Technische Universiteit Eindhoven University of Technology

Aggregation of Spatio-temporal and Event Log Databases for Stochastic Characterization of Process Activities

Rodrigo M. T. Gonçalves, Rui Jorge Almeida, João M. C. Sousa

TKI DINALOG
Dutch Institute for Advanced Logistics

DAIPEX

TÉCNICO LISBOA

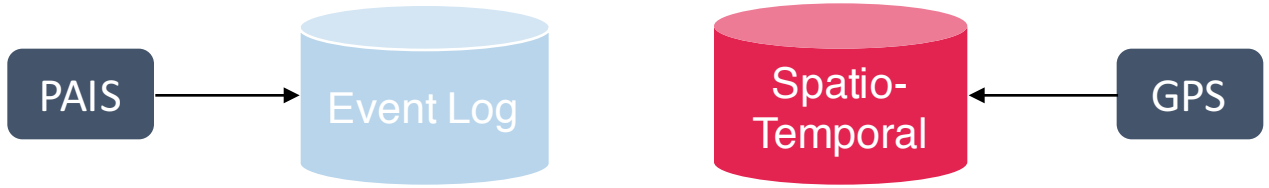
Planning and Scheduling

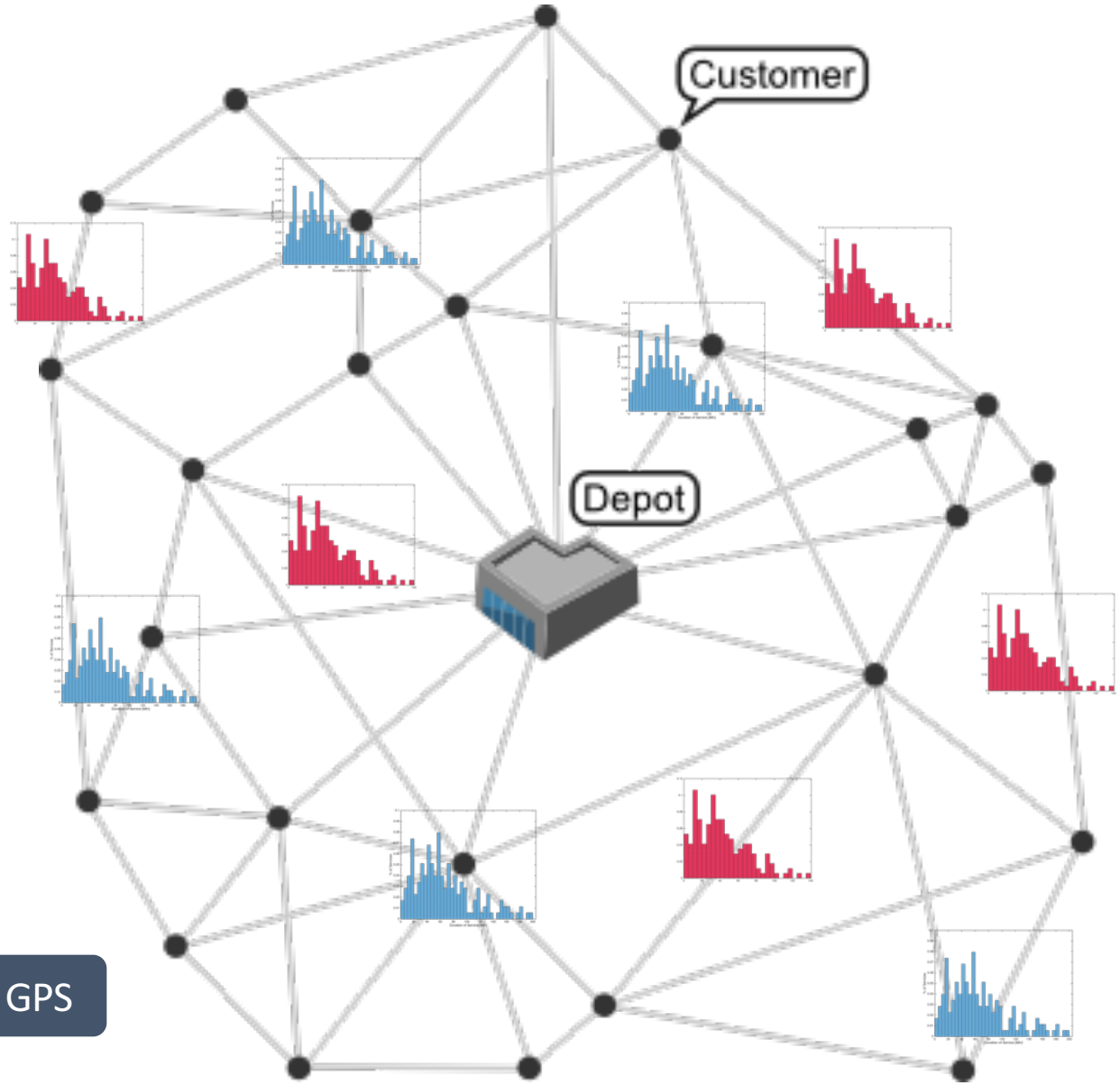
Technische Universiteit Eindhoven University of Technology

In logistic domains, transportation planning and scheduling are made based on a-priori knowledge about processes:

