The MATSim Santiago open data model

Development and first applications

Leonardo Camus
Alejandro Tirachini
Benjamin Kickhöfer

Credits to:
Daniel Hosse
Kai Turner
Felix Zwick

Agenda

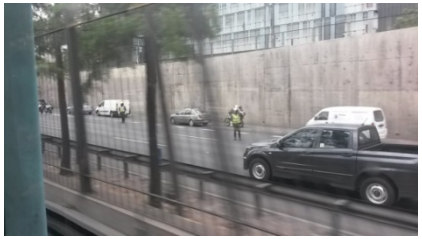
1. History of MATSim Santiago
2. Recent developments
3. Challenges and first applications
4. Summary and outlook

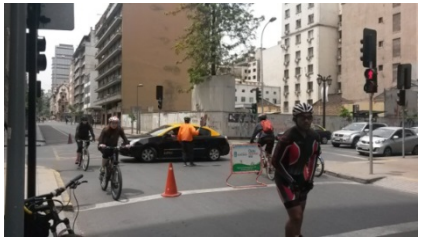
History of MATSim Santiago

How it all started

First impressions











Open data sources

1. Street network from OSM

2. Public transit supply data as GTFS

3. Travel diaries (and other stuff) from EOD2012
   - Exported from Microsoft Access into *.csv
   - Includes some freight traffic (and other information)

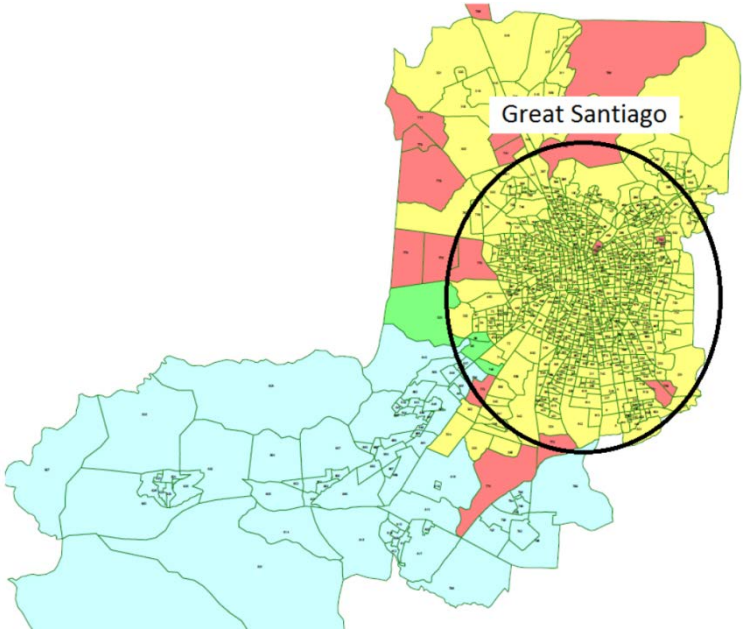
v1: first simulations

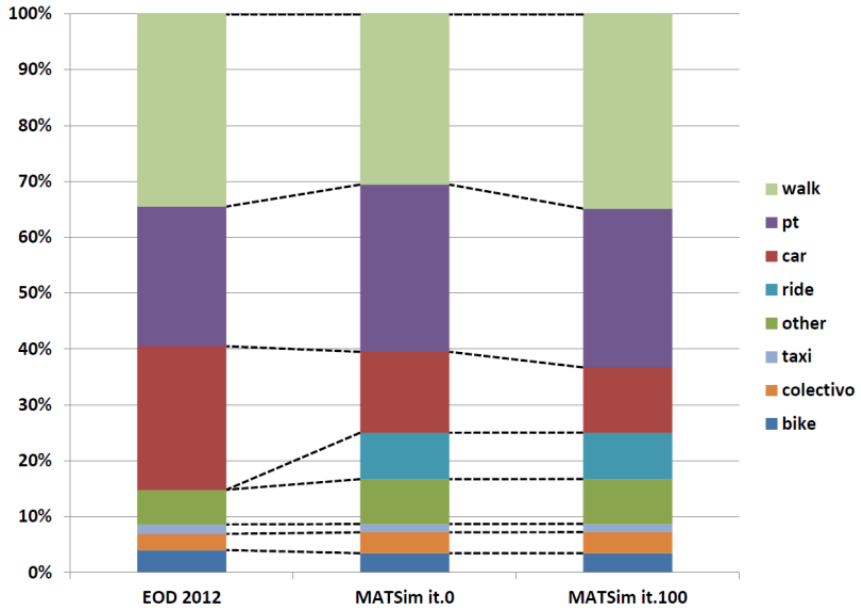
Raw data:

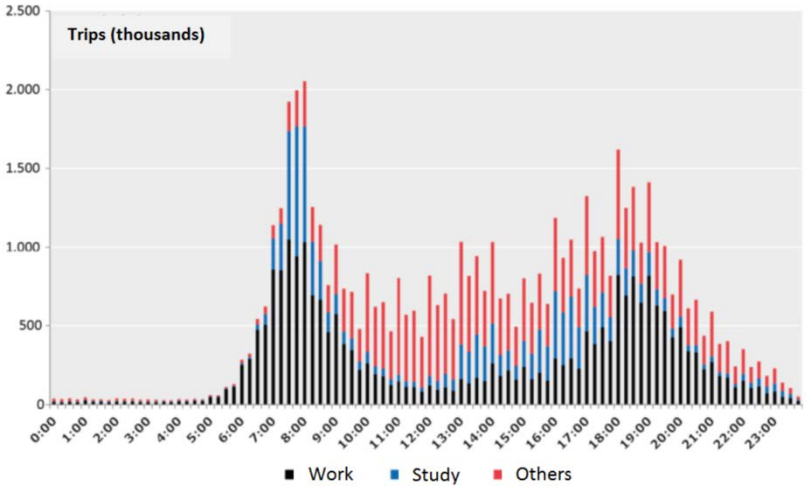
- 60’054 individuals (~1%)
- 113’591 trips

MATSim:

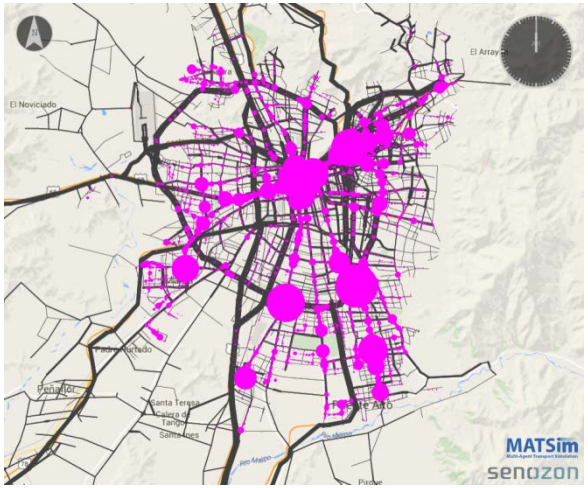
- 42’459 individuals
- 103’055 trips

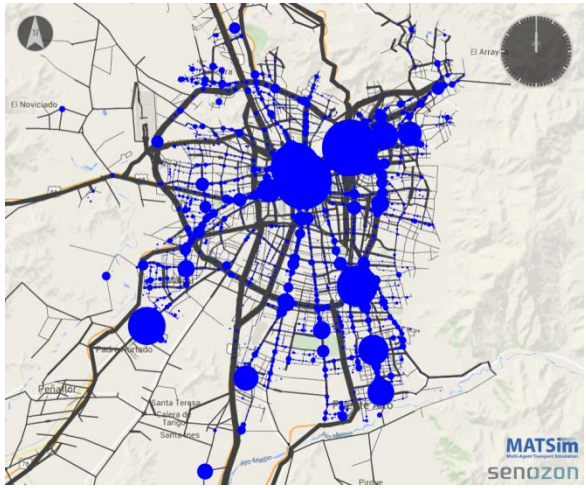


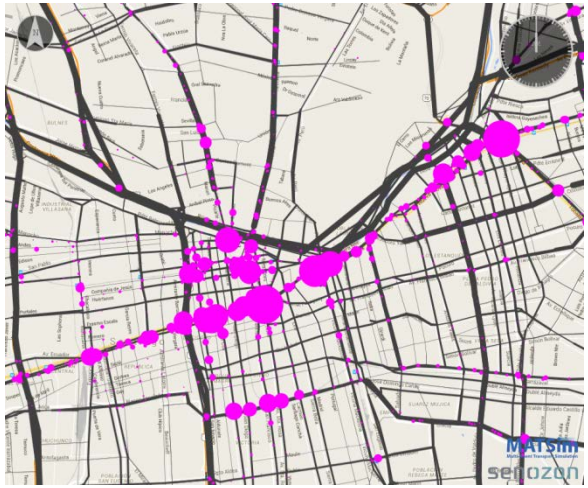


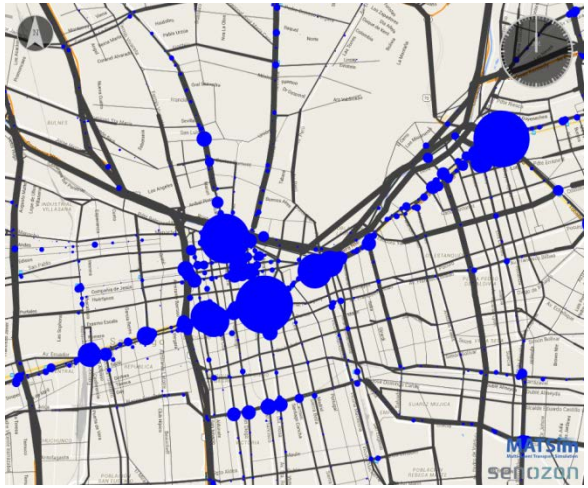


v1: integrating public transit









v1: behavioral parameters and calibration

Table 2: Behavioral parameters.

Parameter	Value	Unit
Source: Munizaga et al. (2008)		
Marginal utility of activity duration (β_{dur})	+ 4.014	$utils/h$
Marginal utility of traveling (β_{trav})	− 1.056	$utils/h$
Marginal utility of money (β_m)	+ 0.0023	$utils/CLP$
Approximate average $VTTS$	+ 2204.35	CLP/h
Results from calibration		
ASC car	+ 1.100	$utils$
ASC PT	− 0.883	$utils$
ASC walk	+ 0.000	$utils$

$$C_{mode,n+1} = C_{mode,n} - log\left(\frac{p_{mode,n}}{p_{mode,it.0}}\right)$$

Table 3: Modal split: comparison between input data and MATSim synthetic population.

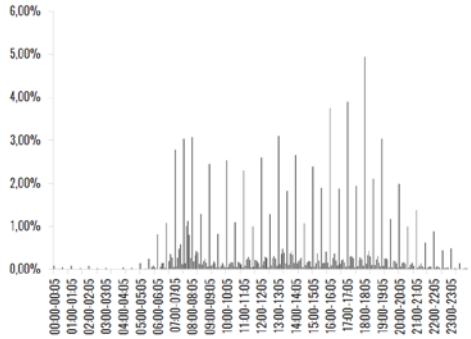
Mode	Sectra (2014)	Raw data	MATSim it.0	MATSim it.200
Bike	4.00	3.41	3.41	3.41
Car	25.70	23.27	14.40	14.28
Colectivo	2.90	3.11	3.73	3.73
Other	6.20	7.74	7.98	7.98
PT	25.00	31.50	29.88	28.19
Ride	in "Car"	in "Car"	8.26	8.26
Taxi	1.70	1.46	1.47	1.47
Train	in "Other"	in "Other"	0.03	0.03
Walk	34.50	29.78	30.83	32.64

