Real-Time City-Level Traffic Prediction for Stockholm City

Matej Cebecauer[a], David Gundlegård[b], **Erik Jenelius**[a], Wilco Burghout[a]

[a]KTH Royal Institute of Technology

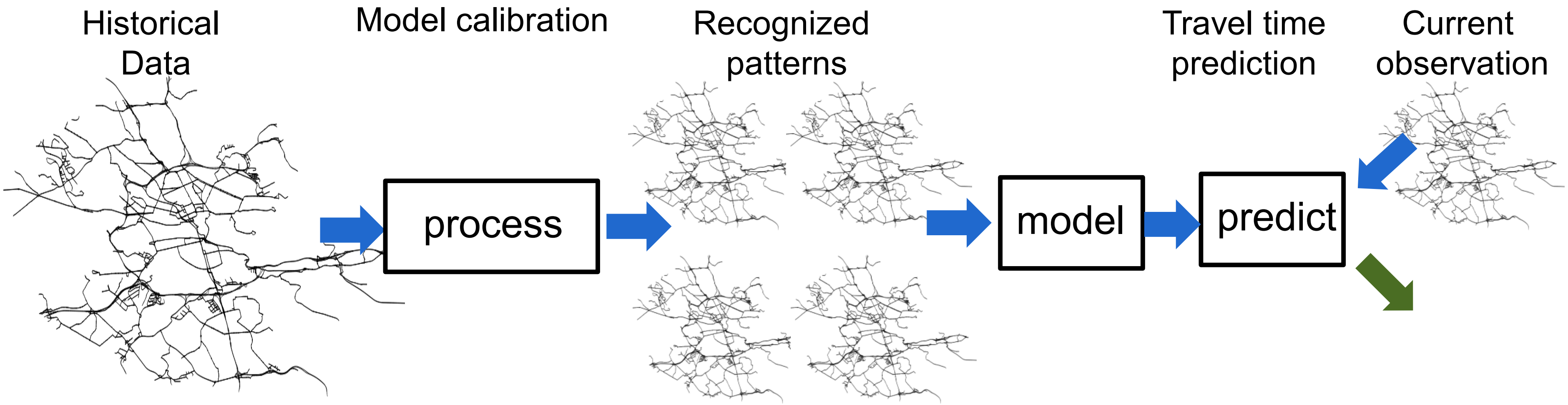
[b]Linköping University

Travel time prediction

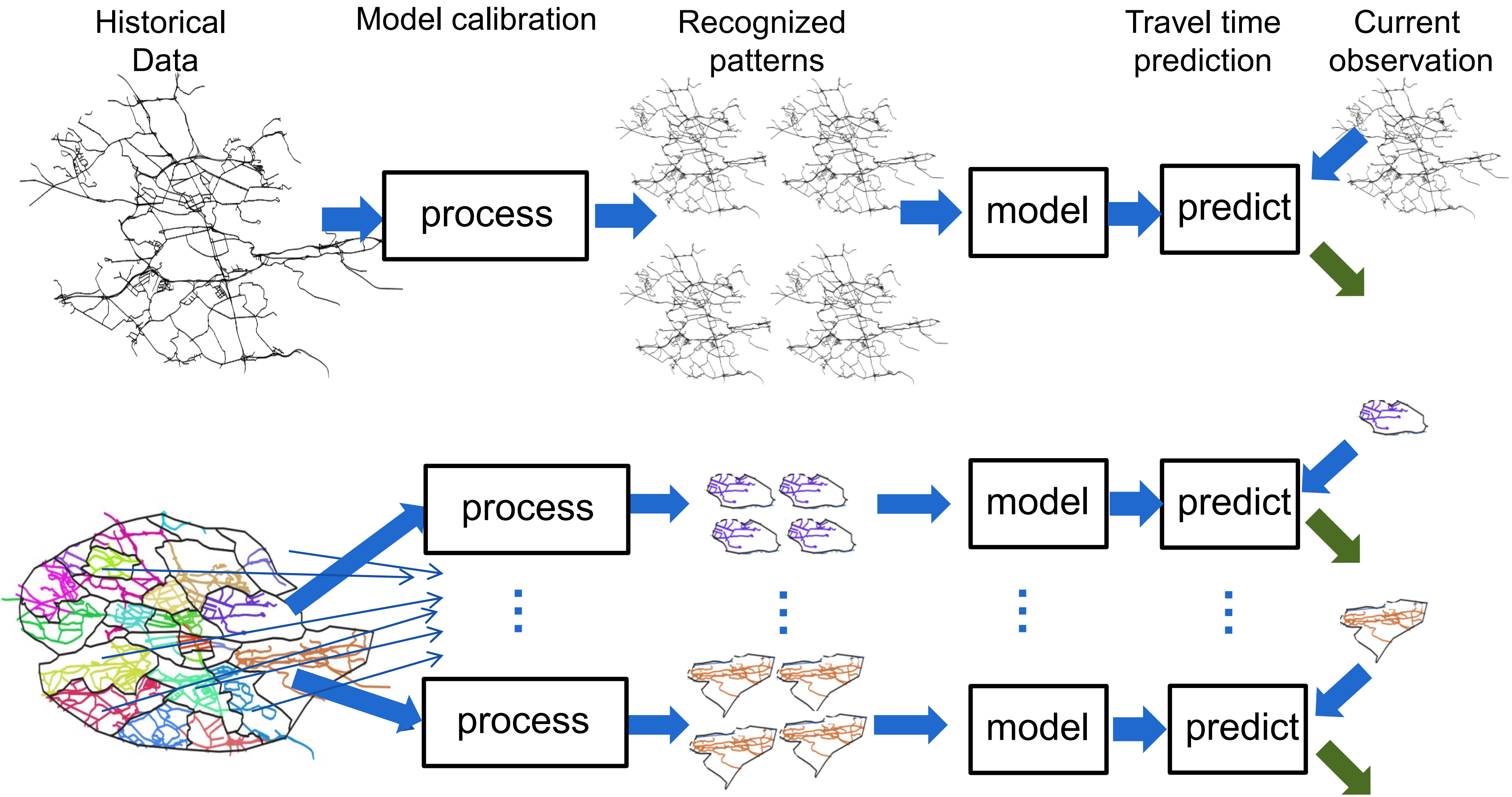
- **Use in traffic management centers and routing applications**
- Research has until recently focused mainly on motorways or main arterials
- **Large-scale urban road networks**
  Forecasting in this context is challenging:
    - Complexity
    - Heterogeneity
    - Network size

  Travel time measurements from floating car data
    - Noisy data
    - Missing data

Network partitioning



Network partitioning



Network partitioning

Bias-variance trade-off:

- **Large neighborhoods** (in the extreme, **the whole network**) can lower the variance
- **Smaller neighborhoods** (in the extreme, **each link individually**) can lower the bias
- **High bias** can lead to under-fitting the prediction model
- **High variance** can lead to over-fitting the prediction model

What partitioning method works best?

What is the effect on computational cost?

Case study

Large-scale urban network

11,340 link segments

Motorways: functional class 0 and 1

Urban roads: functional class 2 and higher

Link travel times (working days of year 2014) are estimated for 15 minute time intervals using GPS probes from 1,500 taxis

Framework for processing GPS probes to travel times on link level:

Cebecauer, M., Jenelius, E., & Burghout, W. (2018) Integrated framework for real-time urban network travel time prediction on sparse probe data. *IET Intelligent Transport Systems* 12(1), 66-74.

