM Model-type	Weight initialization	Proportion of target data for training	Unfrozen/ fine-tuned layers	Test SMAPE (%)	improvement over baseline (%)
Baseline	random	0.65	all	22.24 ±1.96	-
Transfer	pre-trained		None	65.82 ±2.74	-195
			FC3	46.54 ±0.39	-109
			FC2-3	46.55 ±0.34	-109
			FC1-3	45.87 ±0.29	-106
			LSTM, FC	21.24 ±0.99	4
			LSTM	21.49 ±0.90	3
			E	22.75 ±0.73	-2
			E, LSTM, FC	**20.50 ±1.10**	**8**
Baseline	random	0.10	all	55.30 ±16.06	-
Transfer	pre-trained		None	65.82 ±2.74	-19
			FC3	47.15 ±1.30	15
			FC2-3	47.29 ±1.00	15
			FC1-3	47.60 ±1.00	14
			LSTM, FC	29.07 ±1.70	47
			LSTM	**27.41 ±1.68**	**50**
			E	29.14 ±1.61	47
			E, LSTM, FC	30.92 ±1.89	44

Note: FC: fully connected layer, E: Embedding layer.

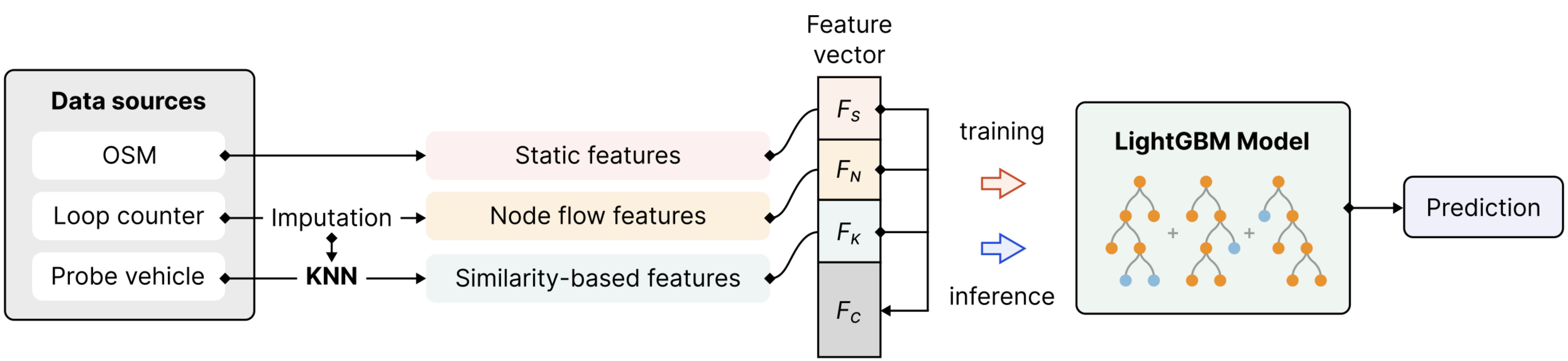


Mahajan, V., Cantelmo, G., Rothfeld, R., Antoniou, C.: Predicting network flows from speeds using open data and transfer learning. *IET Intell. Transp. Syst., 17 (4), 804-824* (2023). https://doi.org/10.1049/itr2.12305

TUM TSE at the NeurIPS 2022 Traffic4cast challenge

NeurIPS 2022 Traffic4cast challenge*

- Predicting the citywide traffic states with publicly available sparse loop count data
- Second-place winning solution to the extended challenge of ETA prediction
- Similarity-based feature extraction method using multiple nearest neighbor (NN) filters