Recent developments

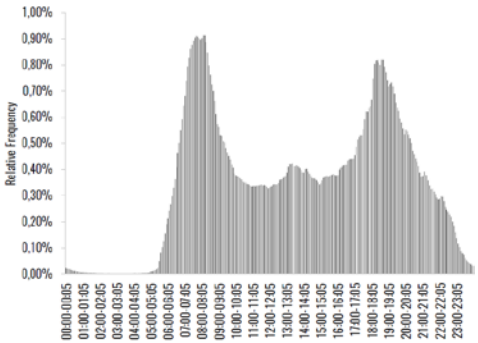
Leo (v2): overview

- Remove agents that were interviewed in summer vacation time and on weekends
- Use expansion factors to create a 1% and a 10% sample
- Add time-dependent tolls to tollways
- Randomize activity end times using smartcard data
- Randomize activity locations using land use data

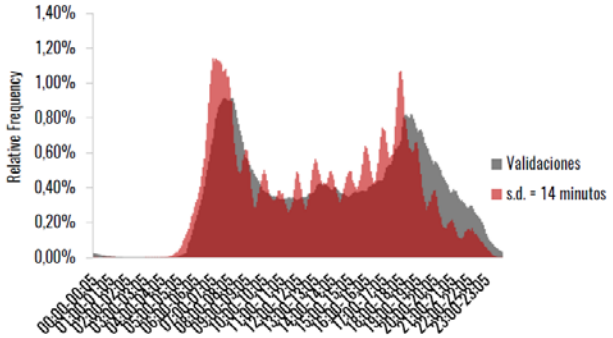
Leo (v2): randomize activity end times using smartcard data



Origin-Destination Survey (2012)

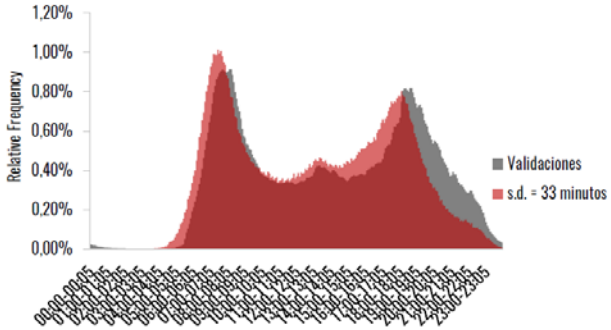


Smartcard data



min MAPE

σ^* = 14 mins.