- **Service duration at each customer;**
- **Travel times between customers;**

Stochastic Data

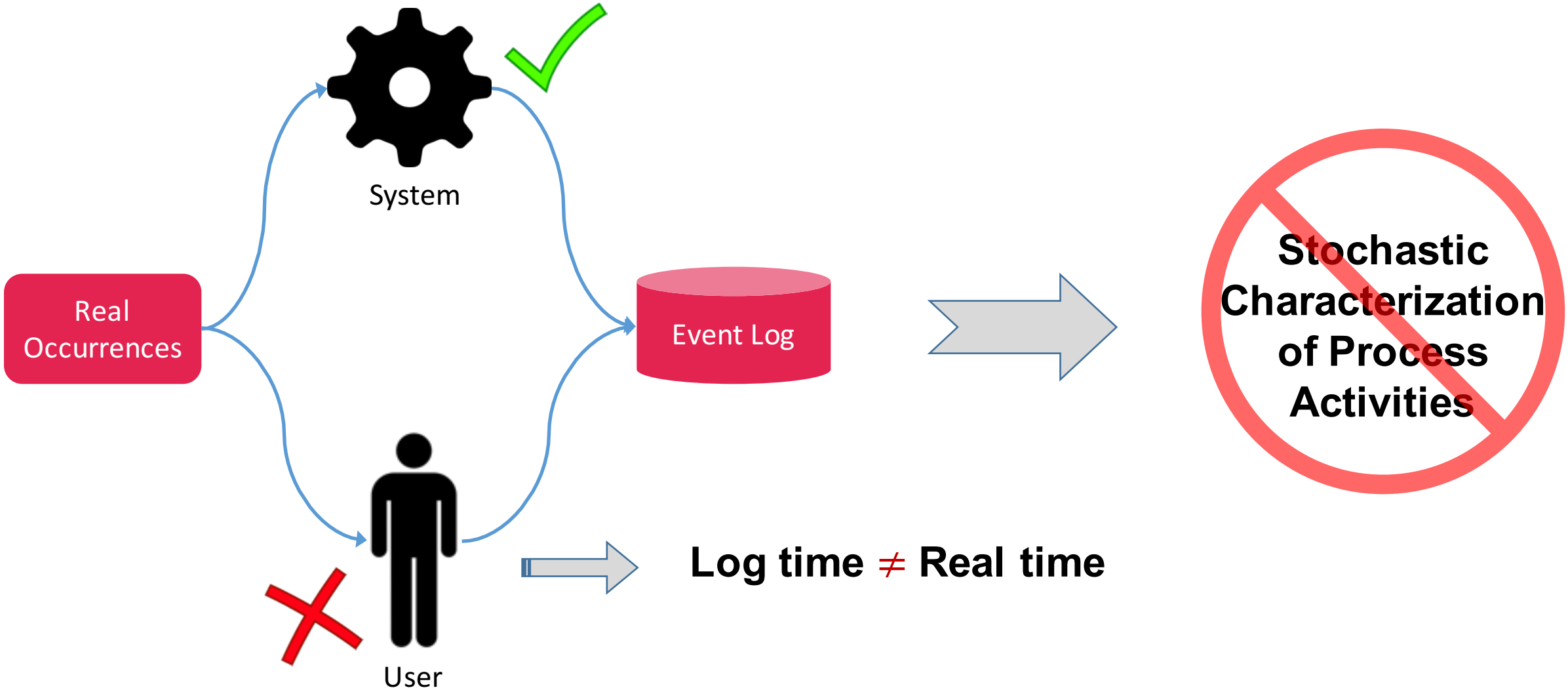




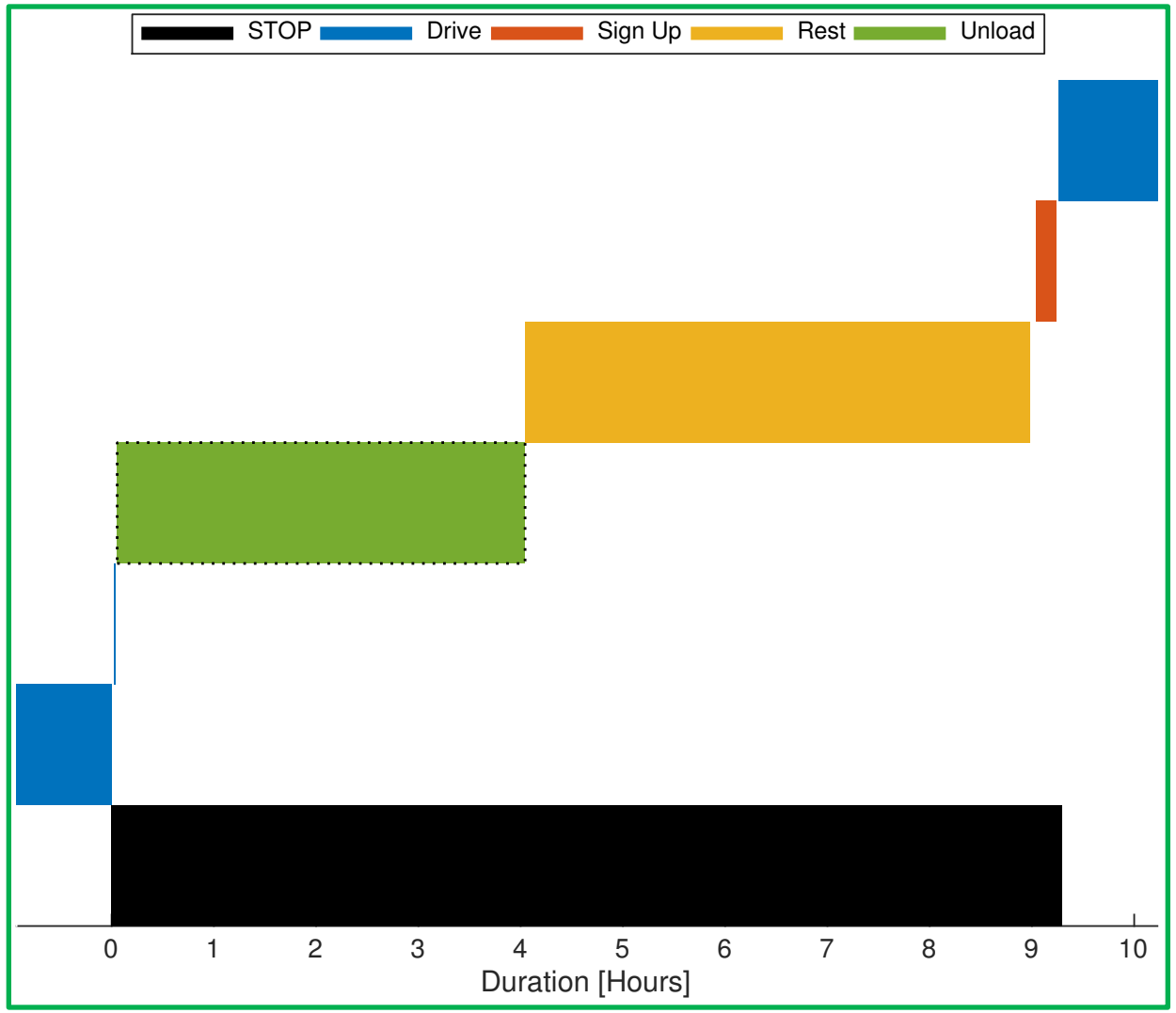
The Problem

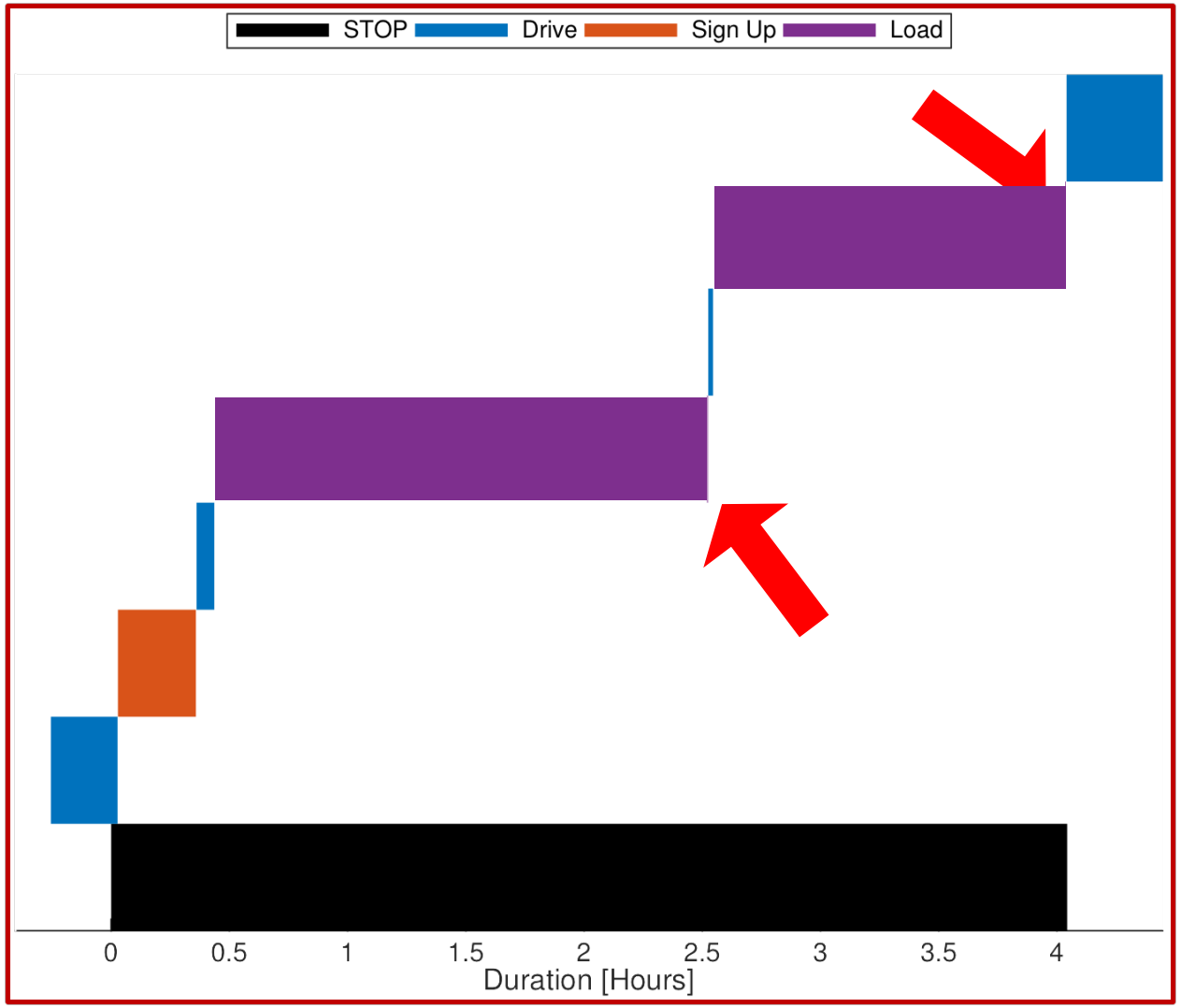
Partially human generated event logs leads to uncertainty related to the time at which events are logged.

The Problem

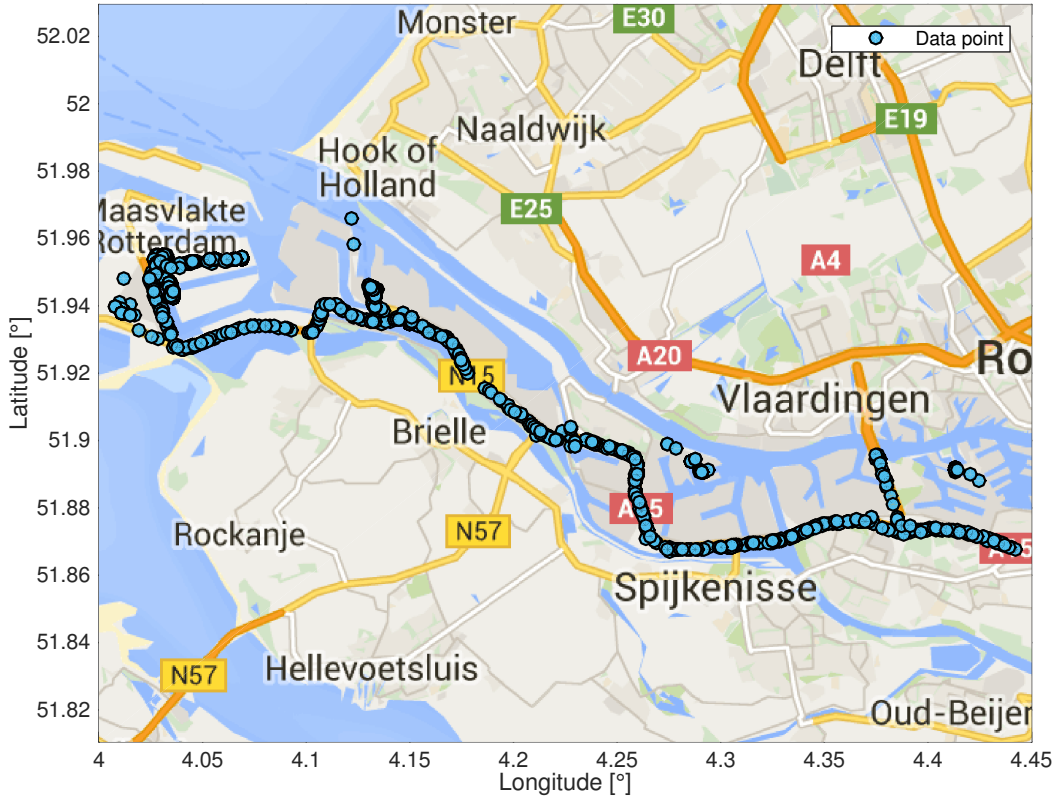

System
Real Occurrences
Event Log
User
Stochastic Characterization of Process Activities
Log time ≠ Real time

The Problem

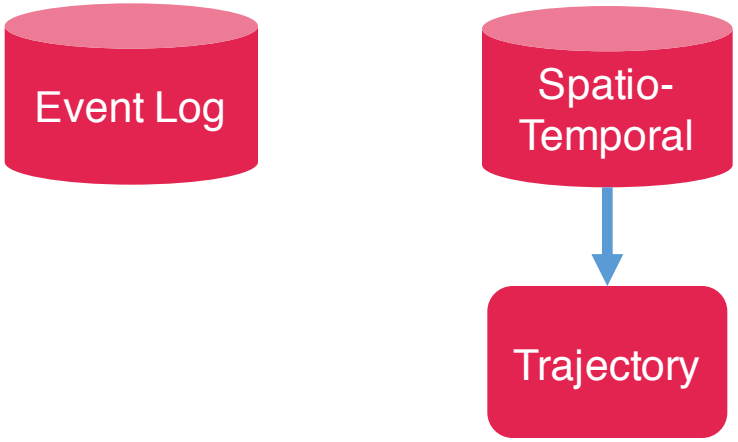

STOP
Drive
Sign Up
Rest
Unload
0
1
2
3
4
5
6
7
8
9
10
Duration [Hours]

