Scaling Traffic Flow Properties of MATSim to correctly cooperate for Population Downsampling

Theresa-Maria Mersini, Theresa Ziemke

June 12th, 2025

Motivation

Method

Results

Questions

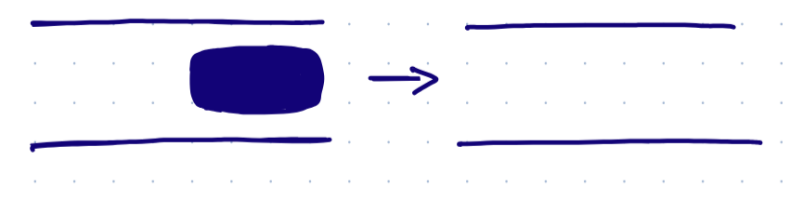
Motivation

Motivation

- Large-scale MATSim scenarios are usually run with a (down-)sampled population
- To model congestion propagation realistically traffic flow properties need to be scaled accordingly
- Currently this means adjusting the flow and storage capacity factors to the sample share, the stuck time parameter is usually not being adjusted.

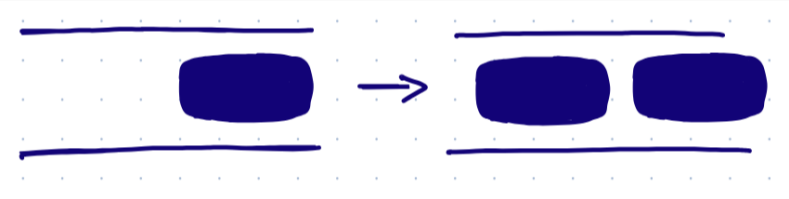
Capacities vs. stuck time in MATSim

- The flow capacity specifies the outflow capacity: number of vehicles that are allowed to leave the link per time step:

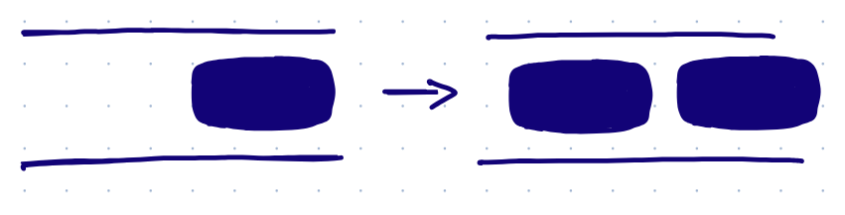


900 vehicles / hour $\widehat{=}$ 1 vehicle every 4 seconds

- The storage capacity is defined as the number of vehicles that can fit on a link.
- The stuck time parameter: A vehicle at the head of the queue gets pushed onto the next link after the stuck time has elapsed even if the storage capacity of the next link is exhausted.



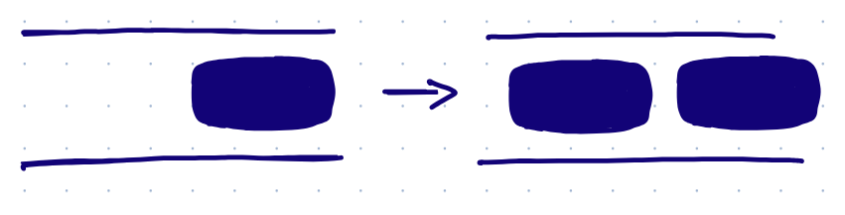
Capacities vs. stuck time in MATSim AND Scaling



A flow capacity of 900 vehicles per hour would mean in case of a sample-size of

- 100 % : 1 vehicle every 4.0 seconds
- 10 % : 1 vehicle every 40.0 seconds
- 1 %: 1 vehicle every 400.0 seconds
- vs. default stuck time parameter: 30.0 seconds

Flow-capacity vs. stuck time in MATSim AND Scaling



A flow capacity of 900 vehicles per hour would mean in case of a sample-size of

- 100 % : 1 vehicle every 4.0 seconds
- ~~10 % : 1 vehicle every 40.0 seconds~~
- ~~1 %: 1 vehicle every 400.0~~
- vs. default stuck time parameter: 30.0 seconds

Method

Method (1/3)