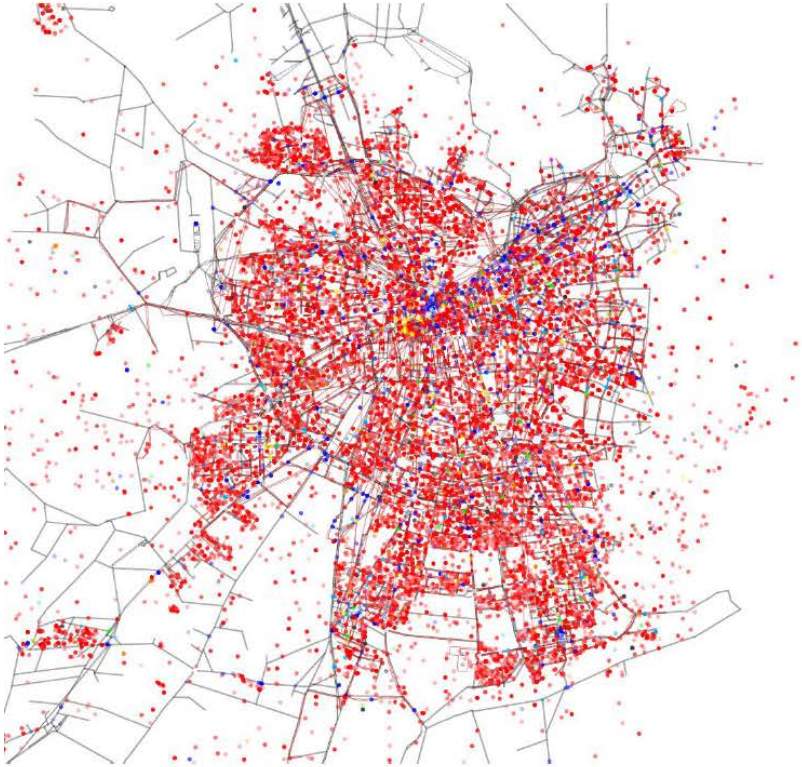


min χ^2

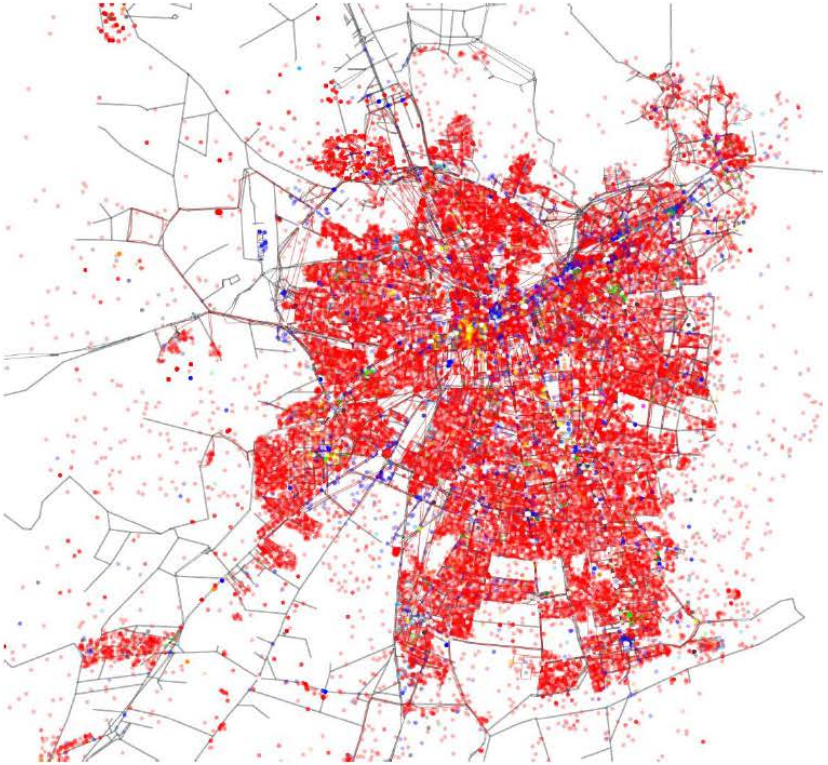
σ^* = 33 mins.