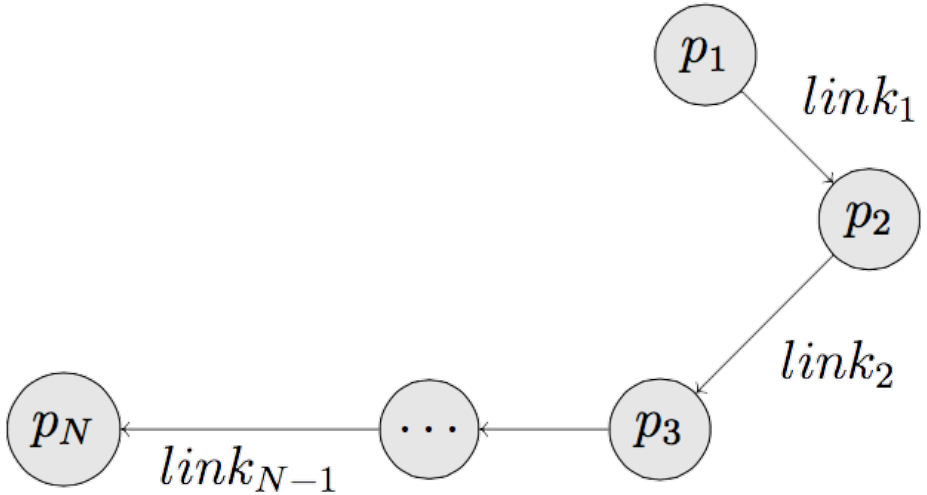

STOP
Drive
Sign Up
Load
0
0.5
1
1.5
2
2.5
3
3.5
4
Duration [Hours]

Framework

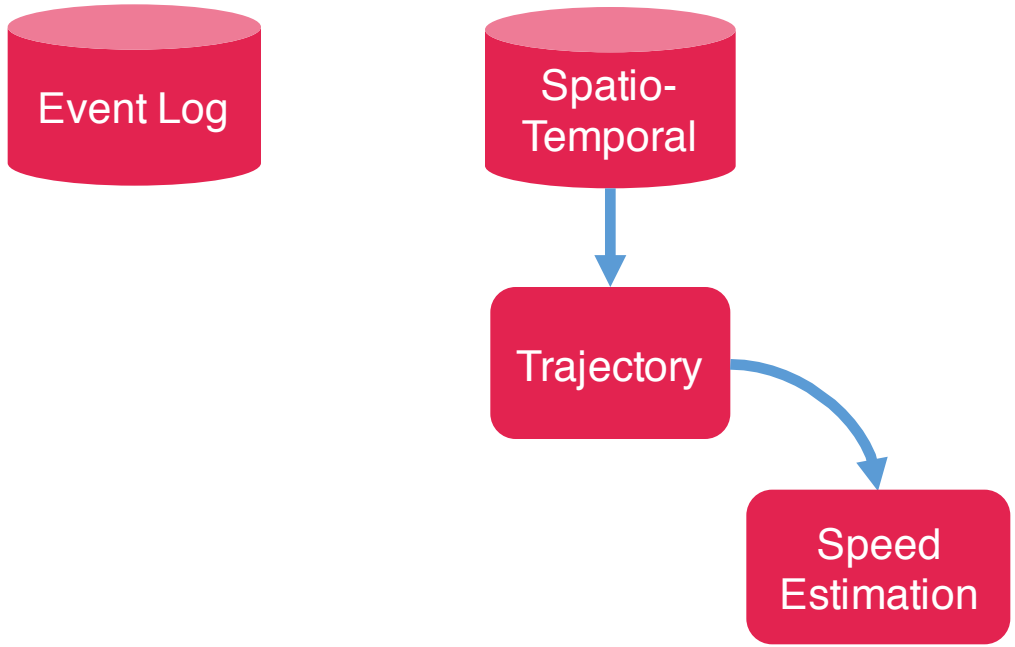


	Latitude	*Longitude*	*Time*
p_1:	ϕ_1,	λ_1,	t_1
p_2:	ϕ_2,	λ_2,	t_2
p_3:	ϕ_3,	λ_3,	t_3
	...		
p_N:	ϕ_N,	λ_N,	t_N





Framework

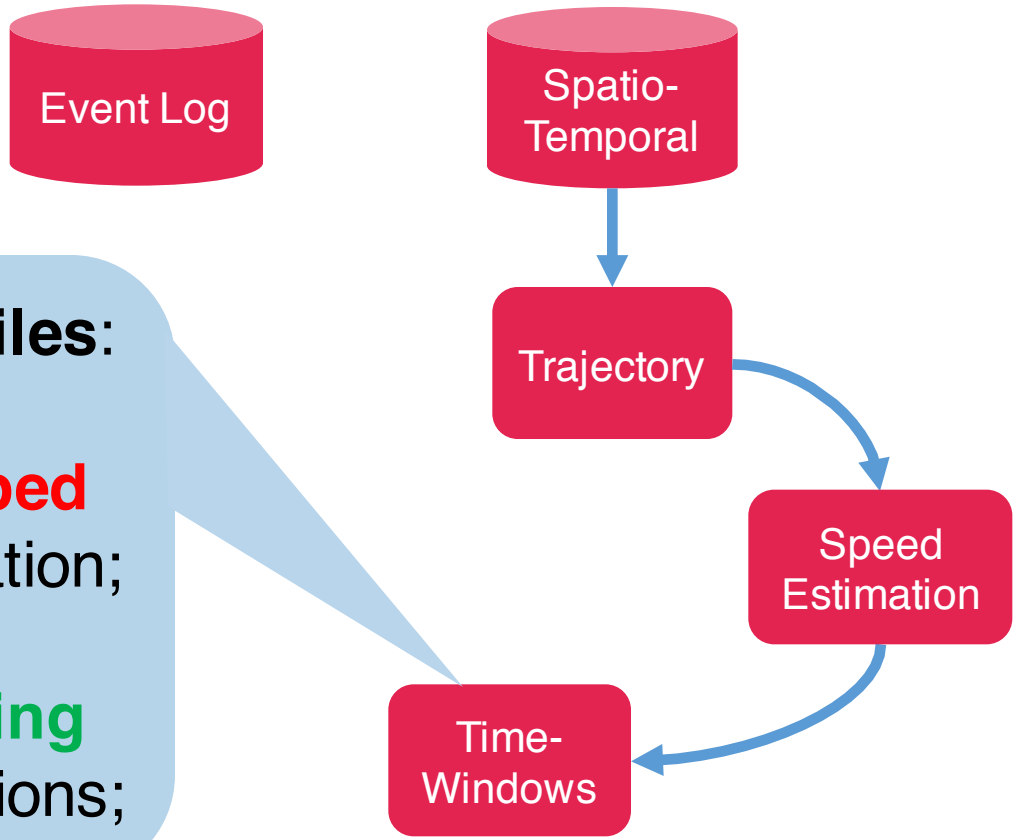


$$\bar{s}_j = \begin{cases} \bar{s}_j & if \quad \Delta t_j \geq \delta \\ \bar{s}_{j-1} & else \end{cases}$$

Framework

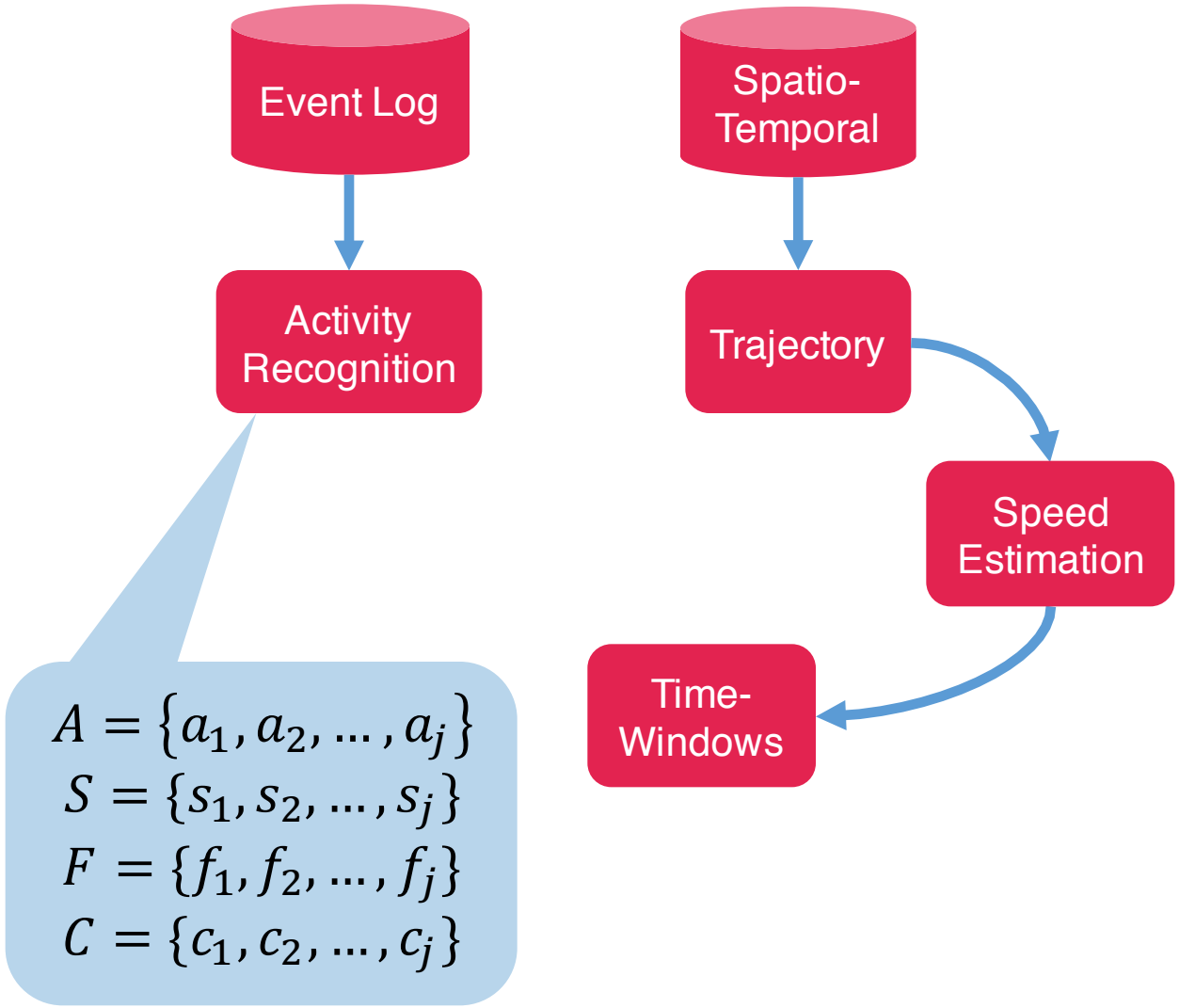
Time-Windows are defined based on **speed profiles**:

- Portions of the trajectory where truck was **stopped**
  → Used to estimate **load & unload** activity duration;
- Portions of the trajectory where truck was **moving**
  → Used to estimate **travel times** between locations;