- Lausitz-Szenario v2024.2, calibrated 100 % plans (1.2 Mio Agents)
- Reduction to (persons) Agents with at least one car leg in their selected Plans: 305 200 Agents
- Mode choice is disabled
- Sample Shares: (0.01, 0.05, 0.1, 0.25, 0.5, 1.0)
- To evaluate variance within sample sizes: Ten 1, 5 and 10 percent samples
- Evaluation of ten different global seeds
- Last Iteration: 500

Method (2/3)

	fC-factor	sC-factor	stuck time	global Seed
case 1	sample-share	sample-share	30.0	default
case 2	sample-share	sample-share$^{0.75}$	30.0	default
case 3	sample-share	sample-share	$\frac{30.0}{\text{sample-share}}$	default
case 4	sample-share	sample-share$^{0.75}$	$\frac{30.0}{\text{sample-share}}$	default
case 5	sample-share	sample-share	30.0	rGs [1]

[1] rGs := (default, 3254, 2306, 6384, 4338, 6003, 5502, 9377, 5621, 9002), simulation run only for 1-pct sample nr. 1 and 5-pct sample nr. 1

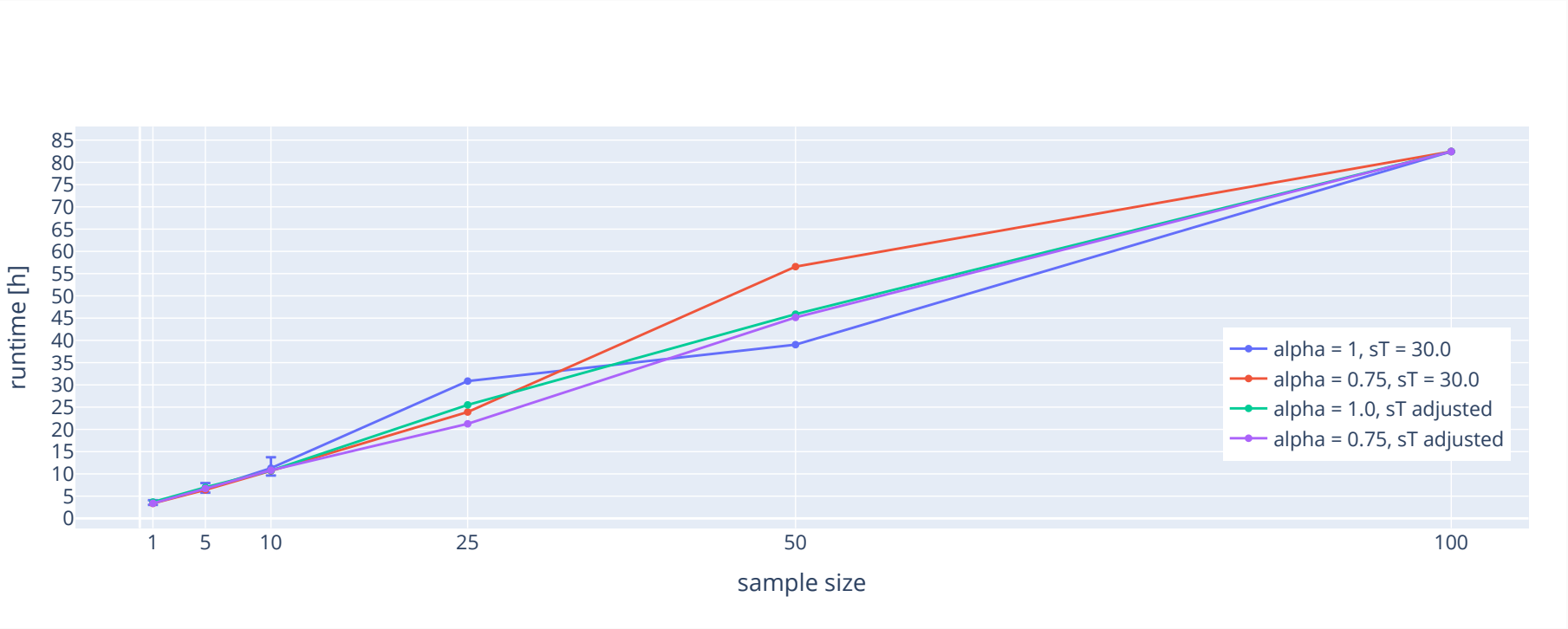
Method (3/3) - Evaluation

The following is being evaluated:

- Runtime
- Average travel time
- Travel time distribution
- MATSim Scores
- Average traveled distances
- Number of stuck time violations