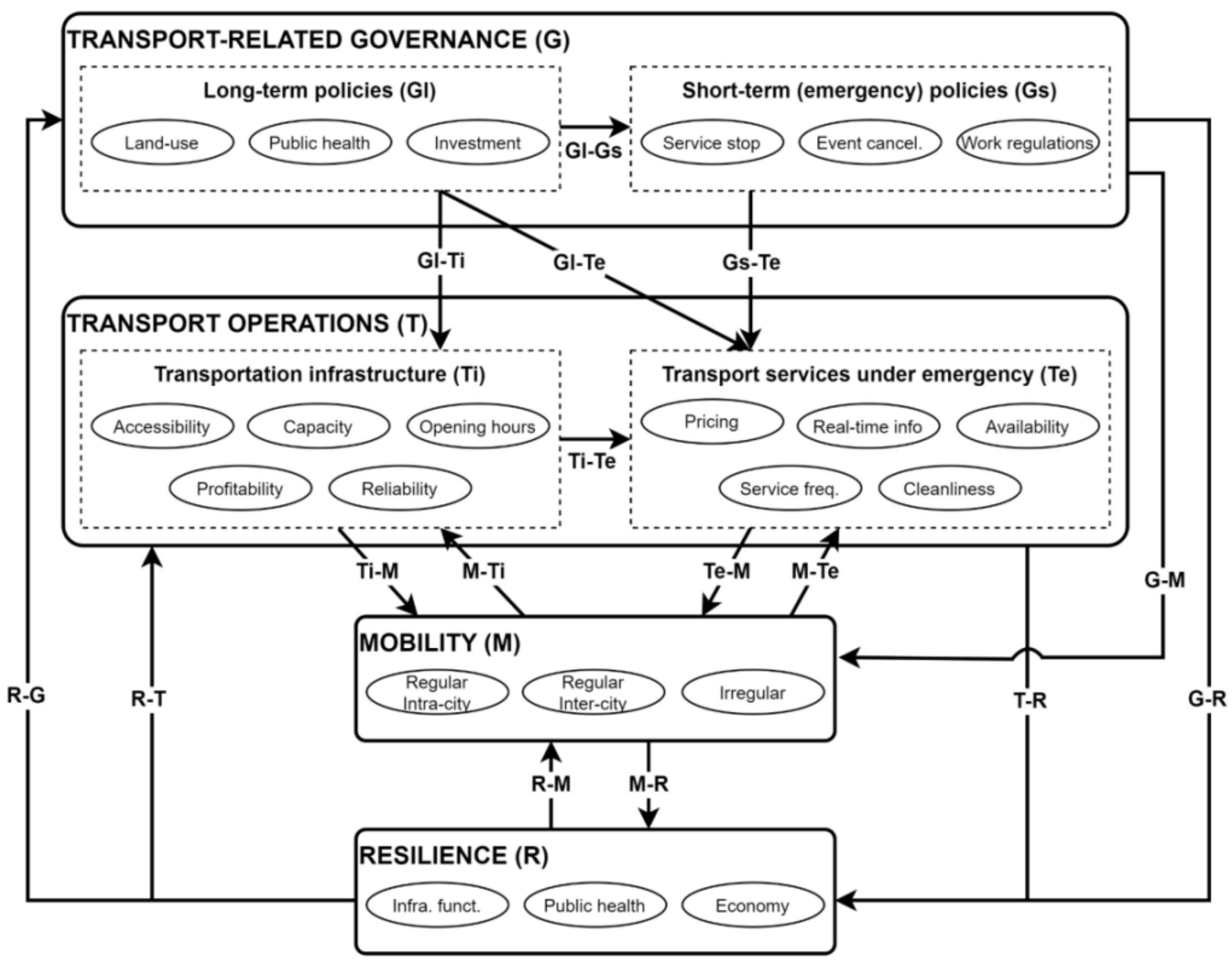
Wu, X., Lyu, C., Lu, Q., & Mahajan, V. (2022). Similarity-based Feature Extraction for Large-scale Sparse Traffic Forecasting. *arXiv*. https://doi.org/https://arxiv.org/abs/2211.07031v1