Leo (v2): randomize activity locations using land use data

Original locations 1% – 20:00



Randomized locations 1% – 20:00



Felix (v3): implementing colectivos

Figure 2: MATSim network and multiline strings

Felix (v3): travel times, slack times, fares

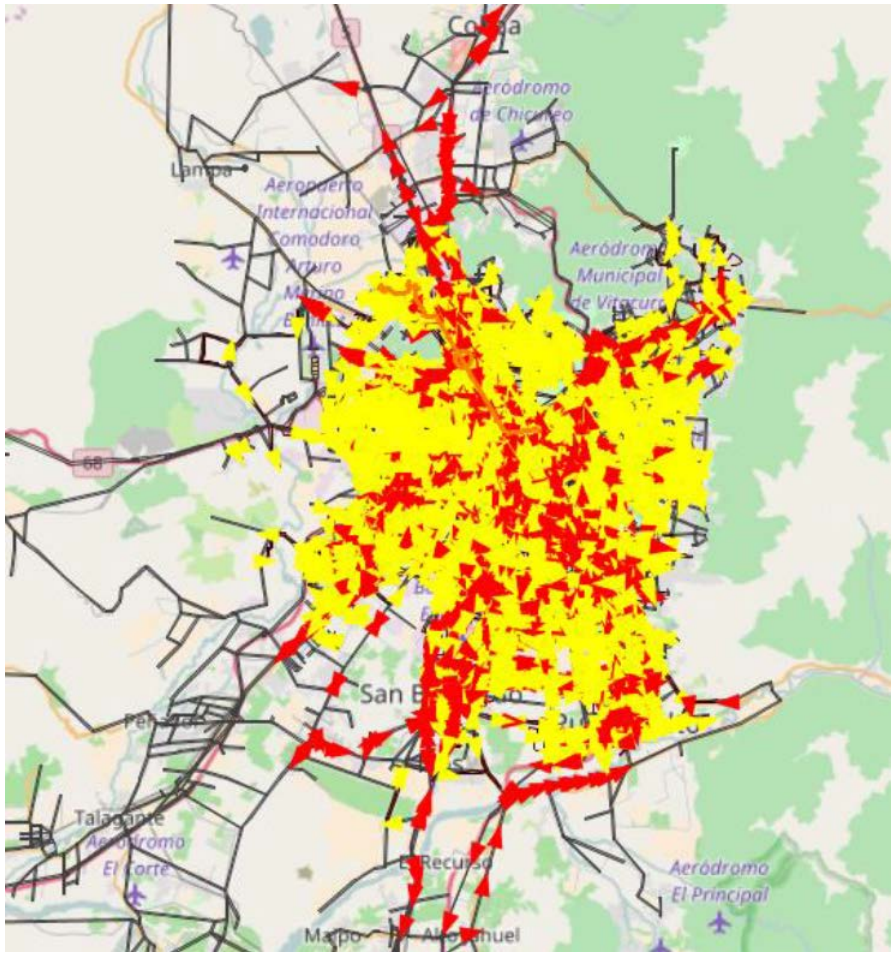
Table 2: Summary of colectivo observations

colectivo line	Slack time	Interval	travel time	fare
5055	00:00:54	00:02:26	00:09:15	300 CLP
5036	00:05:51	00:03:06	00:31:59	450 CLP
9006	00:04:41	00:03:33	00:42:24	800 CLP
8009	00:02:27	00:02:47	00:43:18	600 CLP
7013	00:01:46	00:07:03	00:48:55	1000 CLP
7001	00:02:09	00:04:24	01:02:07	900 CLP
7002	00:01:55	00:04:13	01:05:44	900 CLP
6066 - 6068		00:04:45		1750 – 2100 CLP
8002	00:00:34	00:05:38	01:47:36	1400 CLP
5004	00:02:44	00:03:29	01:48:00	1700 CLP
8020	00:02:10	00:15:11	02:15:00	2300 CLP
Arithmetic mean	**00:02:31**	**00:05:09**	**01:06:45**	**1116 CLP**

Table 3: Colectivo intervals

Colectivo line	Average interval	Peak interval	Off-peak interval
8020	00:15:11	00:15:30	00:14:20
7013	00:07:03	00:07:30	00:06:23
8002	00:05:38	00:06:13	00:05:00
7002	00:04:13	00:04:15	00:04:10
7001	00:04:24	00:04:15	00:04:35
6066 - 6068	00:04:45	00:03:47	00:05:43
5004	00:03:29	00:03:36	00:03:13
9006	00:03:33	00:03:03	00:04:03
5036	00:03:06	00:02:48	00:03:20
8009	00:02:47	00:02:33	00:03:05
5055	00:02:26	00:02:11	00:02:38
arithmetic mean	**00:05:09**	**00:05:04**	**00:05:08**