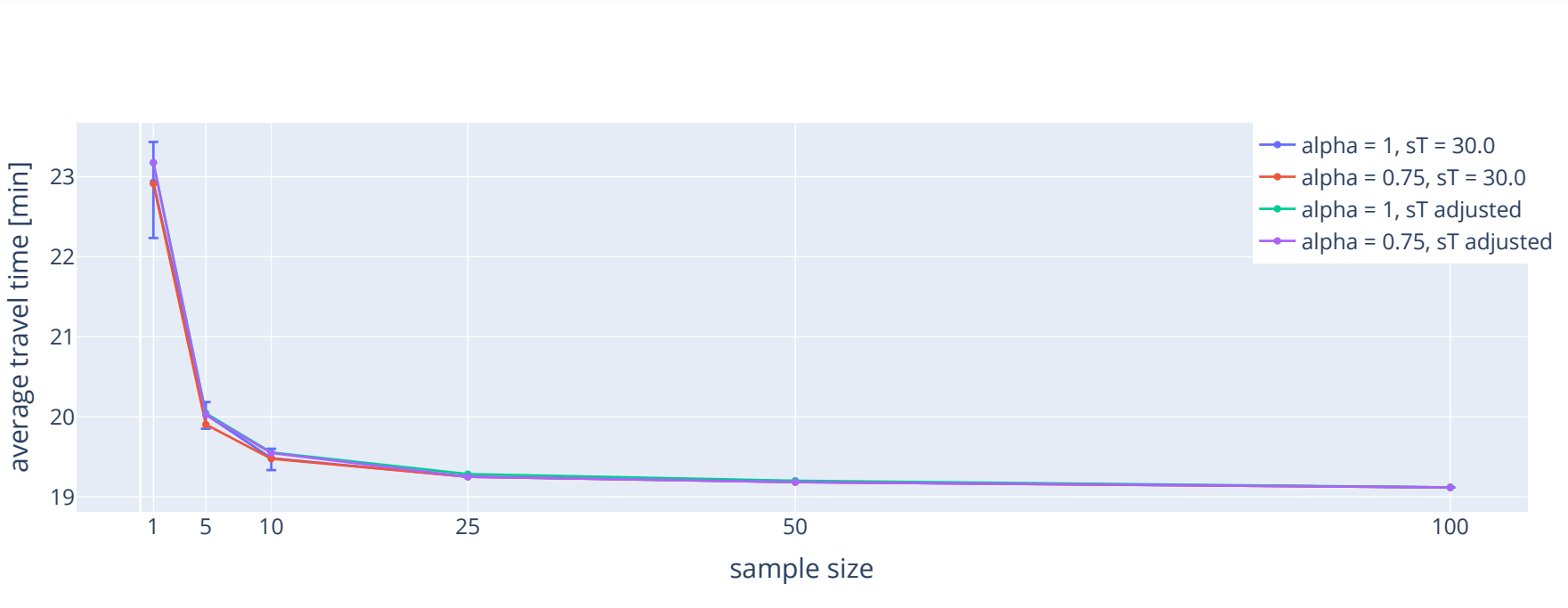
Results

Runtime - (Average) Duration



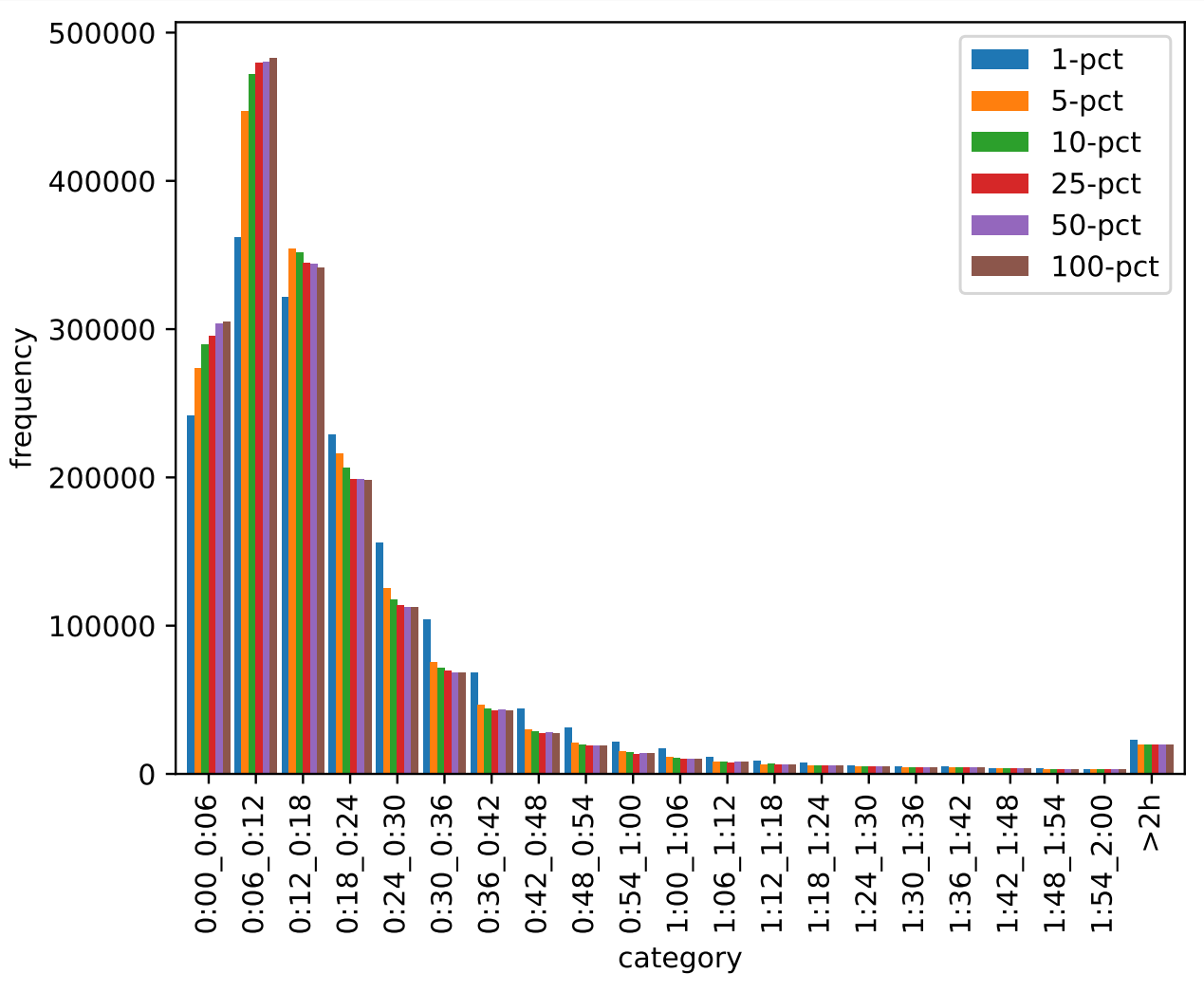
Average Duration of 1, 5 and 10 pct samples nr. 1 to 10, Duration of 25 to 100 pct samples, upper/ lower whiskers denote maximum/ minimum values respectively