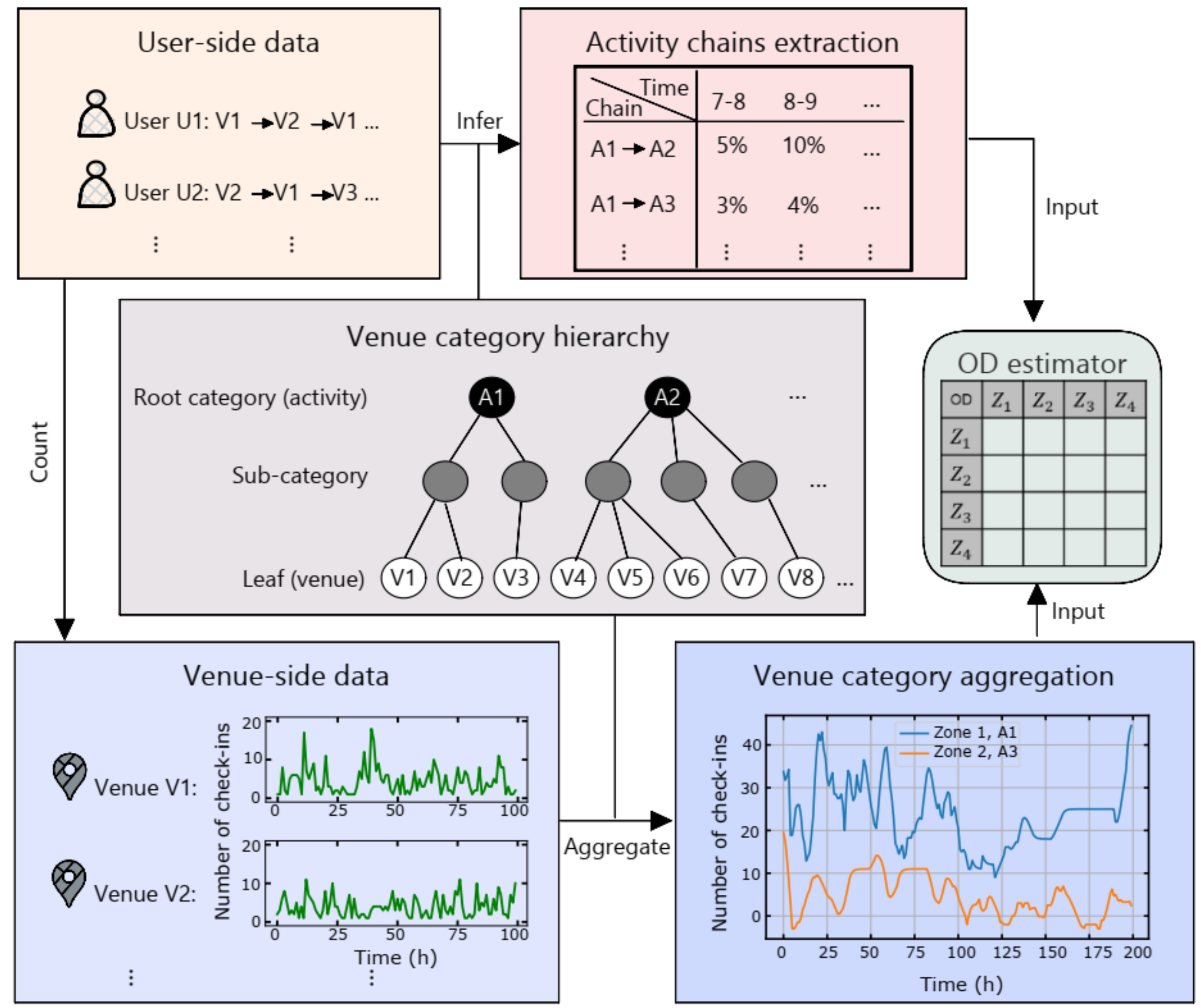
*hosted by the Institute of Advanced Research in Artificial Intelligence (IARAI)

Some ongoing projects

Big Data and Mobility – DARUMA (2021-2024)