Computational experiments

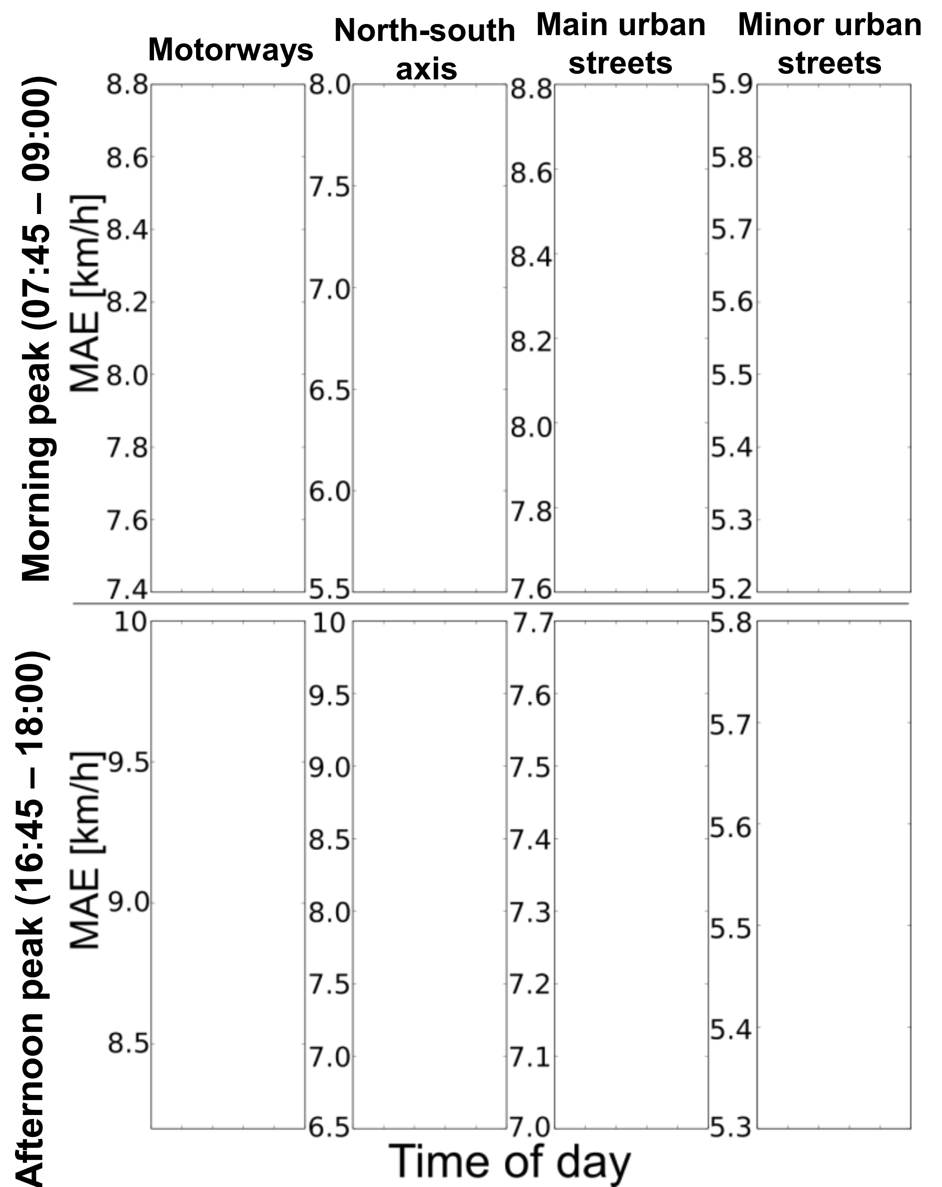
- Several **different clustering methods** and levels of aggregation are used to provide **clusters** (sets of links)
- **To each cluster** the latent factor model (Probabilistic principal component analysis **PPCA**) is applied for **short-term travel time prediction** (15 minutes horizon)
  - Prediction model is calibrated on 30 training days
- Results are evaluated for four groups of links on the 30 evaluation days:
  - Motorways func.class 0,1
  - North-south axis
  - Main urban streets func.class 2
  - Minor urban streets func.class 3,4,5

Computational experiments

Clustering approach

legend(number of clusters)

- Historical mean
- One cluster
- Cluster per link
- Functional
- Spatial
  - Districts
  - Districts & Functional
  - P-median
- Spatio-temporal
  - K-means



Method description

$$\mathrm{MAE}(i) = \frac{1}{K N_E} \sum_{k=1}^{K} \sum_{n=1}^{N_E} |\hat{v}_{ikn} - v_{ikn}|$$

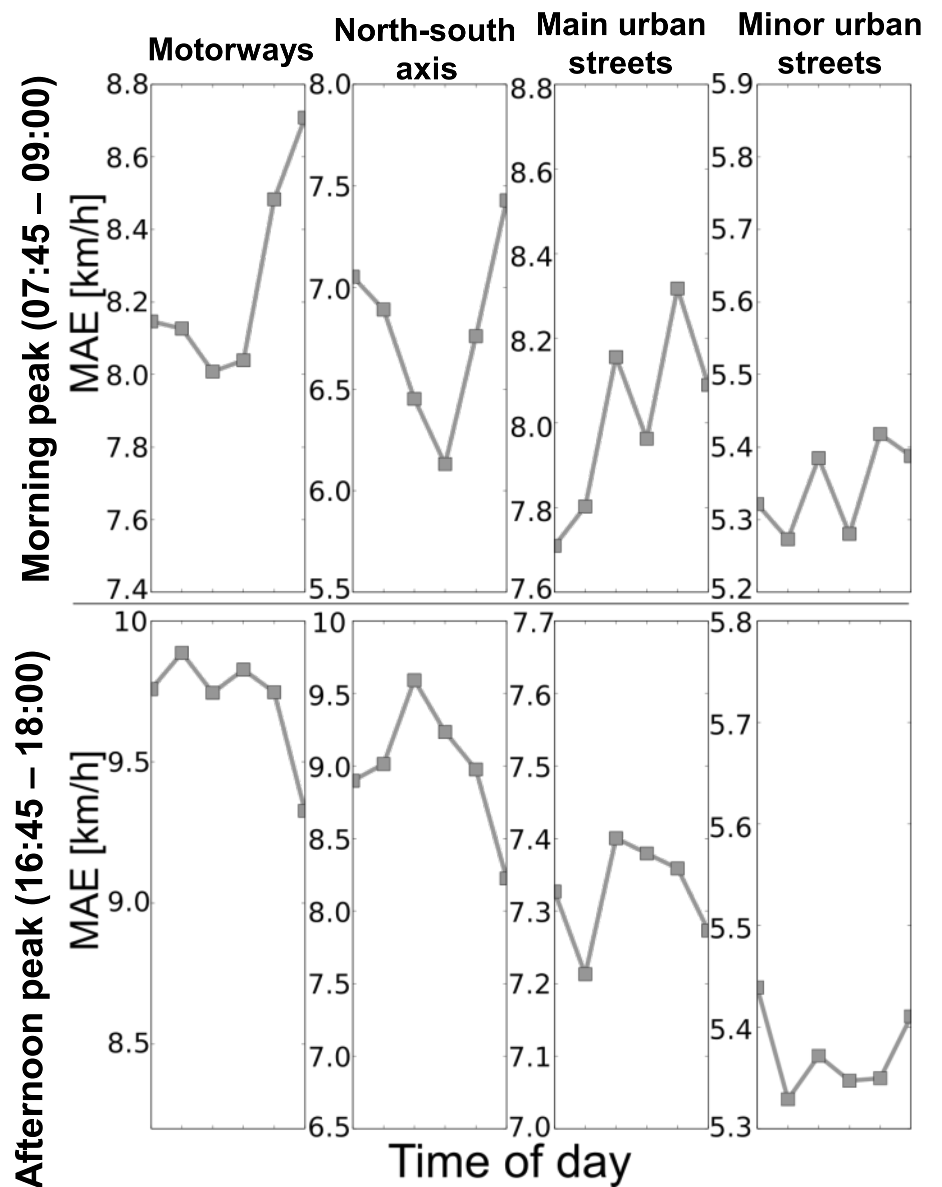
$\hat{v}_{ikn}$ - *predicted speed*
v_{ikn} - *observed speed*
K - *number of links*
N_E - *number of days for evaluation*