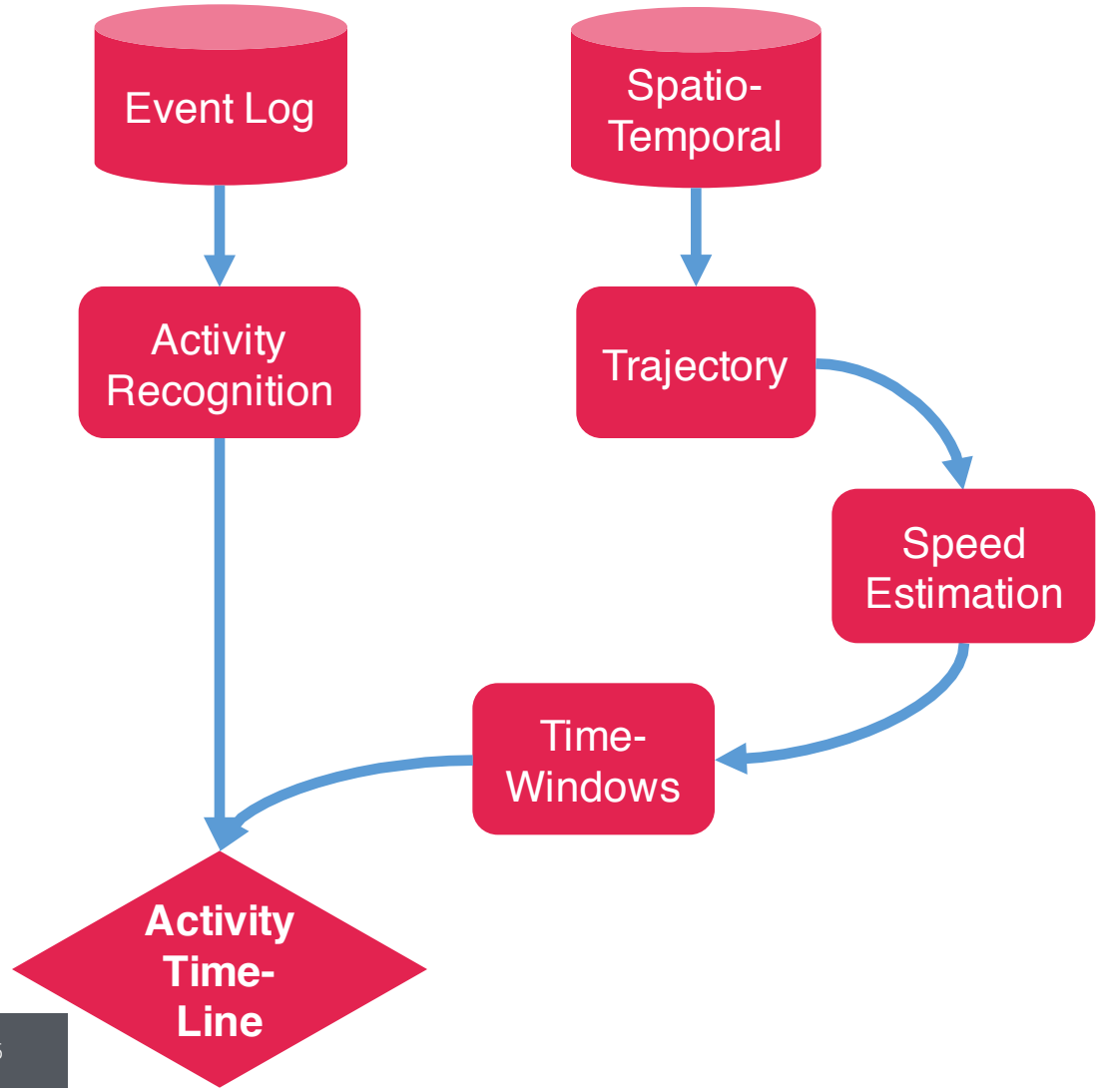
Framework

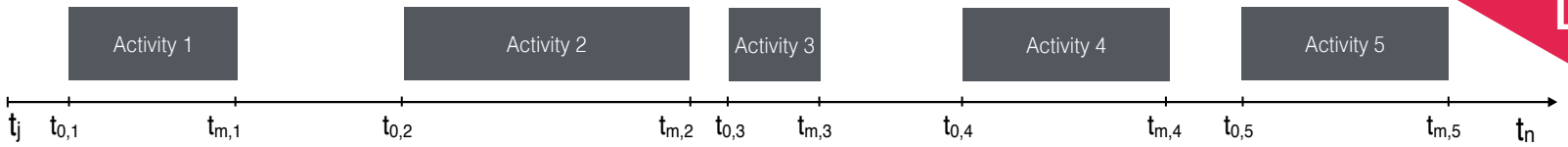
Truck ID	Date & Hour	Event	Activity
1141	2013-05-02 12:57:52	Navigation ETA update	-
1141	2013-05-02 13:57:55	Contact ON	-
1141	2013-05-02 14:57:58	Start of Break	Break
1141	2013-05-02 15:21:41	Cancellation of	Break
1141	2013-05-02 15:21:45	End of Break	Break
1141	2013-05-02 15:21:46	Start of	-
1141	2013-05-02 15:21:46	Cancellation of	-
1141	2013-05-02 15:22:21	Start of Drive	Driving
1141	2013-05-02 15:22:45	Driving times state event	Driving
1141	2013-05-02 15:22:45	Basic record	Driving
1141	2013-05-02 15:23:15	End of Drive	Driving
1141	2013-05-02 15:23:15	Start of	-
…	…	…	…
1141	2013-05-02 18:23:34	Task Busy	-
1141	2013-05-02 18:23:55	Cancellation of	-
1141	2013-05-02 18:24:56	Start of Unload	Unloading
1141	2013-05-02 18:26:29	Contact OFF	Unloading
1141	2013-05-02 18:27:31	Driving times state event	Unloading
1141	2013-05-02 18:27:46	Basic record	Unloading
1141	2013-05-02 18:28:52	Contact ON	Unloading
1141	2013-05-02 18:28:53	Task Finished	Unloading
1141	2013-05-02 18:30:46	End of Unload	Unloading



Framework

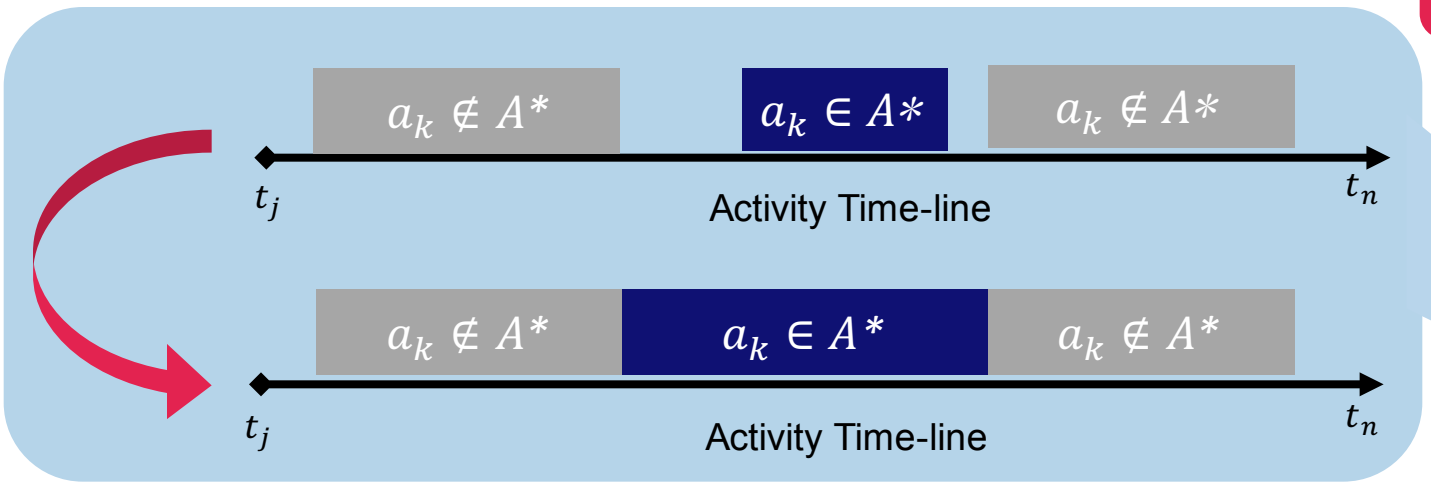
- **Time-windows** define the **upper** and **lower boundary** of the activity time-line and serves as **estimation interval**;
- Activities are assigned to the correspondent time-lines and the activity time-line is built;
- A subset of activities is defined:
  - A^* → **Set of human logged activities**
  - $A^* \subset A$

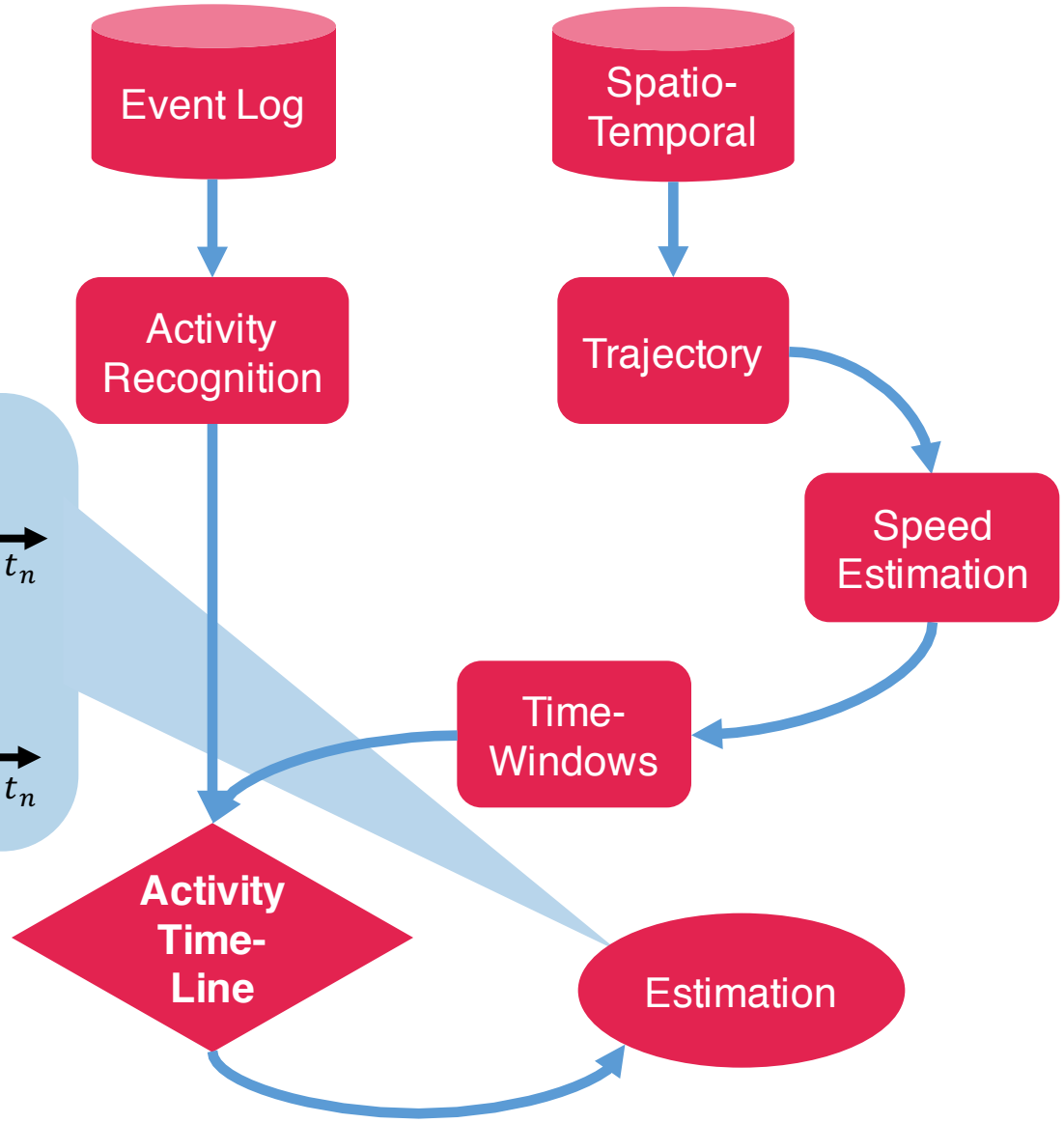




Framework

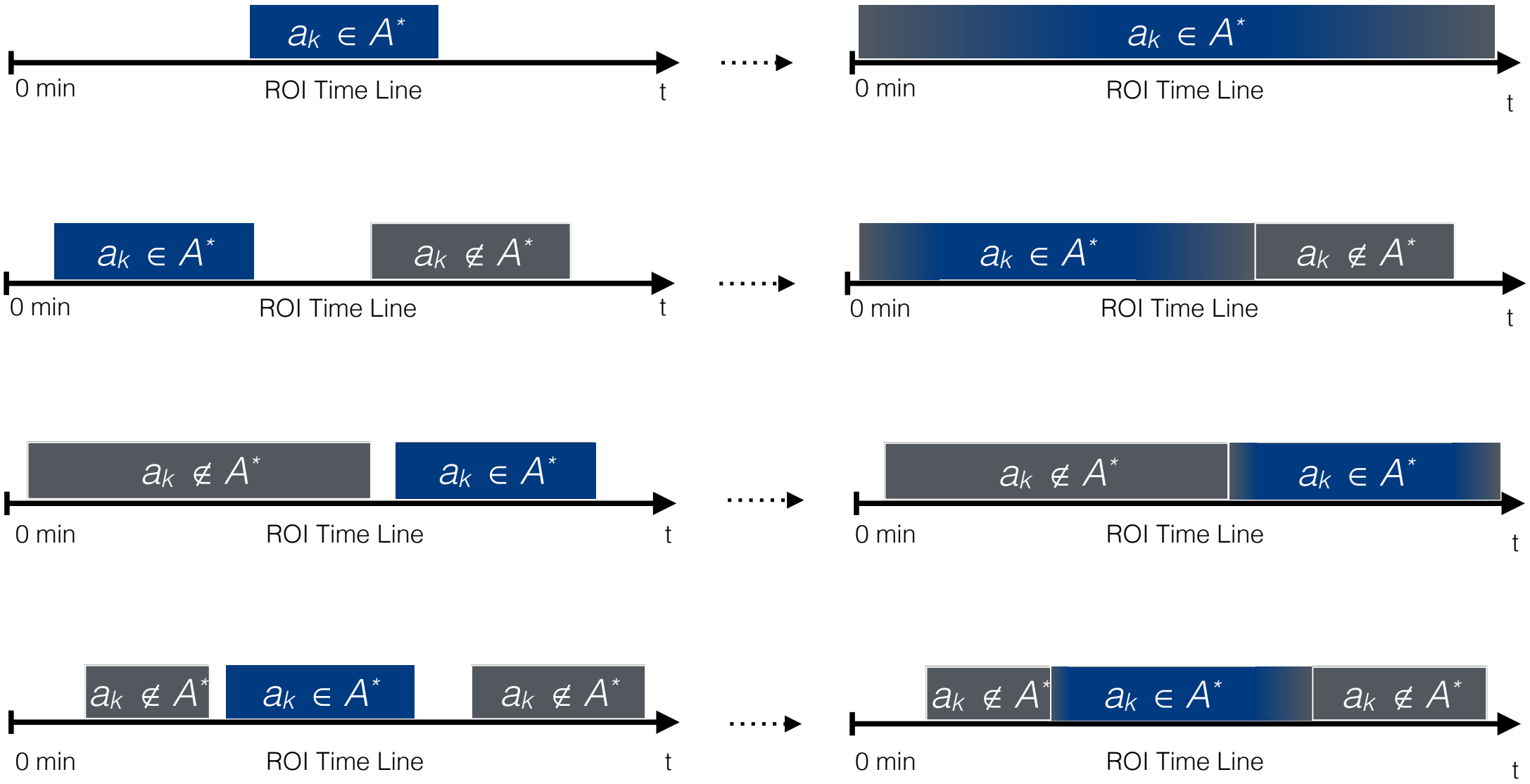
Activity durations are estimated based on the ***empty time*** in the **neighborhood** of such activities:



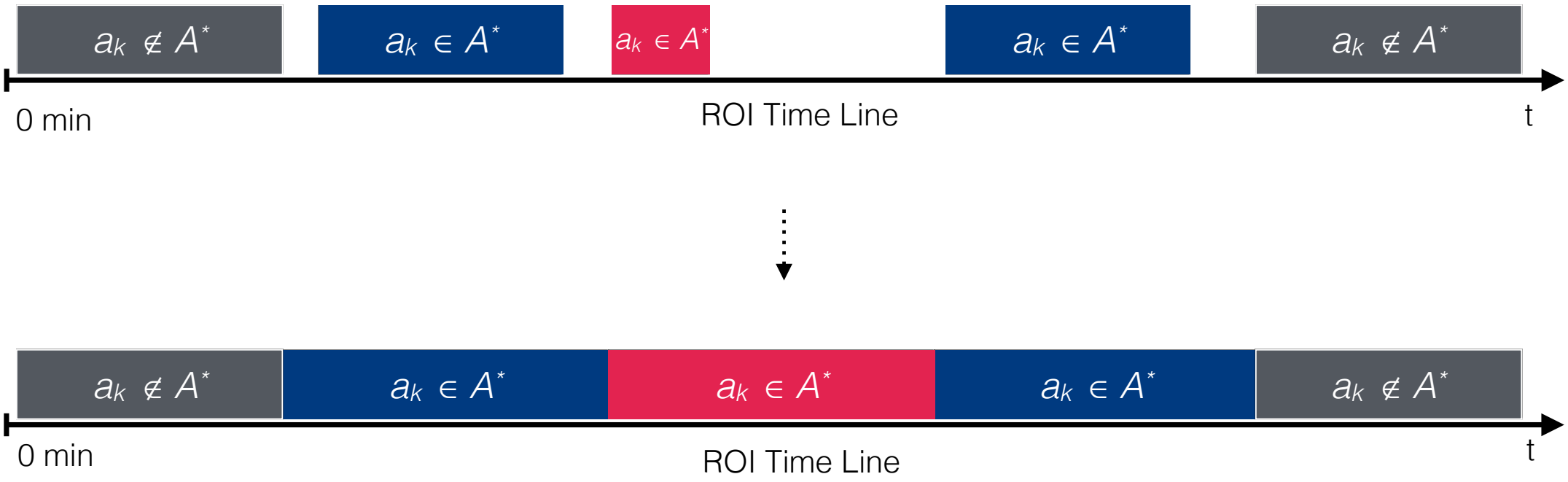


Activity Time-Line

Single activity case



Multiple activities case (ex.)



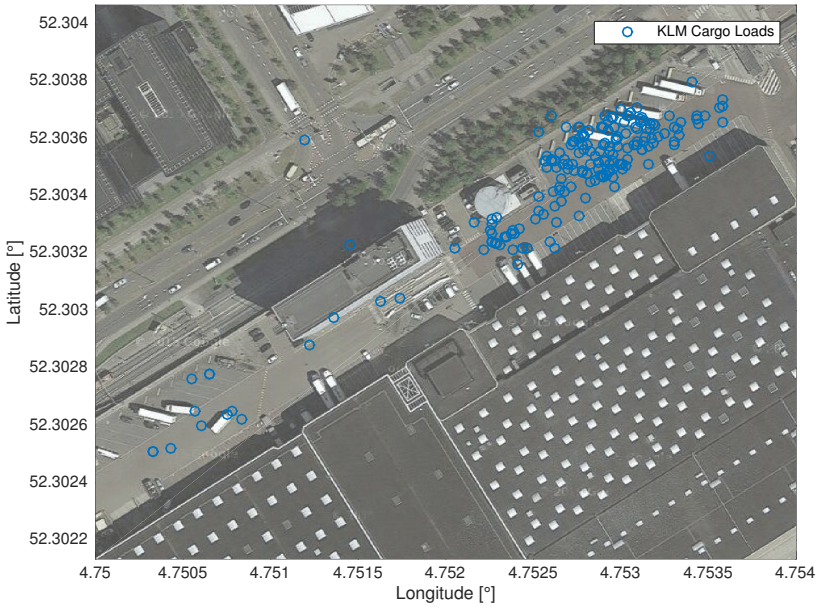
activity duration $\leq$ ε ⟶ Short activity