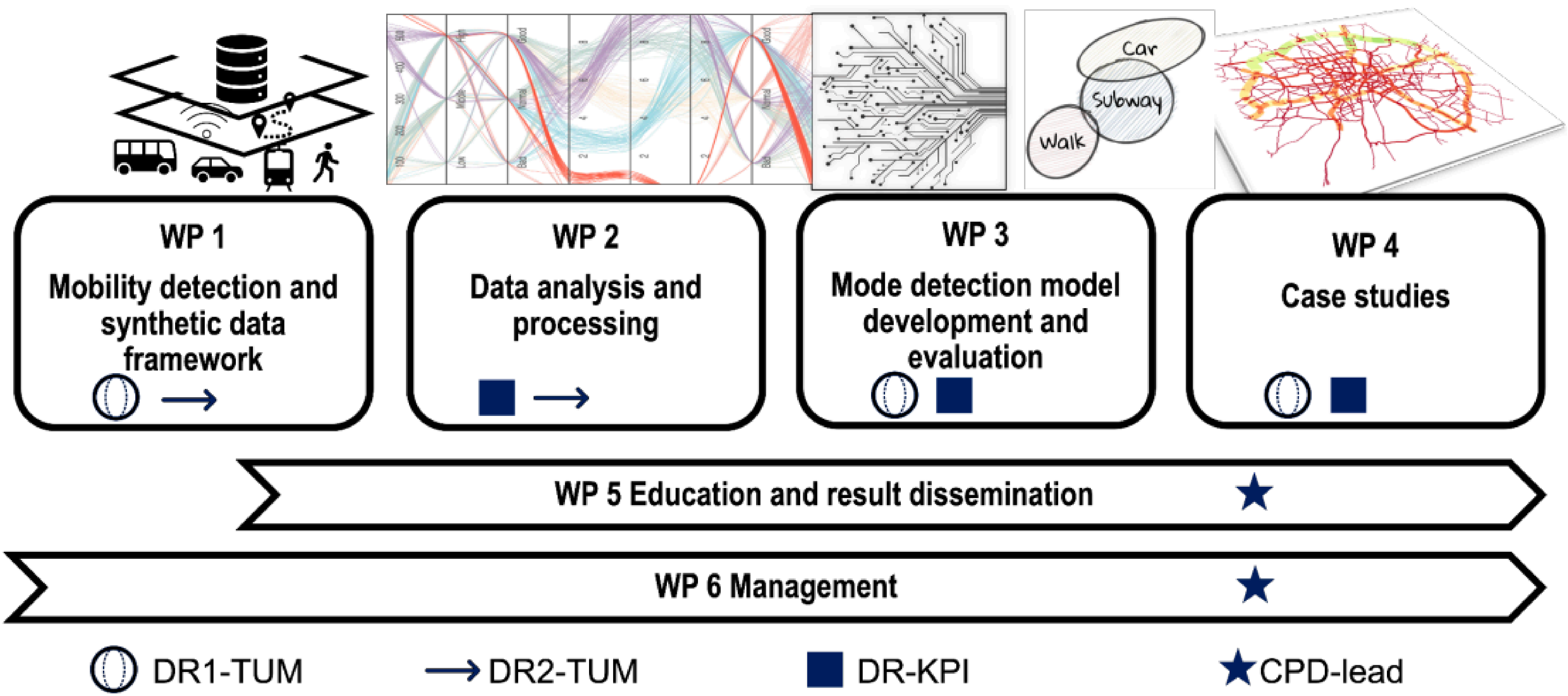
- OD demand estimation using location based social network data
- Evaluating the effectiveness of different policies for mobility peak smoothing





Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

Transport MOde Detection and Analysis (MODA, 2023-2026)



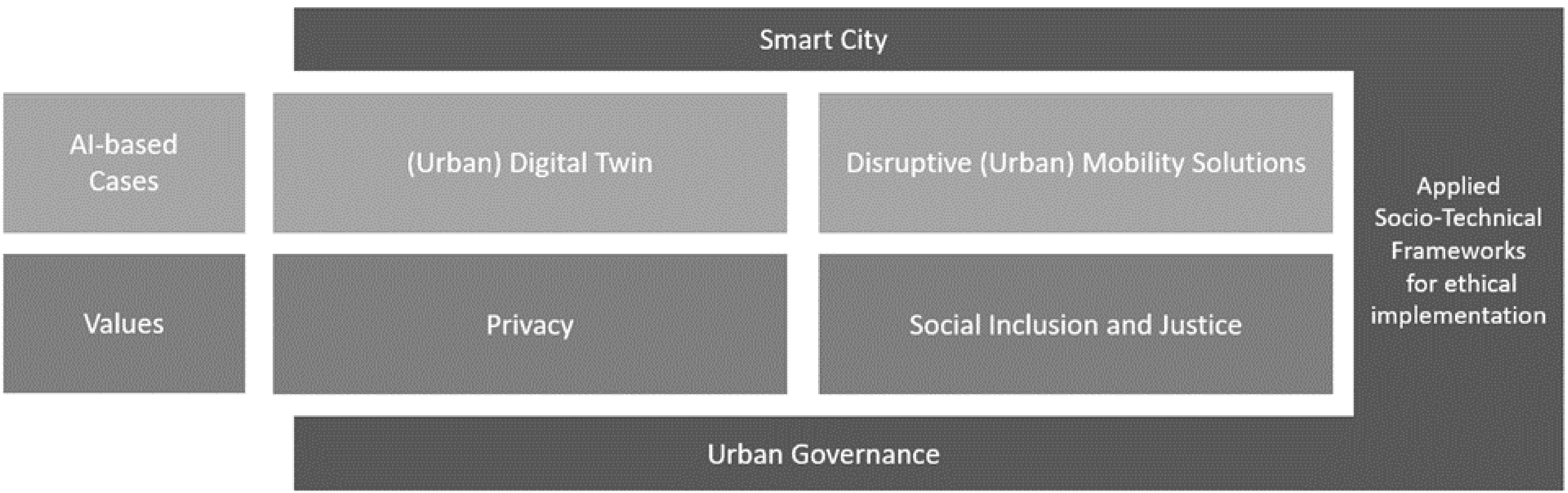
This project has received funding from the International Graduate School of Science and Engineering (IGSSE)

Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

Human Factors – Ethics in the Smart City (2022-2024)