Computational experiments

Clustering approach legend(number of clusters)

- Historical mean (11,430)
- One cluster
- Cluster per link
- Functional
- Spatial
  - Districts
  - Districts & Functional
  - P-median
- Spatio-temporal
  - K-means

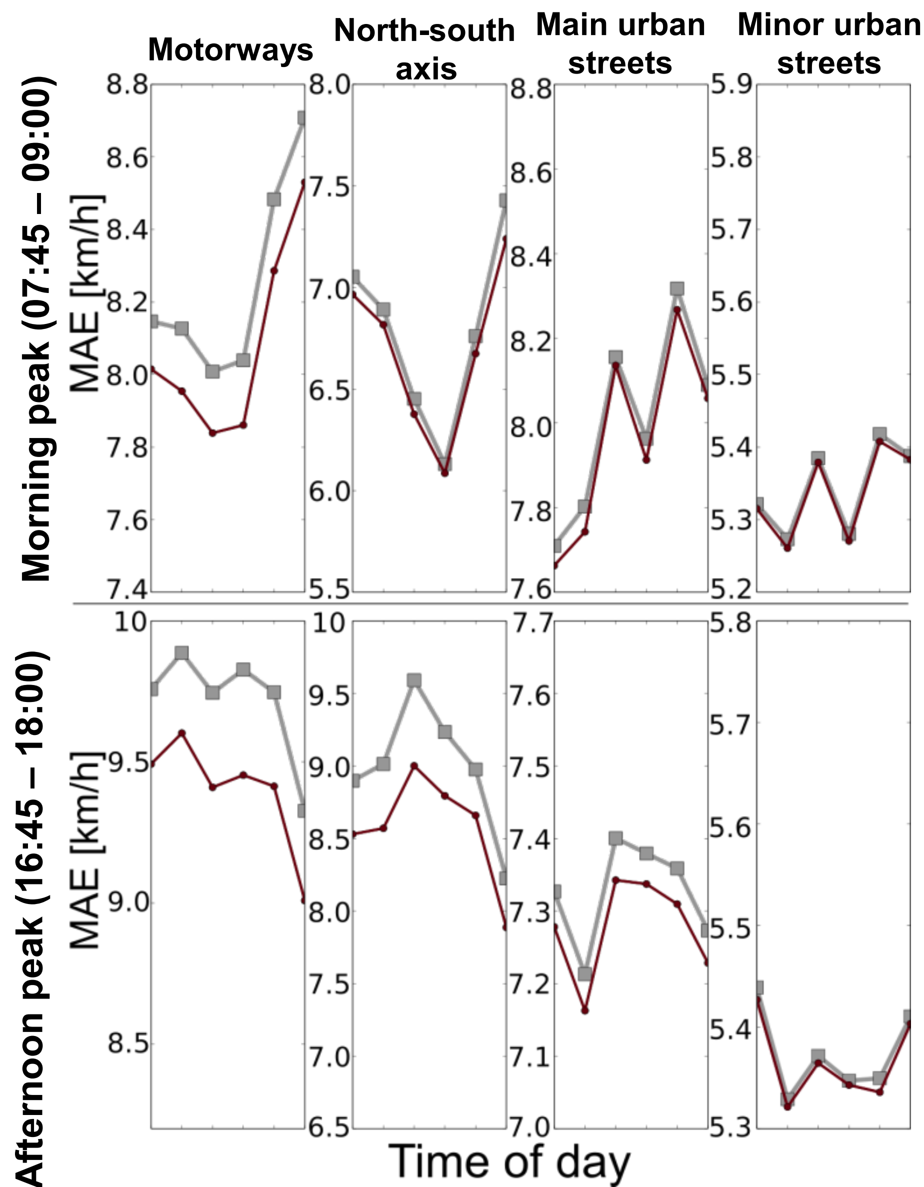
Prediction for link *k* in time interval *f* is the mean value across all historical day observations of link *k* and time interval *f.*



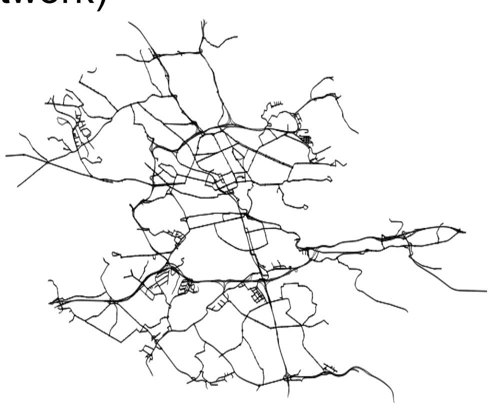
Computational experiments

Clustering approach legend(number of clusters)

- Historical mean (11,430)
- One cluster (1)
- Cluster per link
- Functional
- Spatial
  - Districts
  - Districts & Functional
  - P-median
- Spatio-temporal
  - K-means



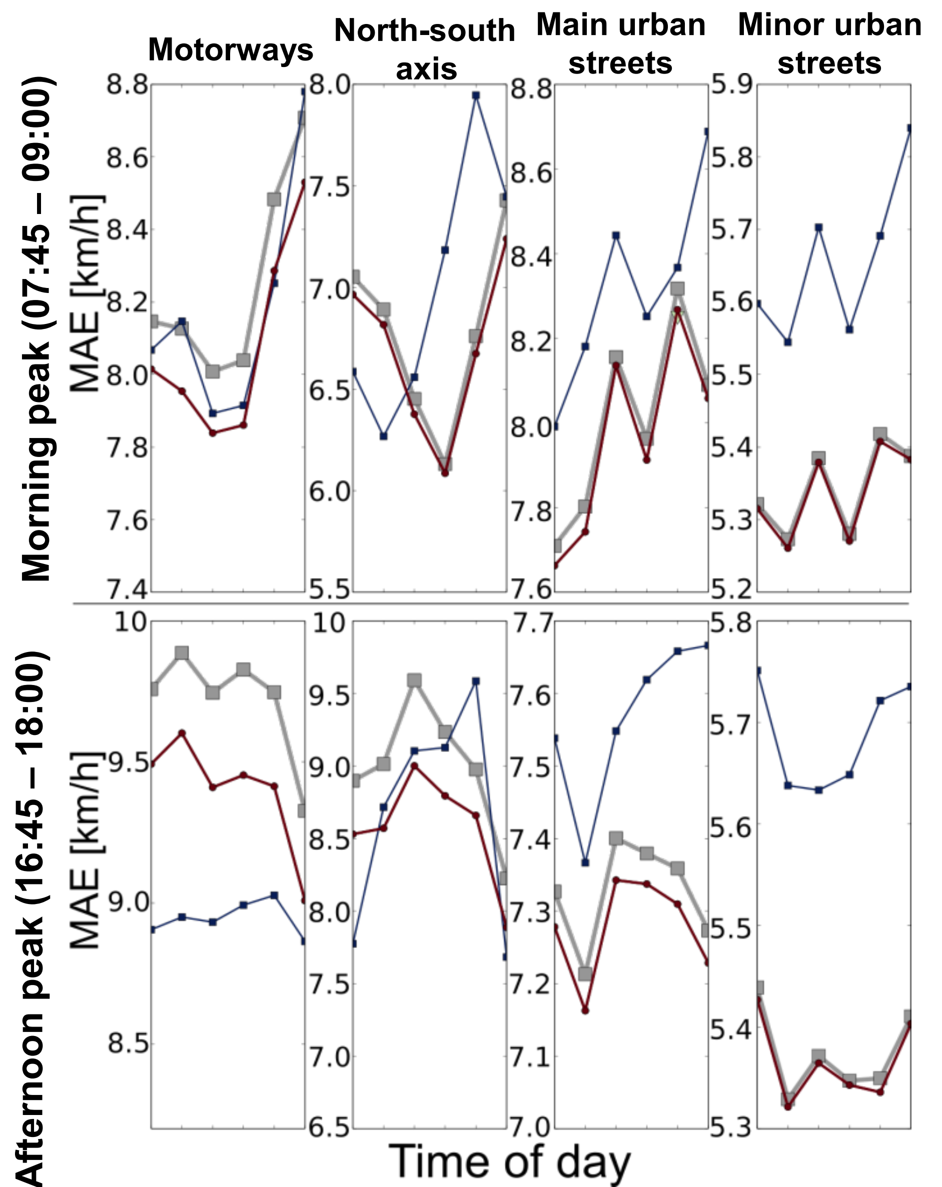
Extremely large neighborhoods
(whole network)



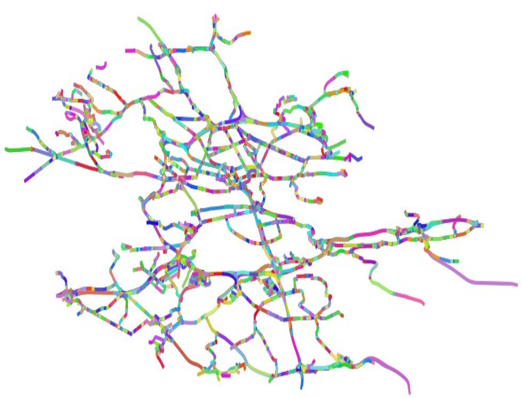
Computational experiments

Clustering approach legend(number of clusters)

- Historical mean (11,430)
- One cluster (1)
- Cluster per link (11,430)
- Functional
- Spatial
  - Districts
  - Districts & Functional
  - P-median
- Spatio-temporal
  - – K-means



Extremely small neighborhoods
(each link individually)

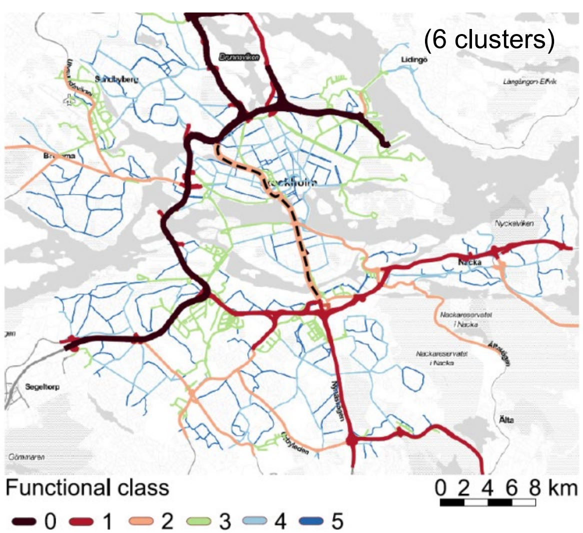


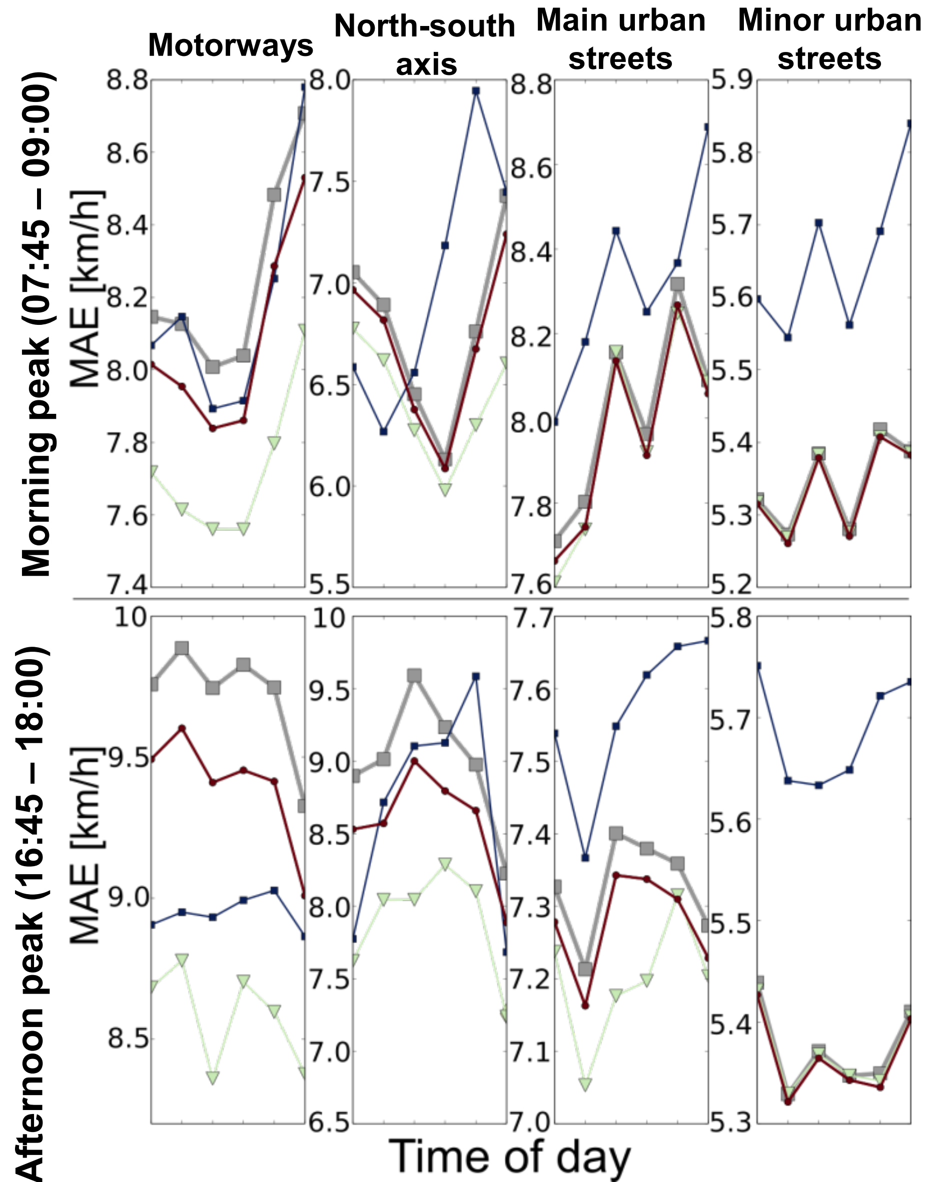
Computational experiments

Clustering approach — legend(number of clusters)