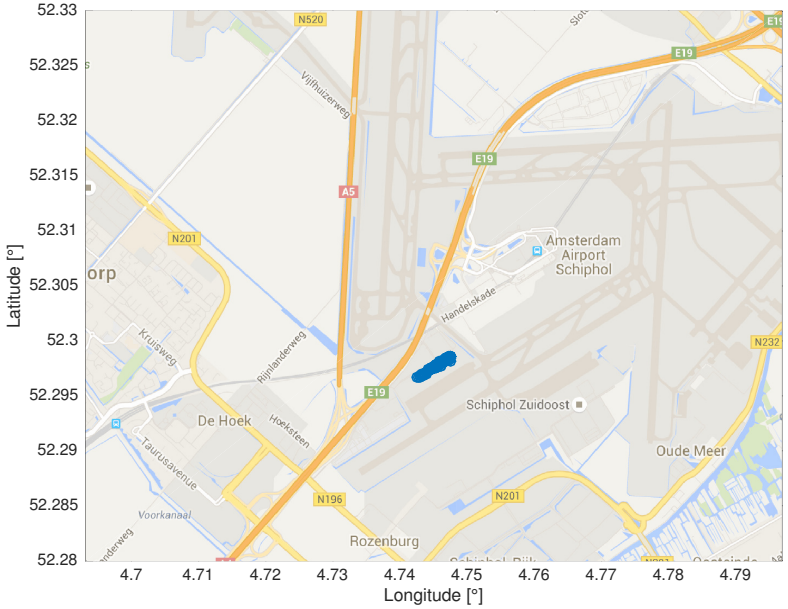
activity duration $>$ ε ⟶ Long activity

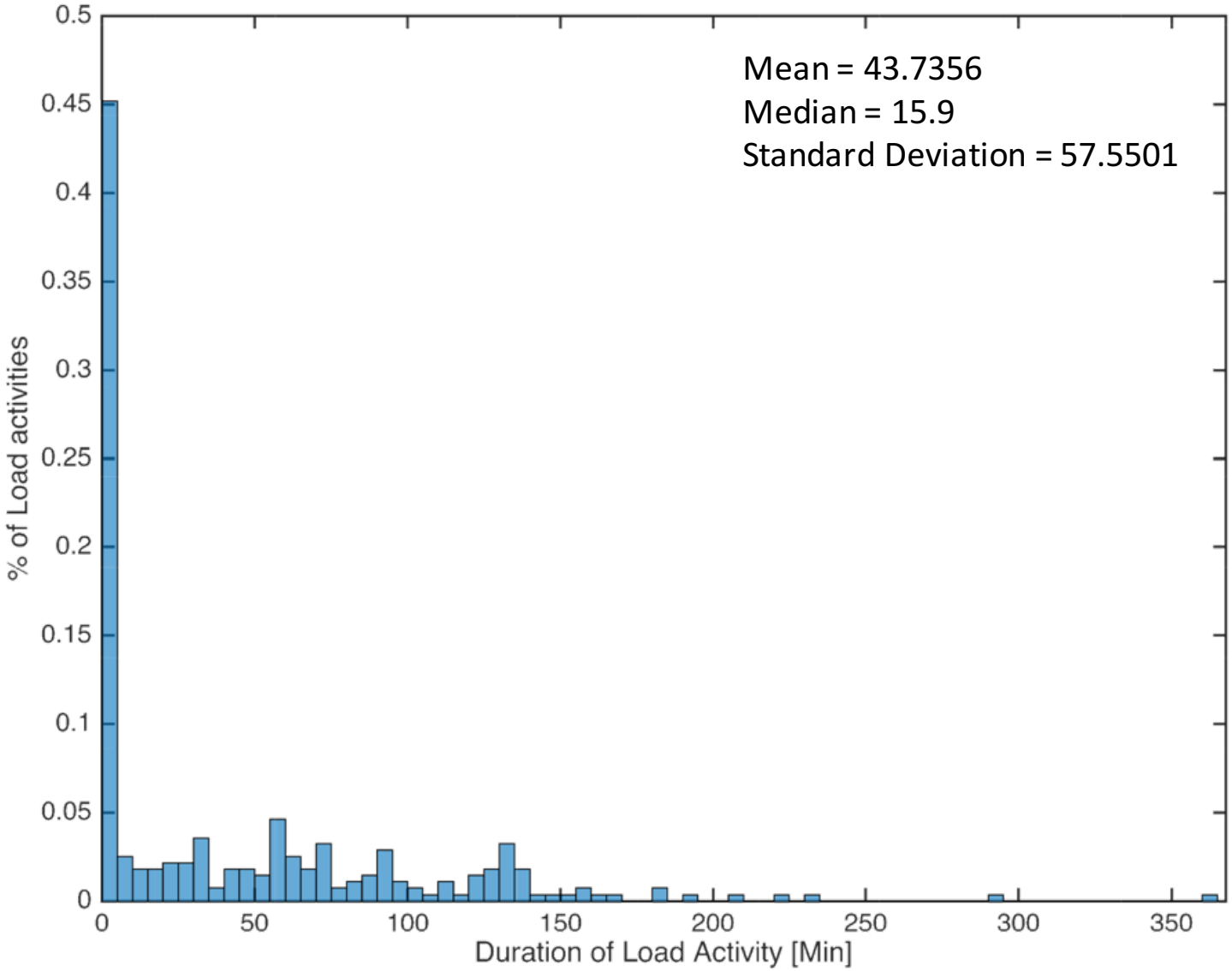
Customer Analysis

Latitude	Longitude	Truck ID	Activity	Estimated Duration
Φ_1	λ_1	TID_1	a_1	Δd_1
Φ_2	λ_2	TID_2	a_2	Δd_2
...	...	...	...	...
Φ_N	λ_N	TID_N	a_N	Δd_N