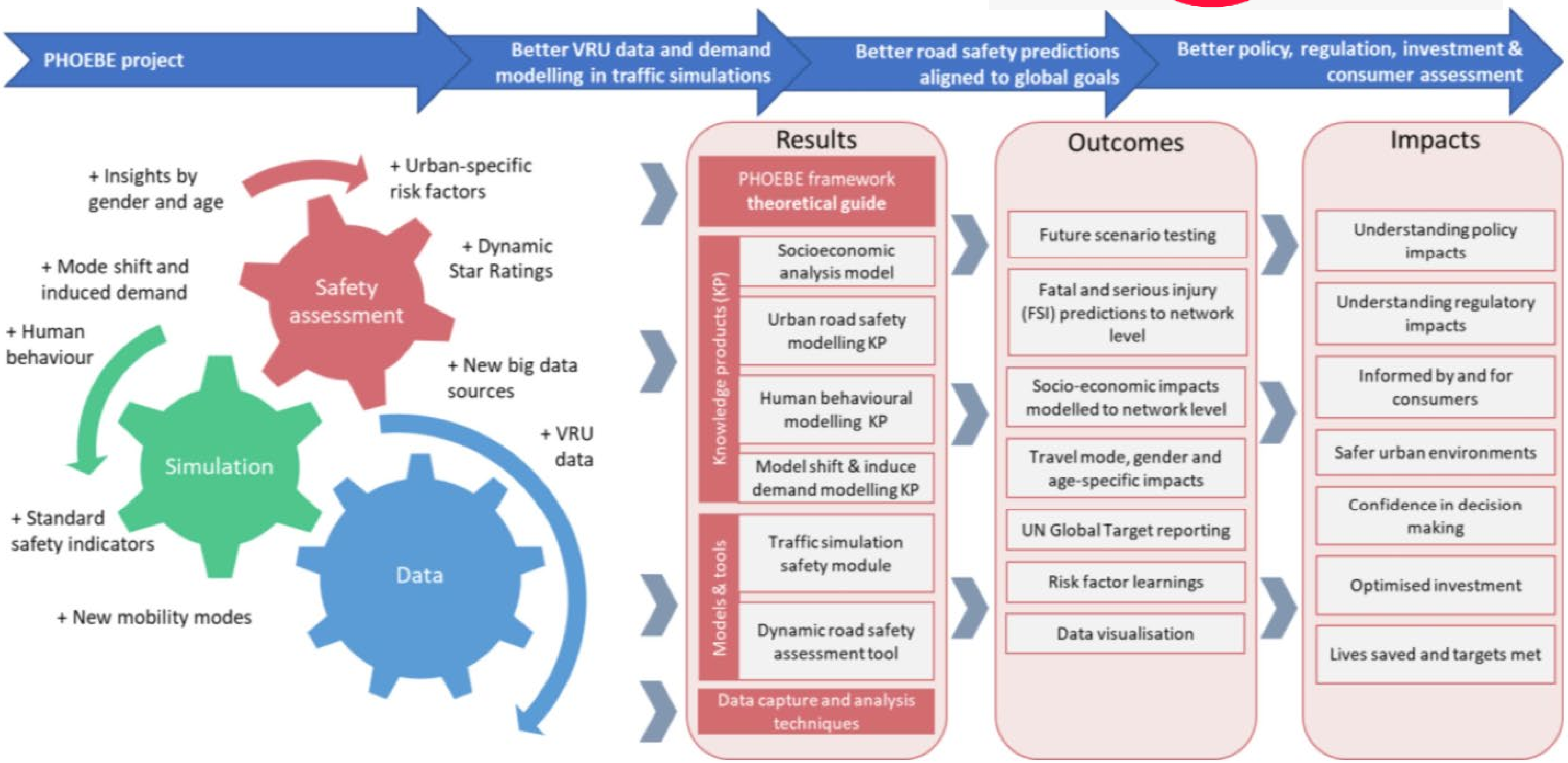
Ethics in AI: Applied socio-technical frameworks to assess the implementation of AI-related solutions.
[https://www.mos.ed.tum.de/en/vvs/forschung/projekte/ethics-for-the-smart-city/]

In this project, funded by the Institute of Ethics in Artificial Intelligence at TUM (IEAI), we collaborate with the Chair of Business Ethics, to bring together an interdisciplinary approach looking at Smart City solutions from both a social sciences and an engineering perspective.



Chair of Transportation Systems Engineering
TUM School of Engineering and Design
Technical University Munich

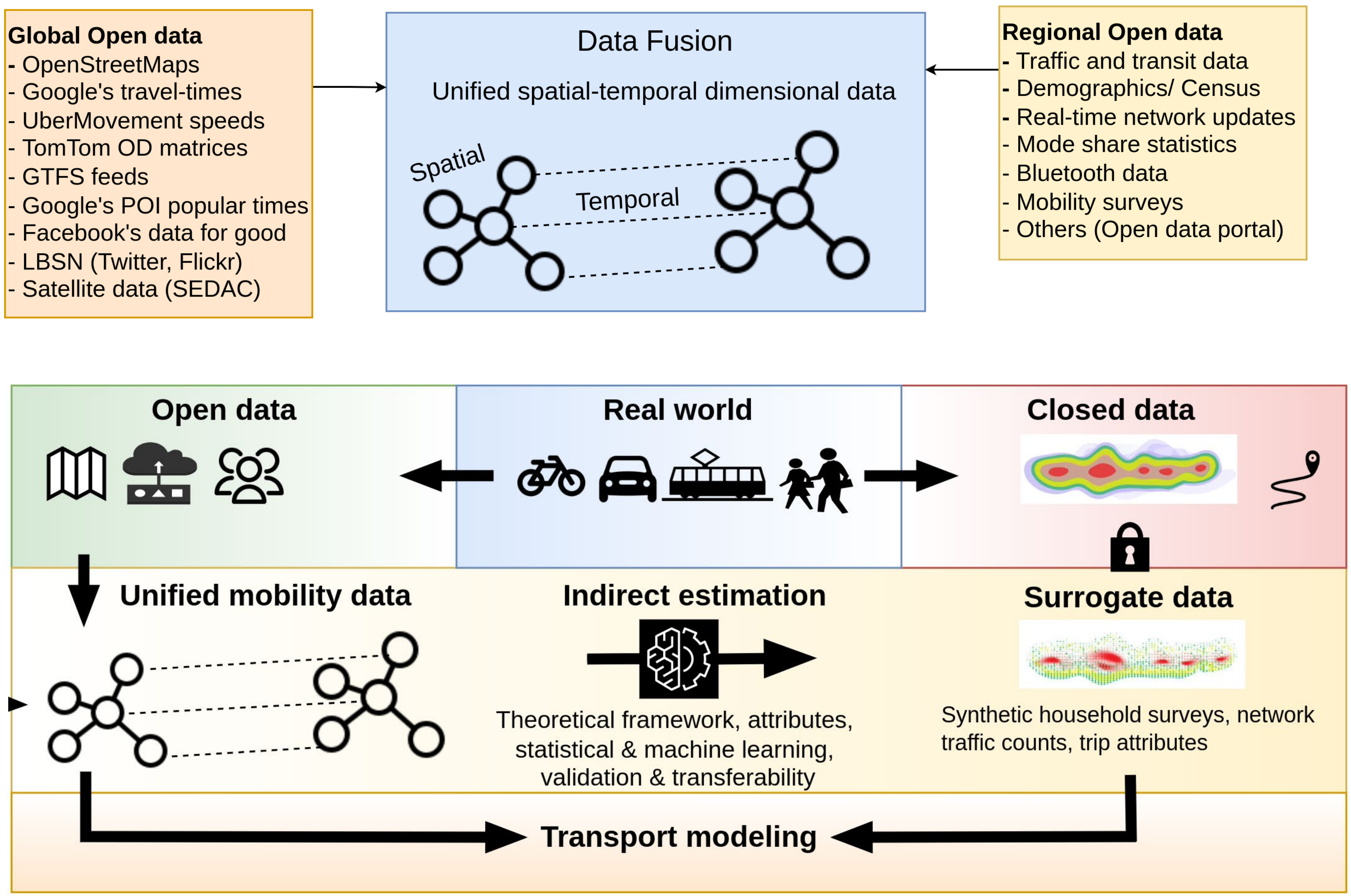
Predictive approaches for safer urban environments PHOEBE (2021-2026)



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101076963

Outlook and conclusion

Future research



Conclusion

- Open data comes with additional burden for consumers, e.g., data quality
- Open data is potentially useful for opportunistic applications, e.g., COVID-19
- Open data and its quality varies geographically

- Transfer learning can help to bridge the lack of data by knowledge transfer from data-rich contexts to data-scarce contexts

References

Mahajan, V., N. Kuehnel, A. Intzevidou, G. Cantelmo, R. Moeckel & C. Antoniou (2021)
Data to the people: a review of public and proprietary data for transport models, Transport Reviews, 42 (4), 415-440. DOI: 10.1080/01441647.2021.1977414

Mahajan, V., Cantelmo, G., Rothfeld, R., Antoniou, C.: Predicting network flows from speeds using open data and transfer learning. IET Intell. Transp. Syst., 17 (4), 804-824 (2023). https://doi.org/10.1049/itr2.12305

Mahajan, V., Cantelmo, G. & Antoniou, C. (2021) Explaining demand patterns during COVID-19 using opportunistic data: a case study of the city of Munich. Eur. Transp. Res. Rev. 13(1), 1-14.
https://doi.org/10.1186/s12544-021-00485-3

Mahajan, V., E. Barmpounakis, M. R. Alam, N. Geroliminis and C. Antoniou, "Treating Noise and Anomalies in Vehicle Trajectories From an Experiment With a Swarm of Drones," in IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2023.3268712.

Timokhin, S.; Sadrani, M.; Antoniou, C. (2020). Predicting Venue Popularity Using Crowd-Sourced and Passive Sensor Data. Smart Cities, 3(3), 818-841.

(Relatively full) list: https://antoniou.mit.edu/

Transportation, open data and artificial intelligence - Challenges and opportunities

Univ.-Prof. Dr. Constantinos Antoniou

Chair of Transportation Systems Engineering

TUM School of Engineering and Design

Technical University of Munich

c.antoniou@tum.de

https://www.mos.ed.tum.de/en/vvs/