- Historical mean (11,430)
- One cluster (1)
- Cluster per link (11,430)
- Functional (6)
- Spatial
  - Districts
  - Districts & Functional
  - P-median
- Spatio-temporal
  - K-means

Clustering based on the links' **functional class attribute**



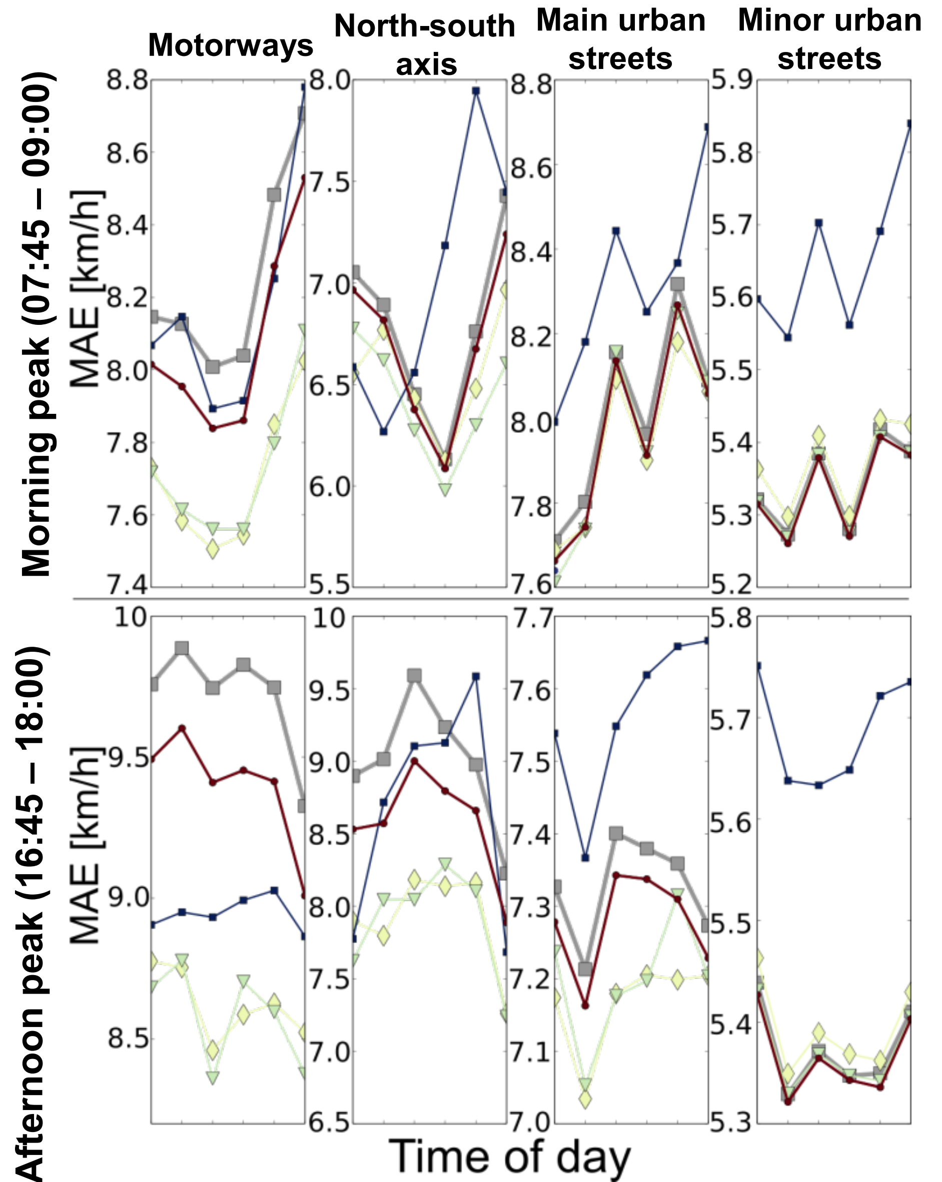


Computational experiments

Clustering approach legend(number of clusters)

- Historical mean (11,430)
- One cluster (1)
- Cluster per link (11,430)
- Functional (6)
- Spatial
  - Districts (25)
  - Districts & Functional
  - P-median
- Spatio-temporal
  - K-means

Clustering based on **administrative districts**

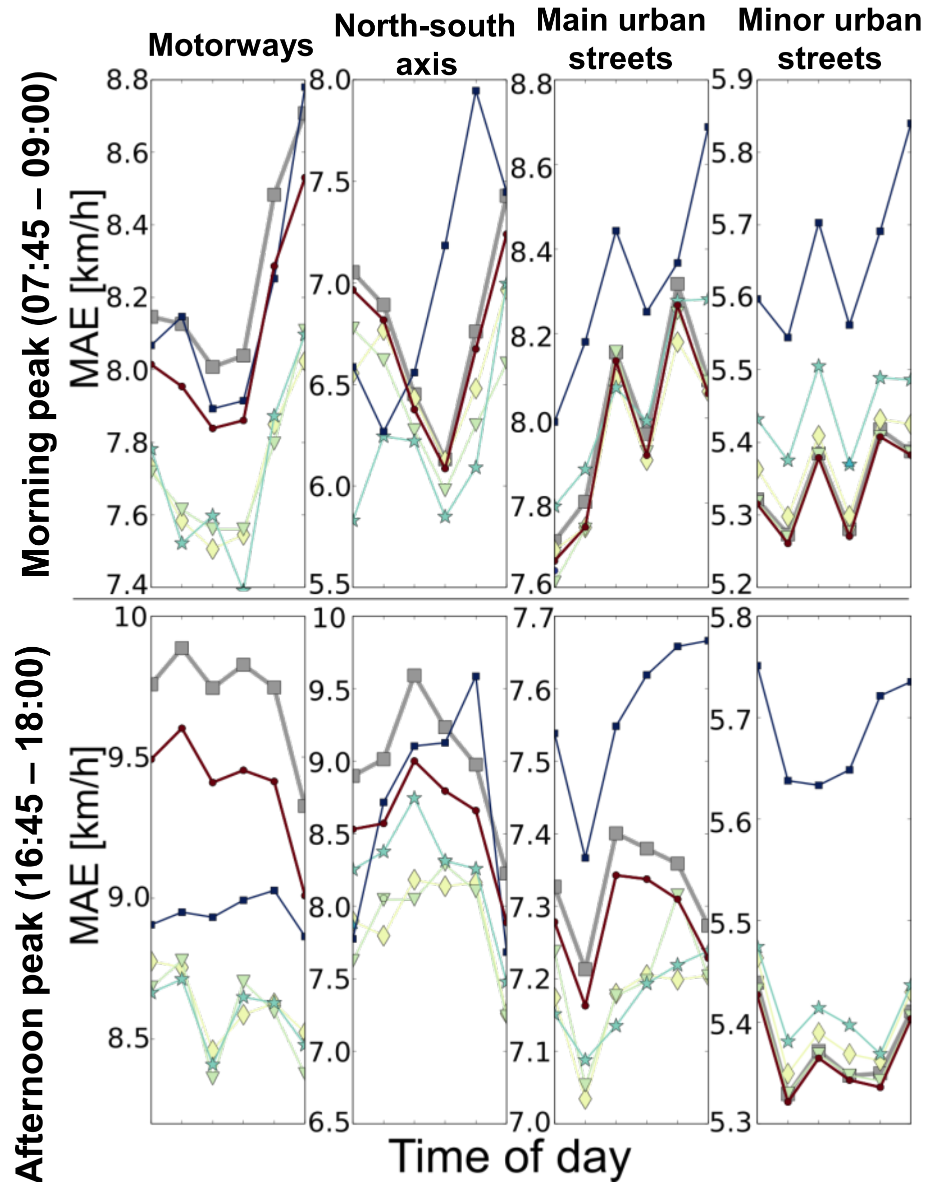


Computational experiments

Clustering approach legend(number of clusters)

- Historical mean (11,430)
- One cluster (1)
- Cluster per link (11,430)
- Functional (6)
- Spatial
  - Districts (25)
  - Districts & Functional (110)
  - P-median
- Spatio-temporal
  - K-means

Clustering based on combining the **functional class** and **administrative districts attributes.** It results in 110 non-empty sets

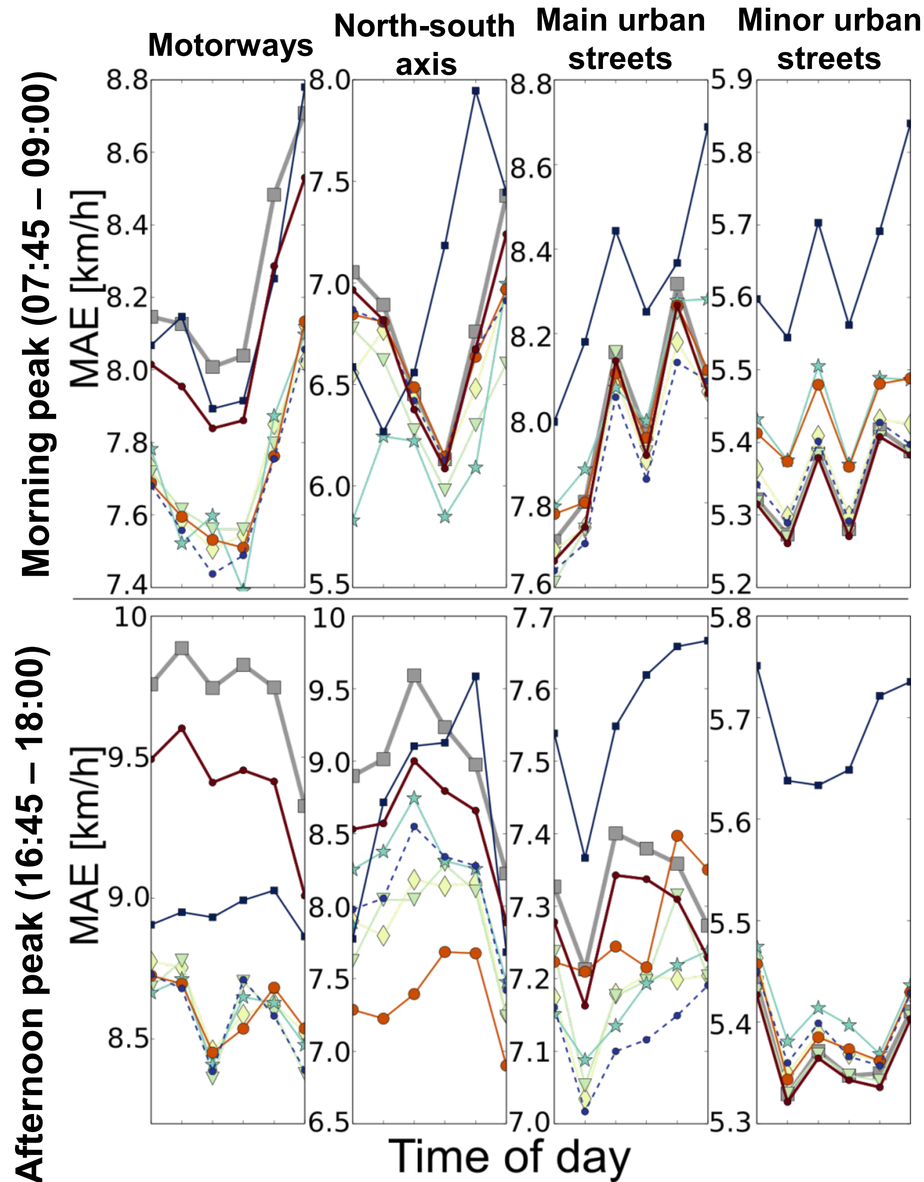


Computational experiments

Clustering approach

legend(number of clusters)

- Historical mean (11,430)
- One cluster (1)
- Cluster per link (11,430)
- Functional (6)
- Spatial
  - Districts (25)
  - Districts & Functional (110)
  - P-median (25) (110)
- Spatio-temporal
  - K-means



Clusters based on **the optimal location of centers** in the case study area, considering network distances

$$x_{kl} = \begin{cases} 1, & \text{if link } k \text{ is assigned to center link } l \\ 0, & \text{otherwise,} \end{cases}$$

$$y_l = \begin{cases} 1, & \text{if link } l \text{ is selected as center,} \\ 0, & \text{otherwise.} \end{cases}$$

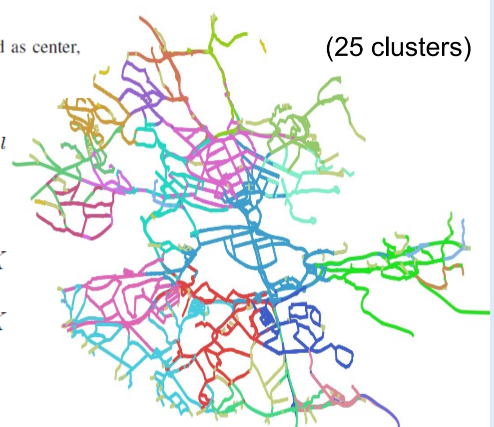
Minimize $\quad F = \sum_{k=1}^{K} \sum_{l=1}^{K} d_{kl} x_{kl}$

subject to

$$\sum_{l=1}^{K} x_{kl} = 1 \qquad k = 1, 2, \ldots, K$$

$$x_{kl} \leq y_l \qquad k, l = 1, 2, \ldots, K$$

$$\sum_{l=1}^{K} y_l = J$$



Computational experiments

Clustering approach

legend(number of clusters)

- Historical mean (11,430)
- One cluster (1)
- Cluster per link (11,430)
- Functional (6)
- Spatial
  - Districts (25)
  - Districts & Functional (110)
  - P-median (25) (110)
- Spatio-temporal
  - K-means (25) (110)

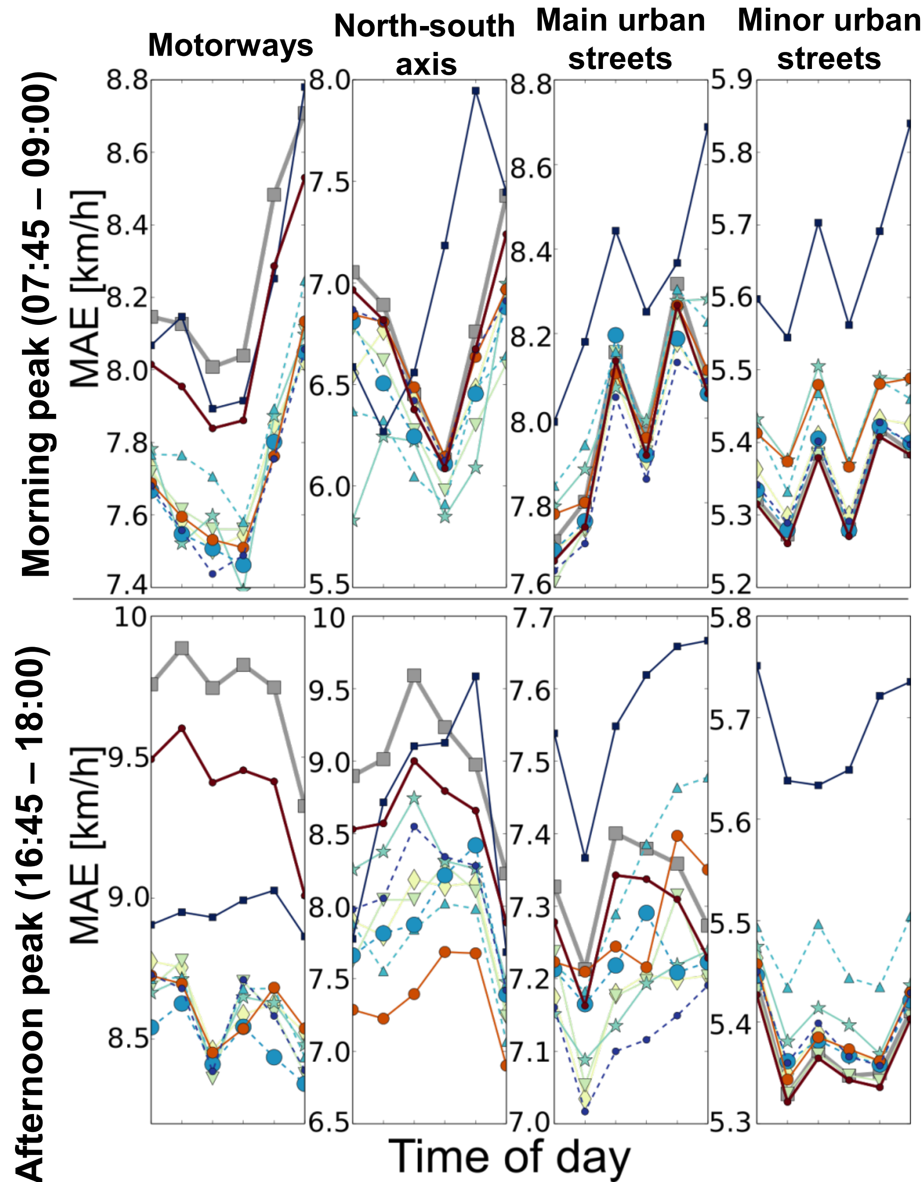
Clusters based on **k-means** consider **spatial coordinates** and **speed on the links in particular time intervals.** It aims to partition the *K* observations to *J* clusters $C = \{C_1, \dots, C_K\}$

$$\arg\min_{C} \sum_{j=1}^{J} \sum_{k \in C_j} \|\mathbf{x}_k - \mathbf{y}_j\|,$$

$$\sum_{l=1}^{K} x_{kl} = 1 \qquad k = 1, 2, \dots, K$$

$$\sum_{l=1}^{K} y_l = J$$

Centroid y_j can be any point in space

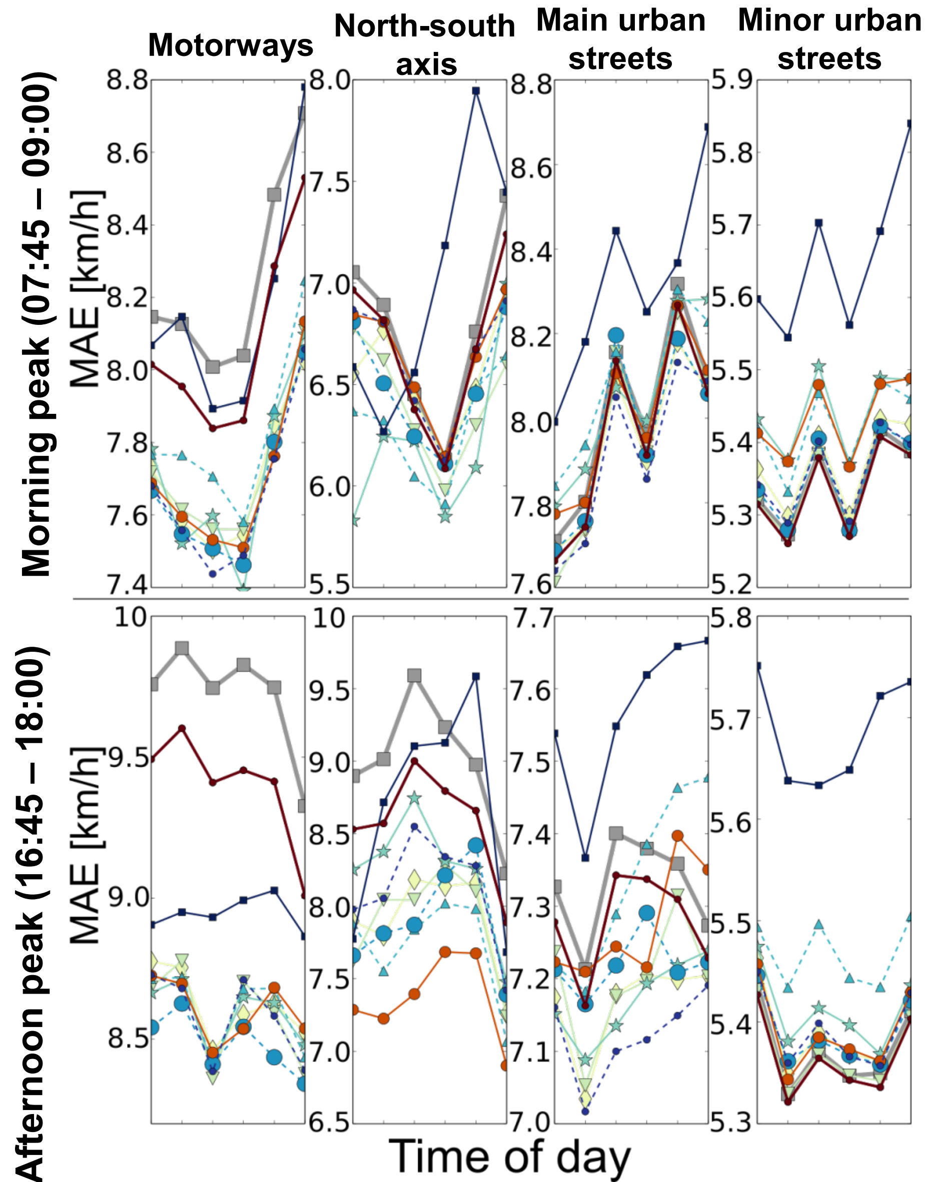


Conclusions

Clustering approach

legend(number of clusters)