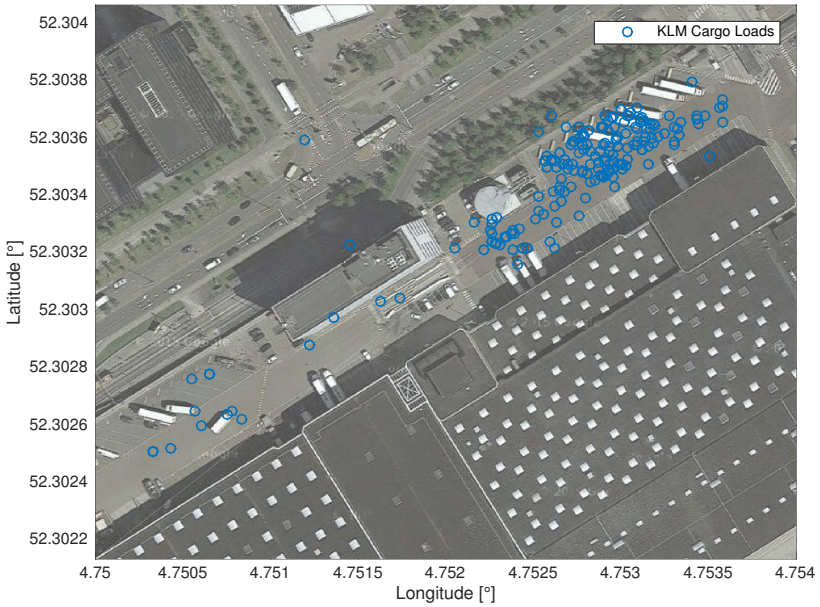
Original Load Activities Duration

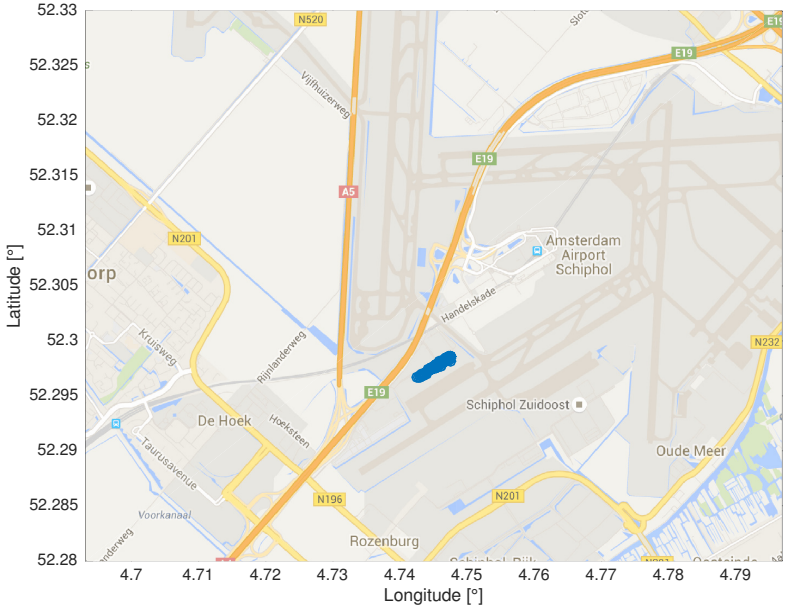


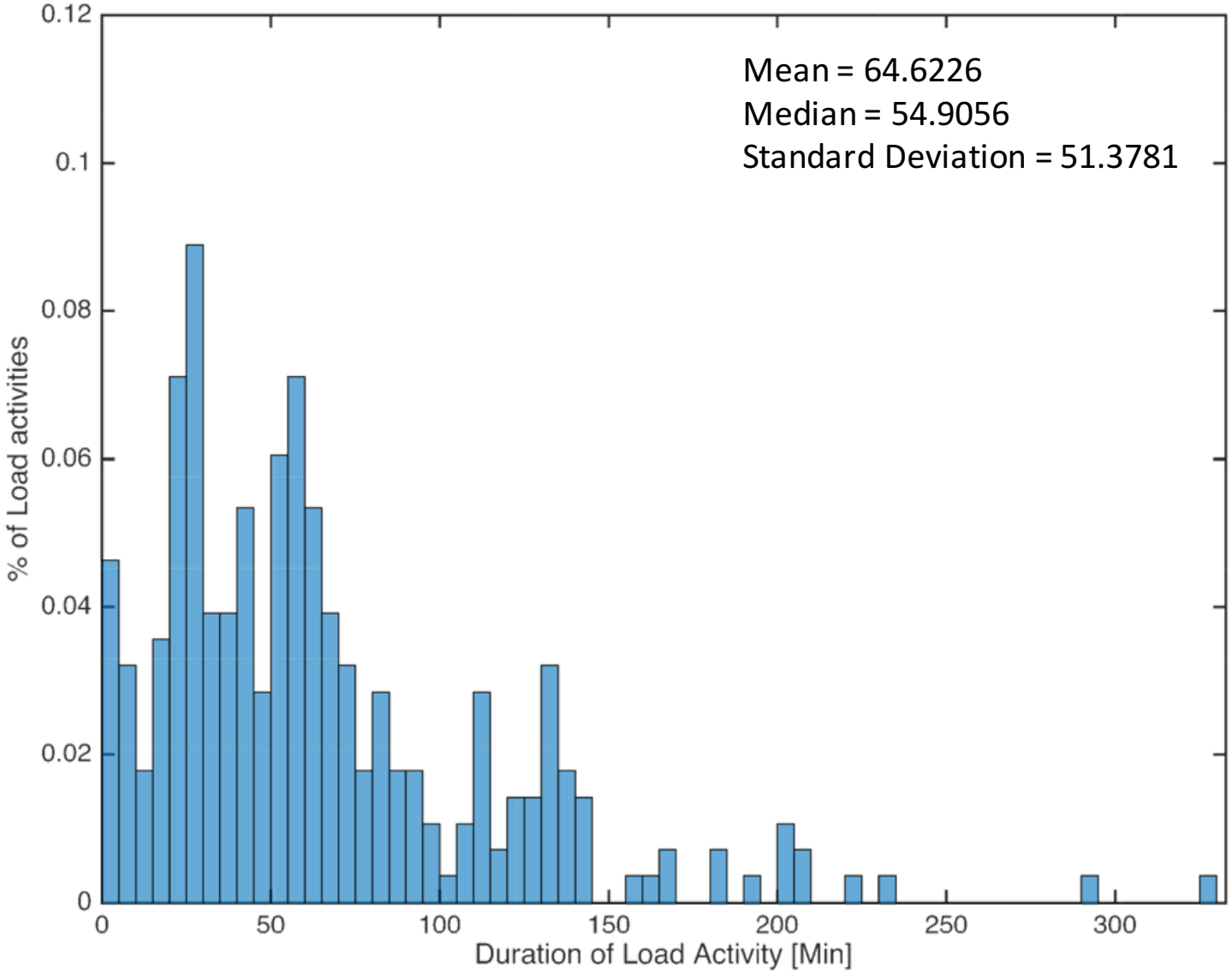




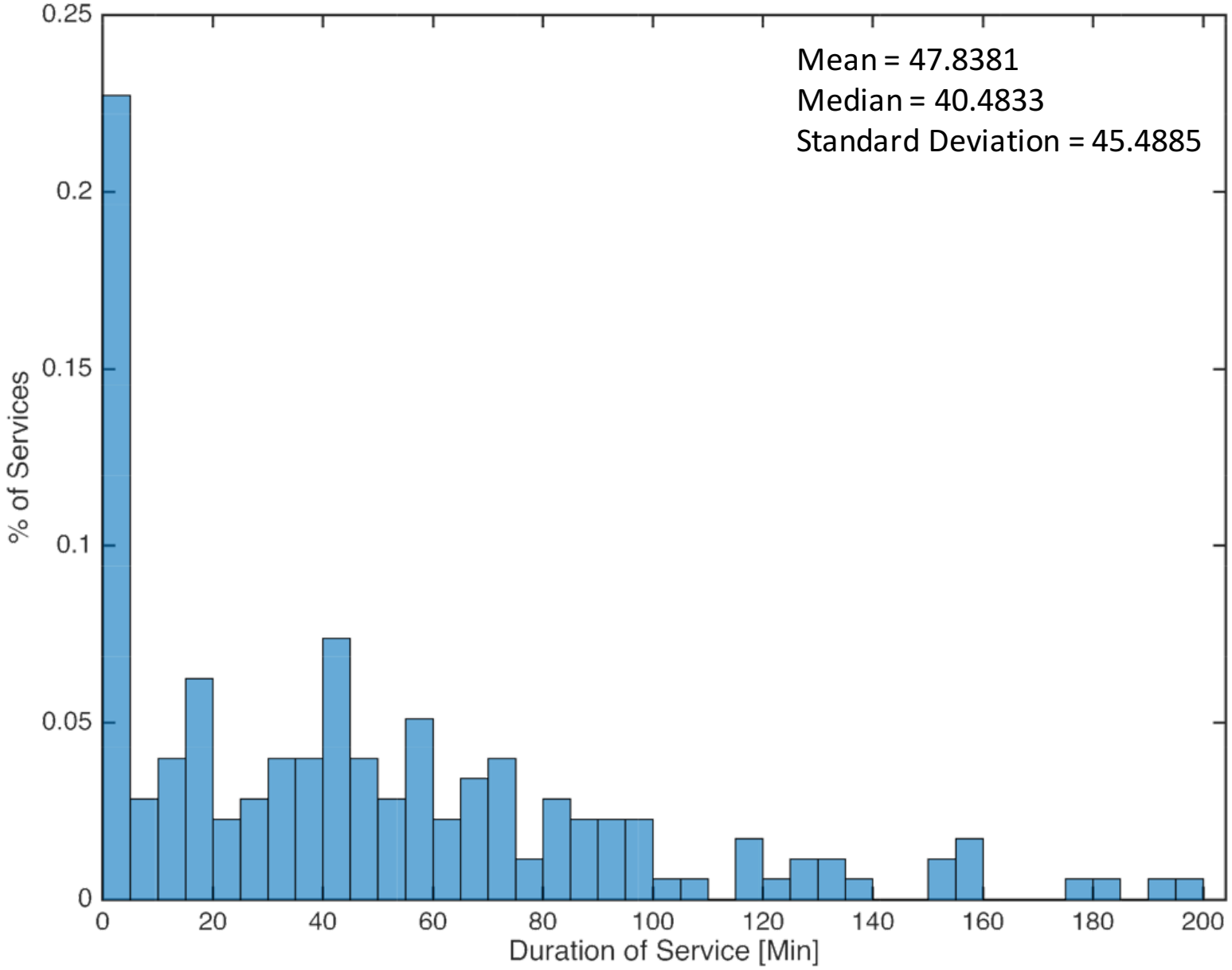
Estimated Load Activities Duration



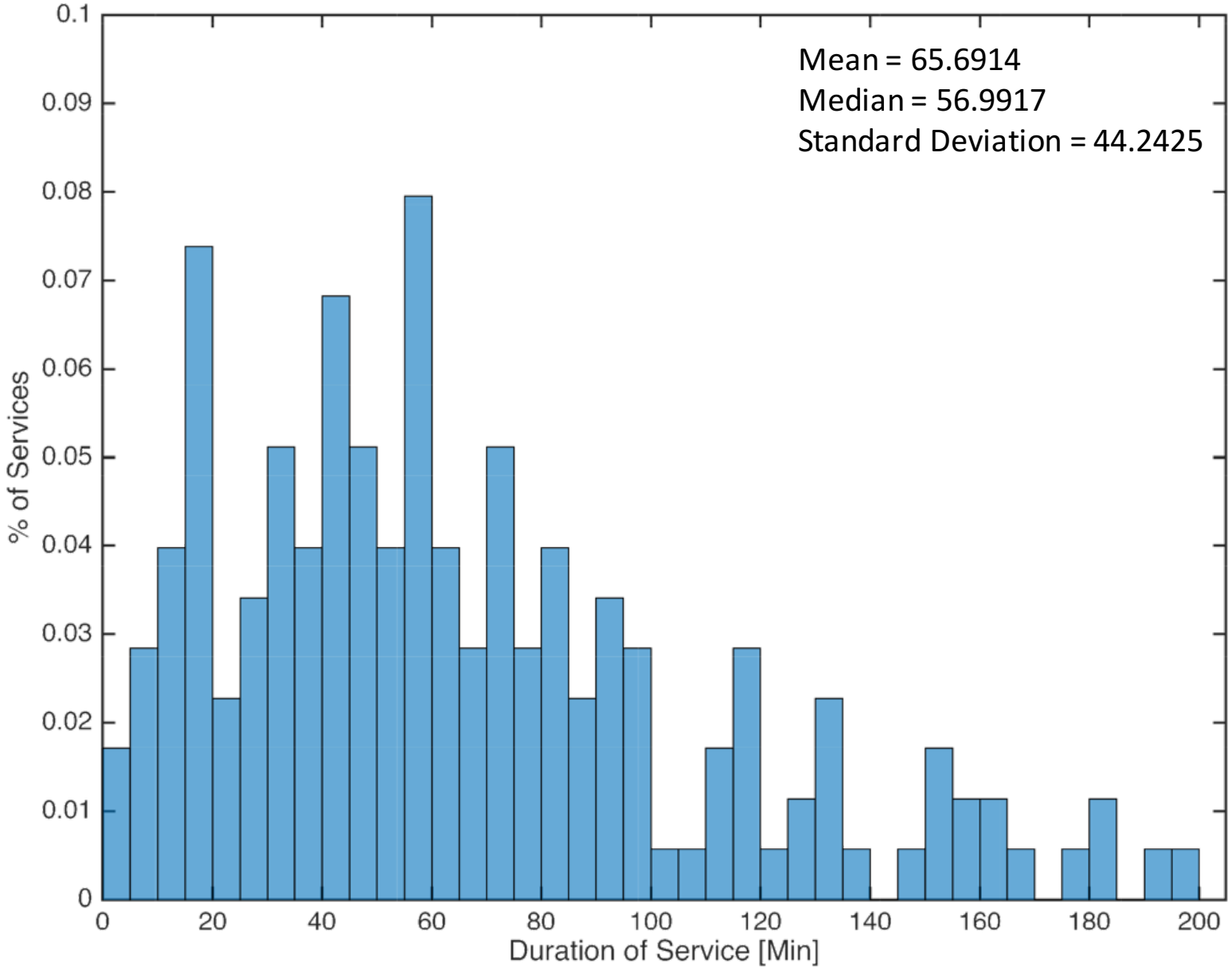




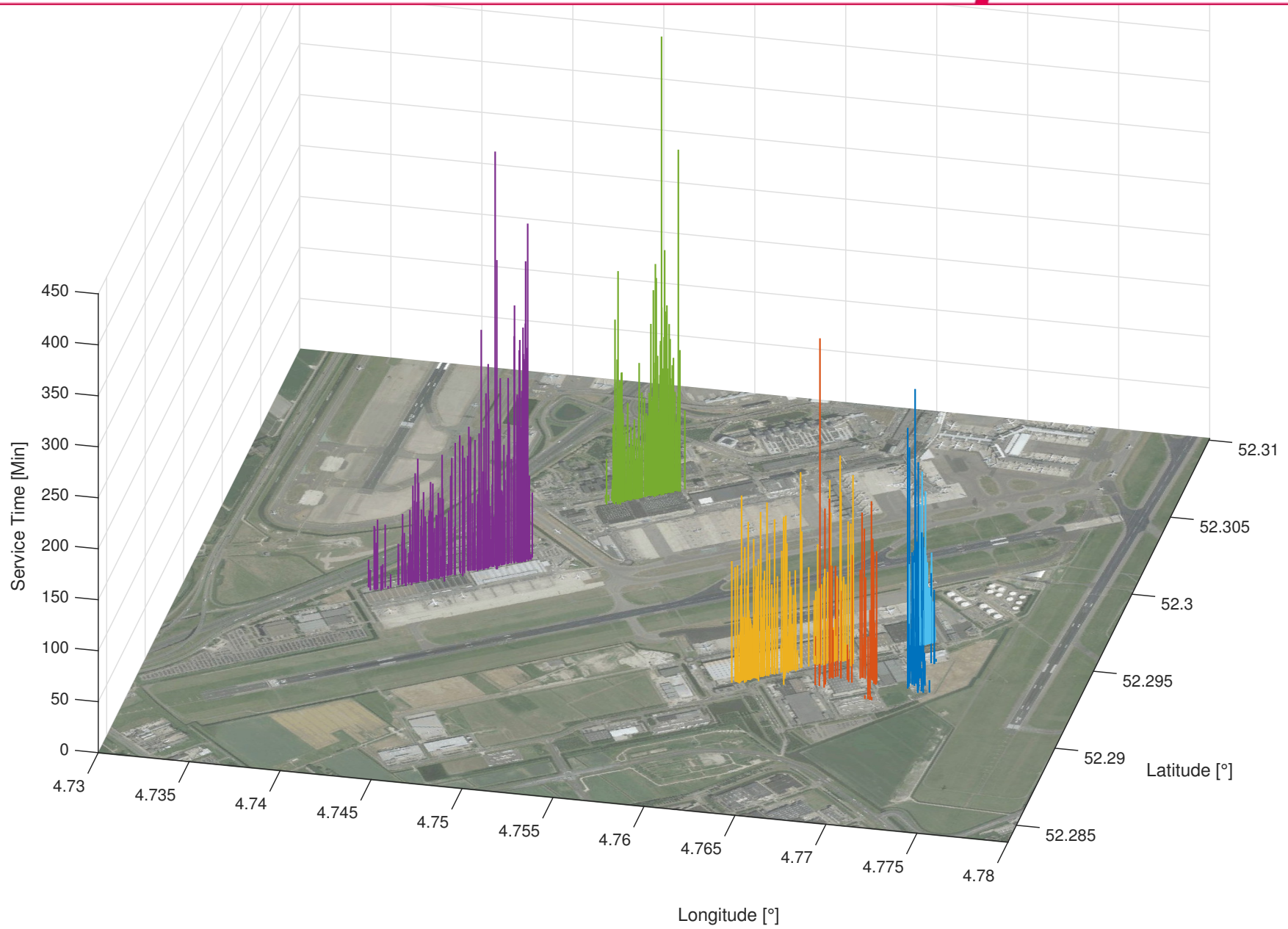
Original Service Duration



Estimated Service Duration



Service times at Amsterdam Airport

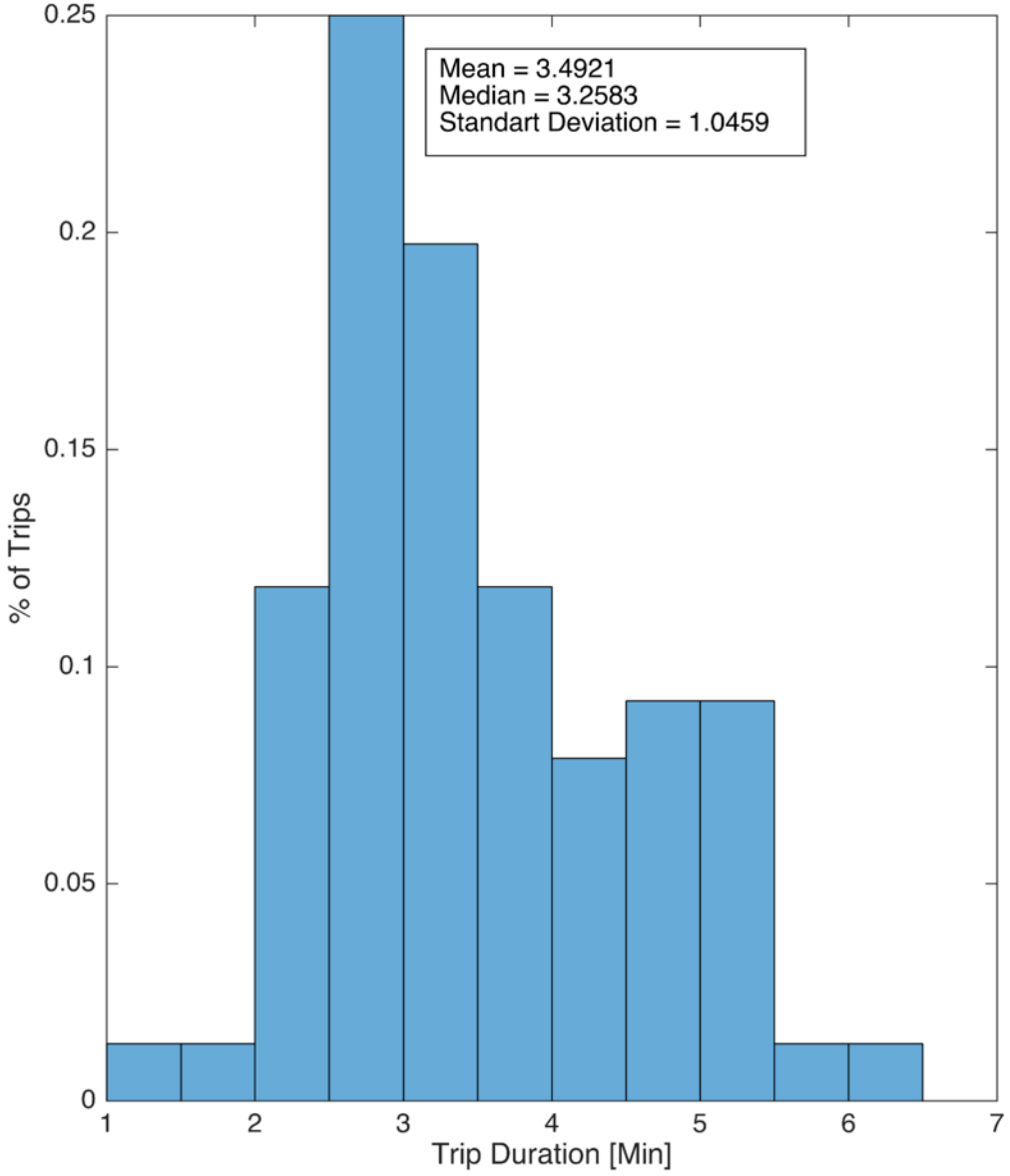


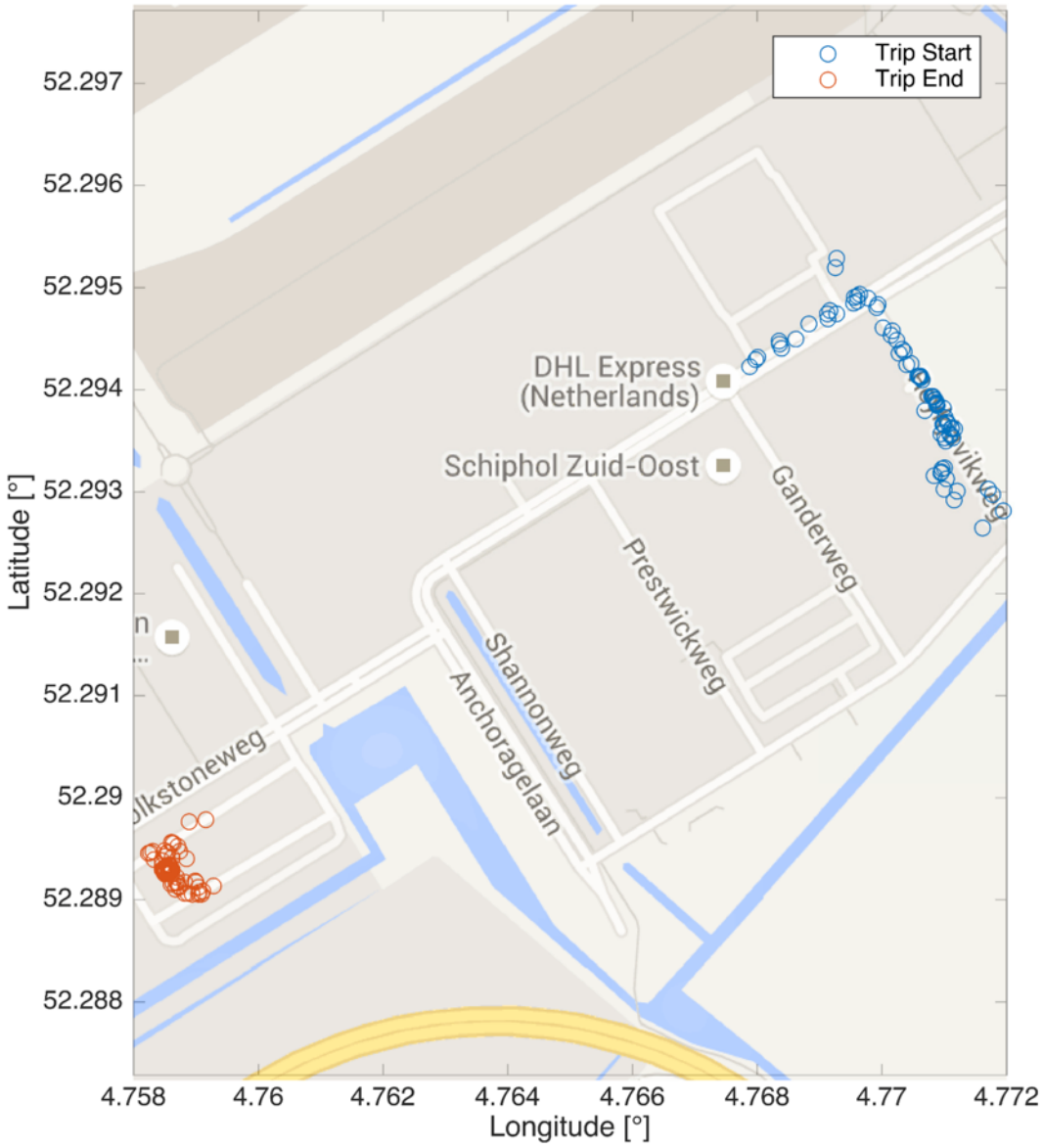
Travel Times

Analyse time-windows corresponding to moving portions of the trajectory

- Time-windows are now related to the moving portions of trajectories:
  - Start of TW = Start of Trip;
  - End of TW = End of Trip;
- To each trip is assigned a TRIP_ID;
- Preform clustering on “start” and “end” locations of trips and intersect clusters results;
- The intersection between a “start” and an “end” cluster gives the IDs from all similar trips.

Travel Times





Conclusions and Future Work

- The aggregation of different types of databases leads to the reduction of uncertainty when preforming stochastic characterization of process activities;
- Enable the prediction of service times at each customer as well as travel times between customers;
- Estimation constrains applied by time-windows and other activities create a well conditioned problem;

Future Work

- Use fuzzy systems for the classification of the trajectory links from spatio-temporal databases to achieve a higher level of detail in event logs. Use additional parameters such as the average link acceleration and link length;