Average of Average Travel Time



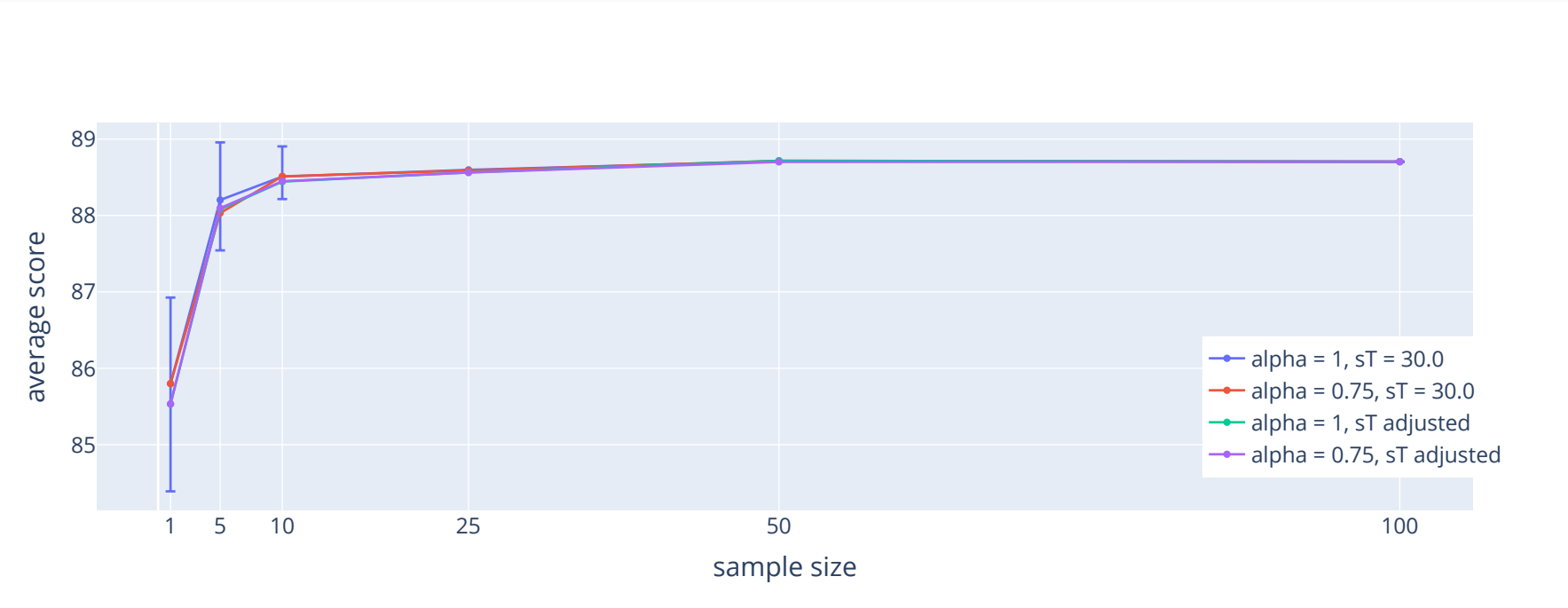
Avg. of avg. travel time of 1, 5 and 10 pct samples nr. 1 to 10, avg. travel time of 25 to 100 pct samples, upper/ lower whiskers denote maximum/ minimum values respectively

Travel Time Distribution



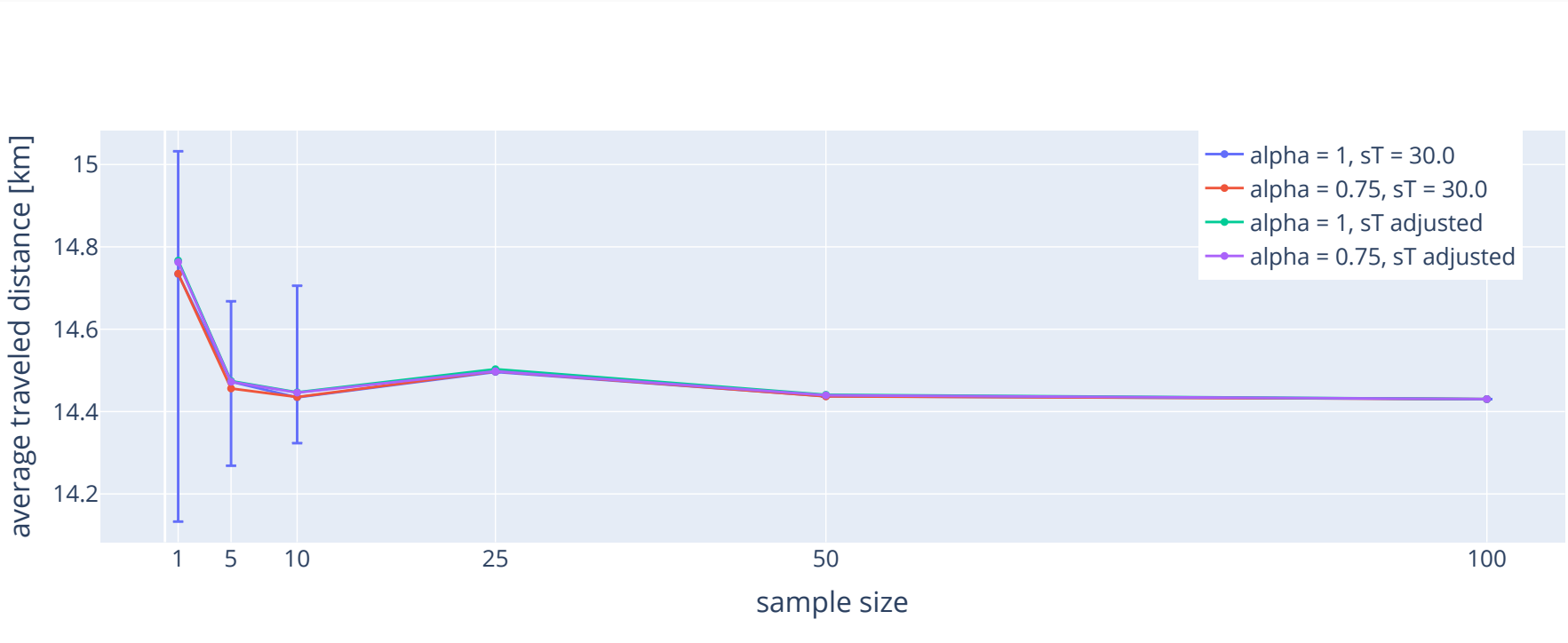
Average frequency by category of 1 to 10 pct samples nr. 1 to 10, frequency by category of sample sizes 25 to 100 pct in six minute categories up to 2 hours, in case of $\alpha = 1$ and stuck time $= 30.0$, frequencies are scaled to 100 %

Experienced Scores of the 500th iteration



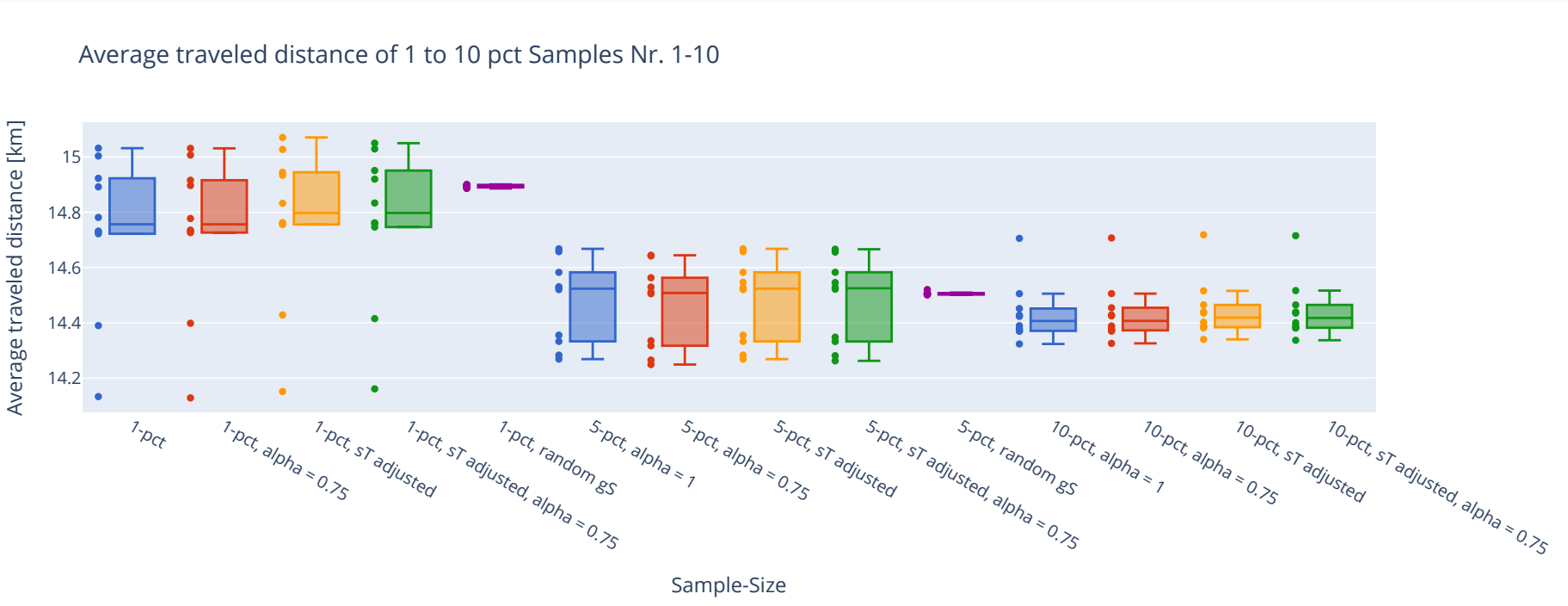
Avg. of avg. experienced scores of 1, 5 and 10 pct samples nr. 1 to 10, avg. exp score of 25 -100 pct, upper/ lower whiskers denote maximum/ minimum values respectively

Average of average traveled distance



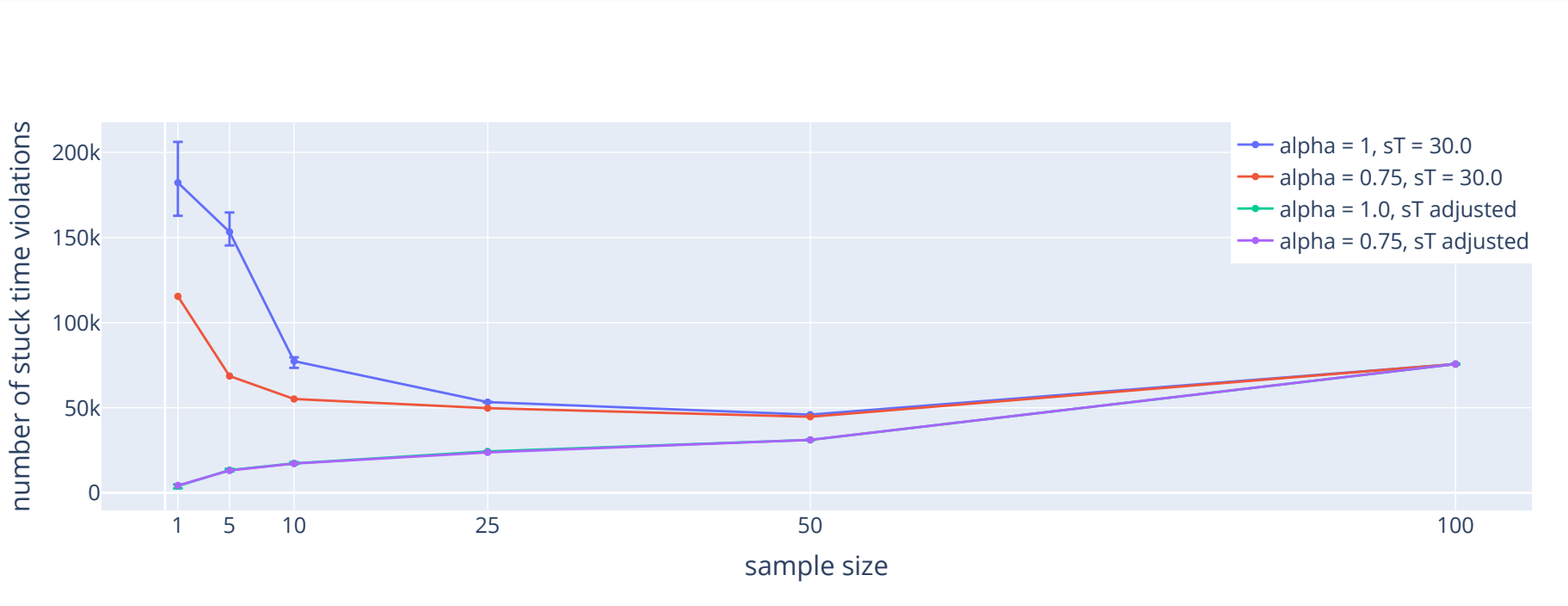
Avg. of avg. traveled distance of 1, 5 and 10 percent samples nr. 1 to 10, avg. traveled distance of sample sizes 25 too 100 pct, upper/ lower whiskers denote maximum/ minimum values respectively

Average traveled distance



Boxplot of avg. traveled distance of 1, 5 and 10 percent samples nr. 1 to 10

Number of Stuck Time Violations



Average number of stuck time violations of sample sizes 1 to 10 of samples nr. 1 to 10, number of stuck time violations of sample sizes 25 to 100, upper/ lower whiskers denote maximum/ minimum values respectively

Conclusion

In case of the Lausitz-Scenario with respect to the (reduced) 100 % sample with mainly car users:

- The results of the 1 and 5 pct samples deviate the most from the larger sample sizes
- Meaningful results require a sample size of at least 10 pct
- For smaller sample sizes: draw and simulate multiple samples and average results, if they vary: increase the sample size
- The influence of different global seeds on aggregated results is not that high
- The influence of the increased storage capacity is also not that high, it should be noted that traffic runs quite smoothly
- Scaling the stuck time should be considered, however $\frac{30.0}{\text{sample-share}}$ overscales it

Questions

Thanks, any Questions?