- Historical mean (11,430)
- One cluster (1)
- Cluster per link (11,430)
- Functional (6)
- Spatial
  - Districts (25)
  - Districts & Functional (110)
  - P-median (25) (110)
- Spatio-temporal
  - K-means (25) (110)



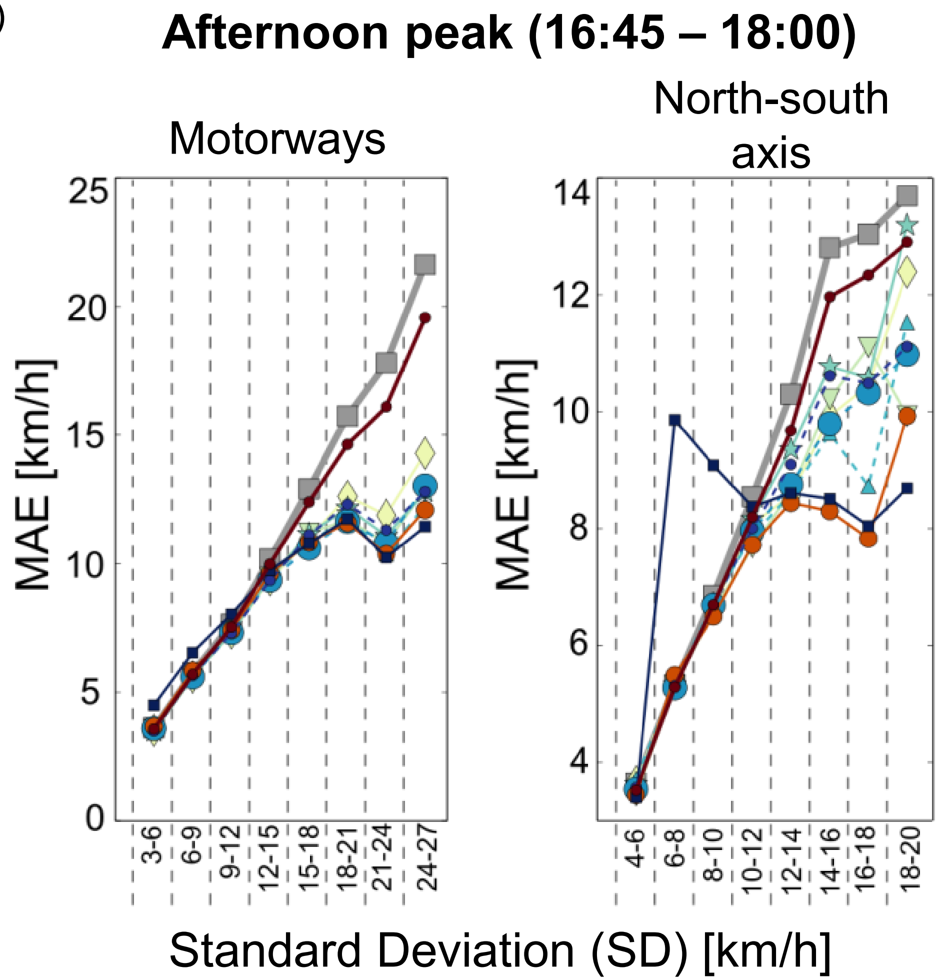
Conclusions

- **Prediction accuracy can be improved** by utilizing multivariate models over time intervals and neighborhoods of links
- **There is a bias-variance trade-off** where using larger neighborhoods can lower the variance but increase the bias
- The **appropriate number of clusters** depends on specific time interval and network region

Conclusions

Clustering approach legend(number of clusters)

- Historical mean (11,430)
- One cluster (1)
- Cluster per link (11,430)
- Functional (6)
- Spatial
  - Districts (25)
  - Districts & Functional (110)
  - P-median (25) (110)
- Spatio-temporal
  - K-means (25) (110)



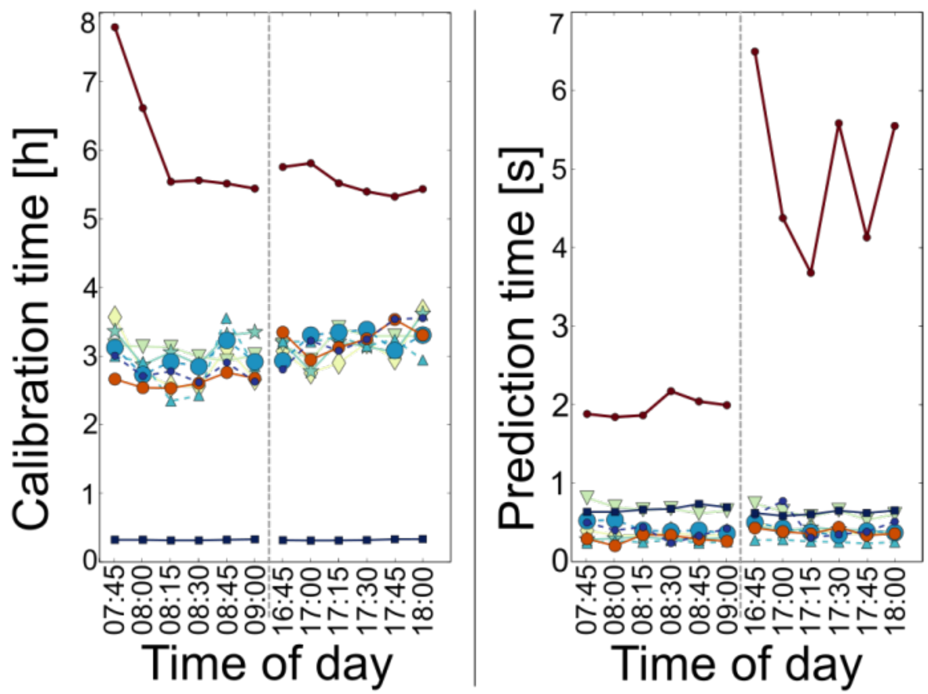
Conclusions

- When link travel time variability grows, the mean prediction error rises as well
- Benefit of partitioning increases with growing variability
- **Positive effect** especially for links with larger variability

Time efficiency

Clustering approach legend(number of clusters)

- Historical mean (11,430)
- One cluster (1)
- Cluster per link (11,430)
- Functional (6)
- Spatial
  - Districts (25)
  - Districts & Functional (110)
  - P-median (25) (110)
- Spatio-temporal
  - K-means (25) (110)



Conclusions

- Decrease of computational cost
- Enables real-time prediction

Real-Time City-Level Traffic Prediction for Stockholm City

Matej Cebecauer[a], David Gundlegård[b], **Erik Jenelius**[a], Wilco Burghout[a]

[a]KTH Royal Institute of Technology

[b]Linköping University