Felix (v3): colectivo (also) as PT feeder

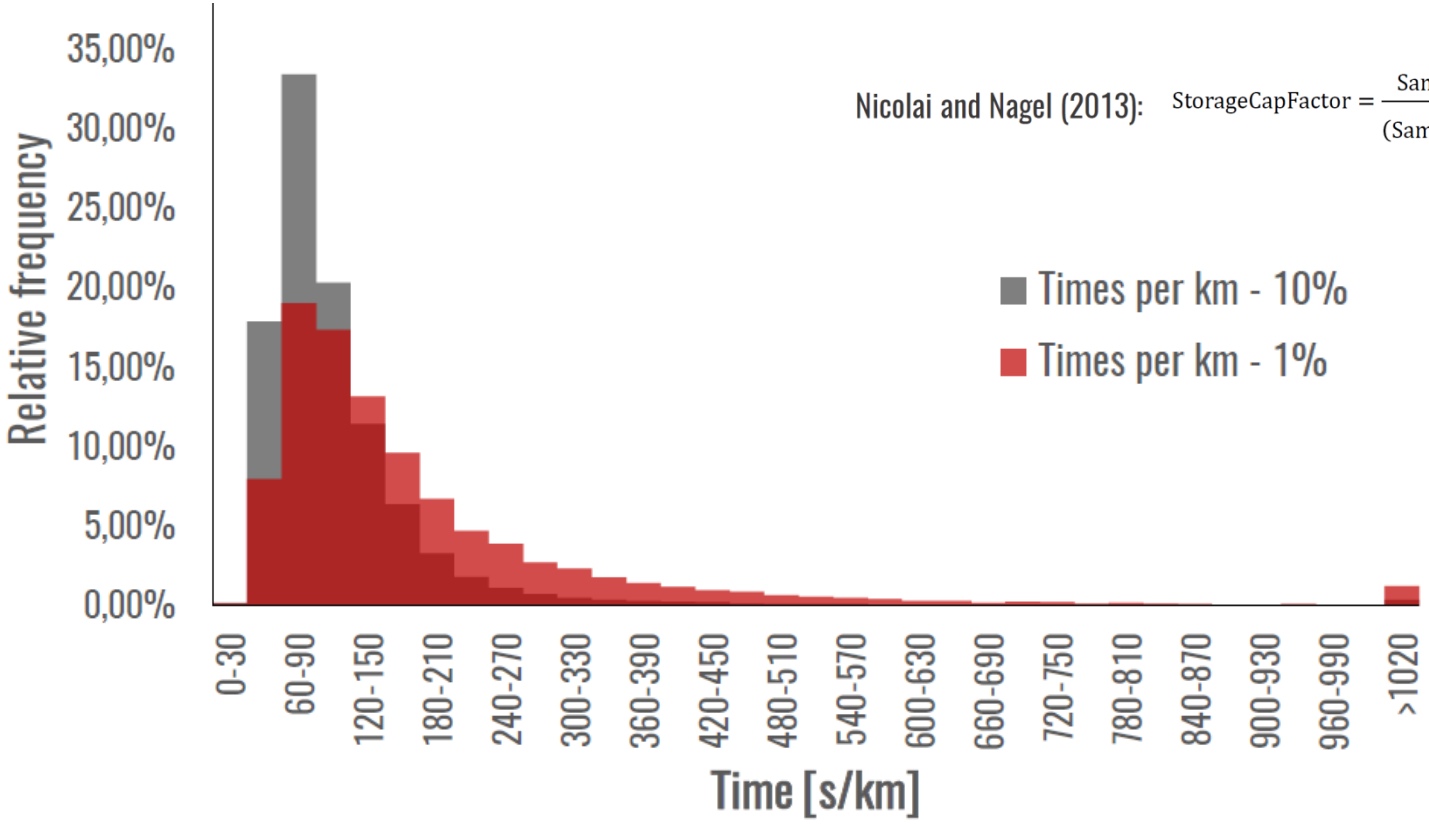


Challenges and first applications

Influence of sample size on travel times

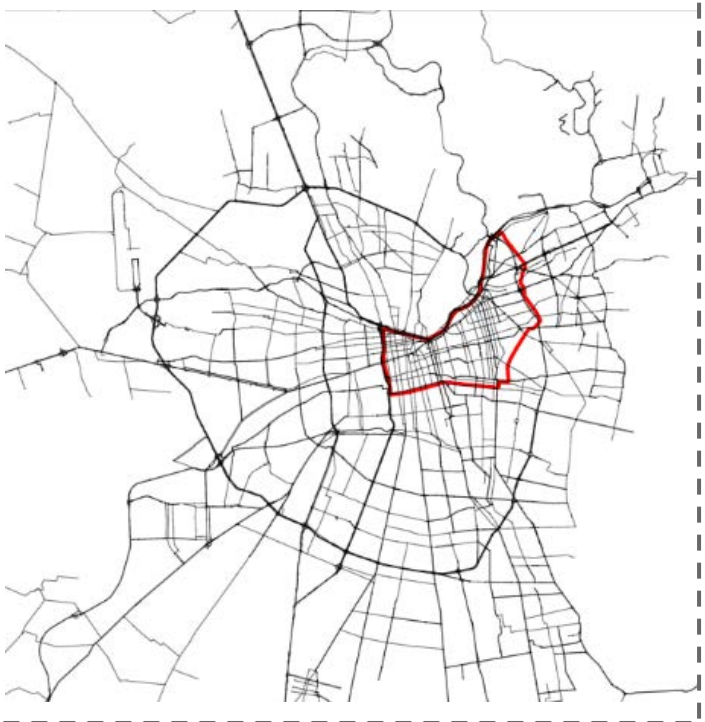
Sample size rate	C_f factor	N_{veh} factor
1%	0,01	0,032
10%	0,1	0,18

Nicolai and Nagel (2013): $\text{StorageCapFactor} = \frac{\text{SamplingRate}}{(\text{SamplingRate})^{\frac{1}{4}}}$



Cordon pricing schemes: setup

Outer cordon



$\tau_{in} = \$6.000$
$\tau_{out} = \$3.600$

Inner cordon

$\tau_{in} = \$6.000$
$\tau_{out} = \$2.650$

Cordon pricing schemes: elasticities 7:30-8:30

In- and outflow

Links	1% case [%]	10% case [%]	SDG (2009) [%]
Out	-83,69	-86,16	~-70
In	-83,03	-83,48	~-50

Total travel distance car

Scenario	1% case [%]	10% case [%]	SDG (2009) [%]
Exterior Cordon	-23,46	-22,68	~-5
Triangle Cordon	-4,44	-4,10	~-1,5

Number of car trips

Scenario	Caso 1% [%]	Caso 10% [%]	SDG (2009) [%]
Exterior Cordon	-20,79	-20,32	-5,8
Triangle Cordon	-3,90	-3,51	-1,4

Summary and outlook

Open data scenario

- Documentation: https://svn.vsp.tu-berlin.de/repos/public-svn/publications/vspwp/2016/16-02/ (more to come for v2, v3, …)
- Different versions of the runnable MATSim scenarios (no MATSim installation necessary): https://svn.vsp.tu-berlin.de/repos/public-svn/matsim/scenarios/countries/cl/santiago/
- Code: https://github.com/matsim-org/playgrounds/tree/master/santiago

Further steps

- Ideal for BSc or MSc theses; the idea is to integrate every improvement into the current state and create a new version.

- Possible improvements:
  - ~~Synthesize a 10% or 100% population~~
  - ~~Calibrate travel times, counts~~
  - ~~Add tolled roads~~
  - ~~Activity distribution according to land use data~~
  - Network corrections with automatic feedback to OpenStreetMap
  - Add freight transport (important for emissions!)
  - Add ~~colectivos~~, taxis
  - Include capacity constraints for PT vehicles
  - Map PT routes to the road network (interaction with cars)
  - Include bike as transport mode
  - …

Backup

Opportunities

- Social value: better transport planning through competition (identification of weak spots of the system, policy studies, environmental/social analysis, provision of public infrastructure)
- Commercial value: creating a platform for innovative mobility-based services (car sharing systems, supply chain/location planning, delivery/logistic planning, navigation, …)
- Transparency:
- Participation/engagement: