Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

WORKSHOP UCD

FUZZY LOGIC METHOD TO DEAL WITH UNCERTAINTY IN ENGINEERING PROBLEMS

AUTOR:
Penélope Gómez Jiménez

Dublin, March the 24th

INDEX

1. INTRODUCTION AND OBJECTIVES

1.1. IN TRANSPORT PLANNING…

- ✓In traffic operation management, enormous quantities of data are required, one of the most important data are $\Rightarrow$ **traffic volume data.**
- ✓Traffic data obtained in the field usually **have some errors** $\Rightarrow$ **inconsistent.**
- ✓If there are inconsistencies **data cannot be used** as input data in traffic planning algorithms.
- ✓In the past, certain methods were applied to adjust the observed values **classical methods**.
- ✓**Missing data processing** is another common issue $\Rightarrow$ **Imputation Techniques**.

1.2. CLASSIC MODELS TO ADJUST FIELD DATA TO REACH CONSISTENCY

EXISTING METHODS

1.2.1. Manual method

- 1.2.2.Least-squares method
- 1.2.3. Maximum-likelihood method
- 1.2.4. Fuzzy-regression method
- 1.2.5. Fuzzy-optimization method
- 1.2.6. Necessity-interval-regression

All methods, except for Manual M., use the optimization approach to adjust the values as close to the observed values, subject to the constraints.

1.3. THE AIM OF THIS METHOD IS

✓**Optimize** the solution.

✓Be able to introduce **subjective** analyst perception $\Rightarrow$ **fuzzy logic**

The **specific objectives already reached** are:

1) Obtaining a **consistent set** of values while preserving the integrity and include their reliability;

2) Identifying faulty TCS without the need for additional information;

3) Imputing values for data collected in the field whose information is missing;

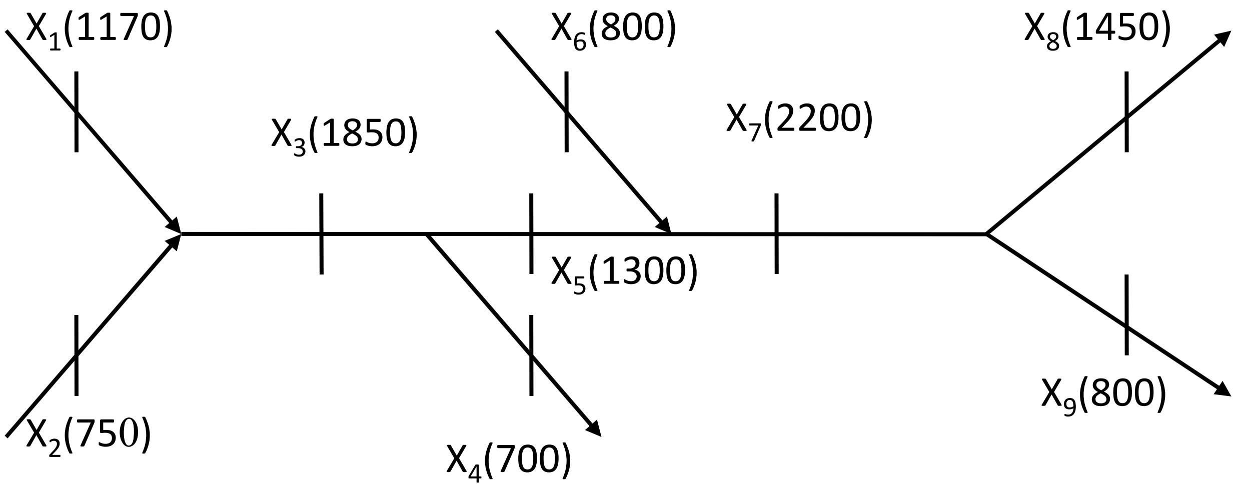
4) Adjusting the numbers of **boarding and alighting** passengers on a public transport transit line, without requiring counts at all stops.

5) Estimate OD matrices in a network

Research Going on: Minimize the settling time of the Frequency in the Primary Control

Bilevel fuzzy optimization to pre-process traffic data to satisfy the law of flow conservation, (2011). Transportation Research Part C, 19 (1), pp. 29-39.

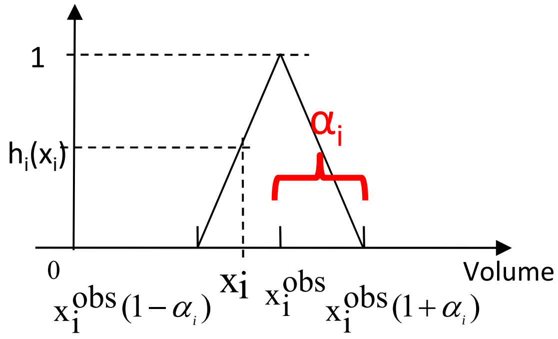
MATERIALS:



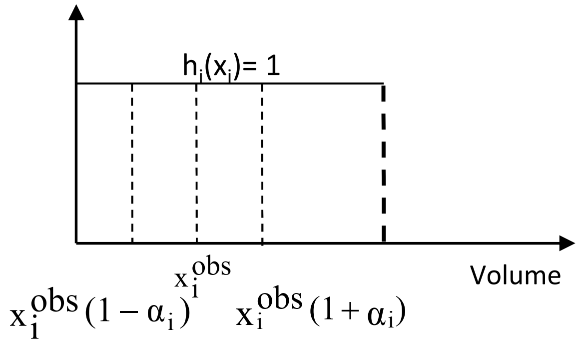
x_i (observed value)

$$x_1+x_2 \neq x_3$$
$$x_3-x_4 \neq x_5$$
$$x_5+x_6 \neq x_7$$
$$x_8+x_9 \neq x_7$$

DATA: TYPES OF DATA



Crisp Number



Missing Value

MATHEMATICALLY

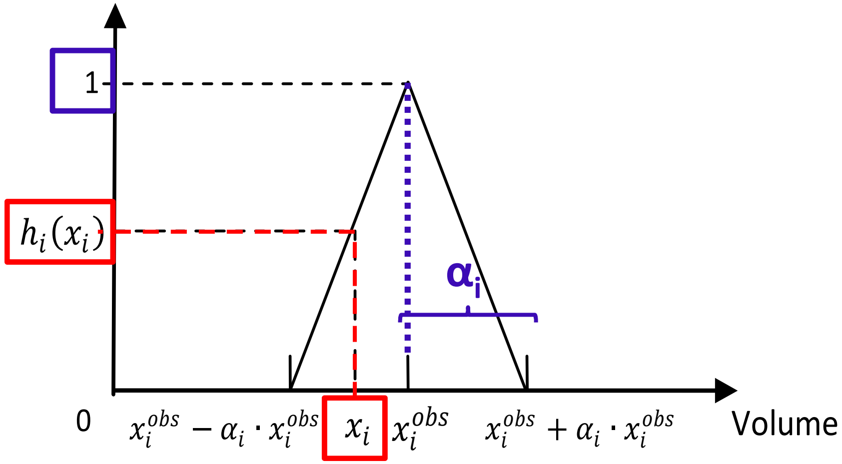
Two existing methods of optimization (Kikuchi and Miljkovic, 1999):

1. **Maximizing the Minimum (MM) of h_i method**:

$$\max \min_i \left\{ h_{x_i^{obs}} (x_i) \right\}$$

2) **Maximizing the Sum (MS) of h_i method**:

$$\max \left\{ \sum_i h_{x_i^{obs}} (x_i) \right\}$$



Subject to

Constraints related to the membership functions:

$$x_i^{obs} - \alpha_i \cdot x_i^{obs} \le x_i \le x_i^{obs} + \alpha_i \cdot x_i^{obs}$$

$$0 \le h_{x_i}^{obs} \le 1$$

Constraints related to the conservation of flow at each control point

$$x_1 + x_2 = x_3$$
$$x_3 - x_4 = x_5$$
$$x_5 + x_6 = x_7$$
$$x_8 + x_9 = x_7$$

MATHEMATICALLY

The proposed method of optimization **Bilevel Optimization Method (BO)**

Two steps:

1st . MM method:

$$\max \ \min_{i} \left\{ h_{x_i^{obs}}(x_i) \right\}$$

h=h* is recorded

2nd Among all combination getting h=h*, **Maximizing the Sum (MS):**

$$\max \ \left\{ \sum_i h_{x_i^{obs}}(x_i) \right\}$$

Subject to

Constraints related to the membership functions: $h(x_i^{obs}) \geq h^*$

Constraints related to the conservation of flow at each control point

$$x_1 + x_2 = x_3$$
$$x_3 - x_4 = x_5$$
$$x_5 + x_6 = x_7$$
$$x_8 + x_9 = x_7$$

Results

Certain indicators should be borne in mind when the results are analysed:

- The **first indicator** is **the lowest value of *h***,

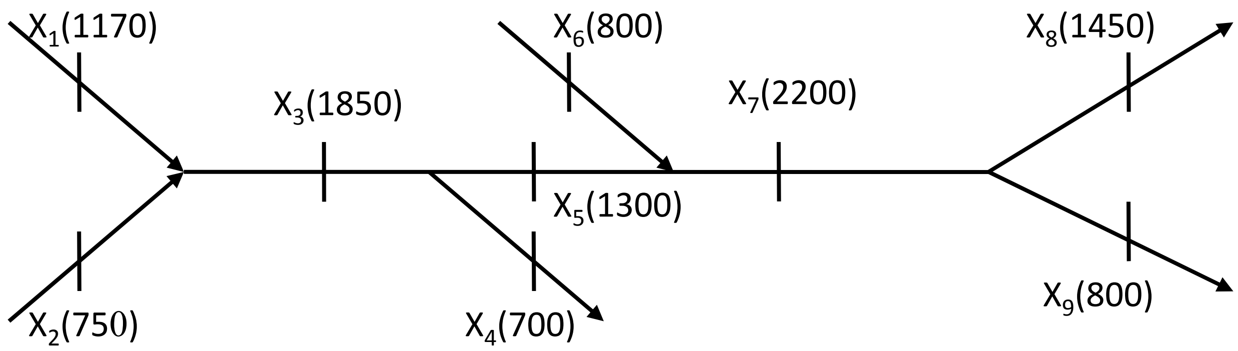
- The **second indicator** is **the sum of h_i**.

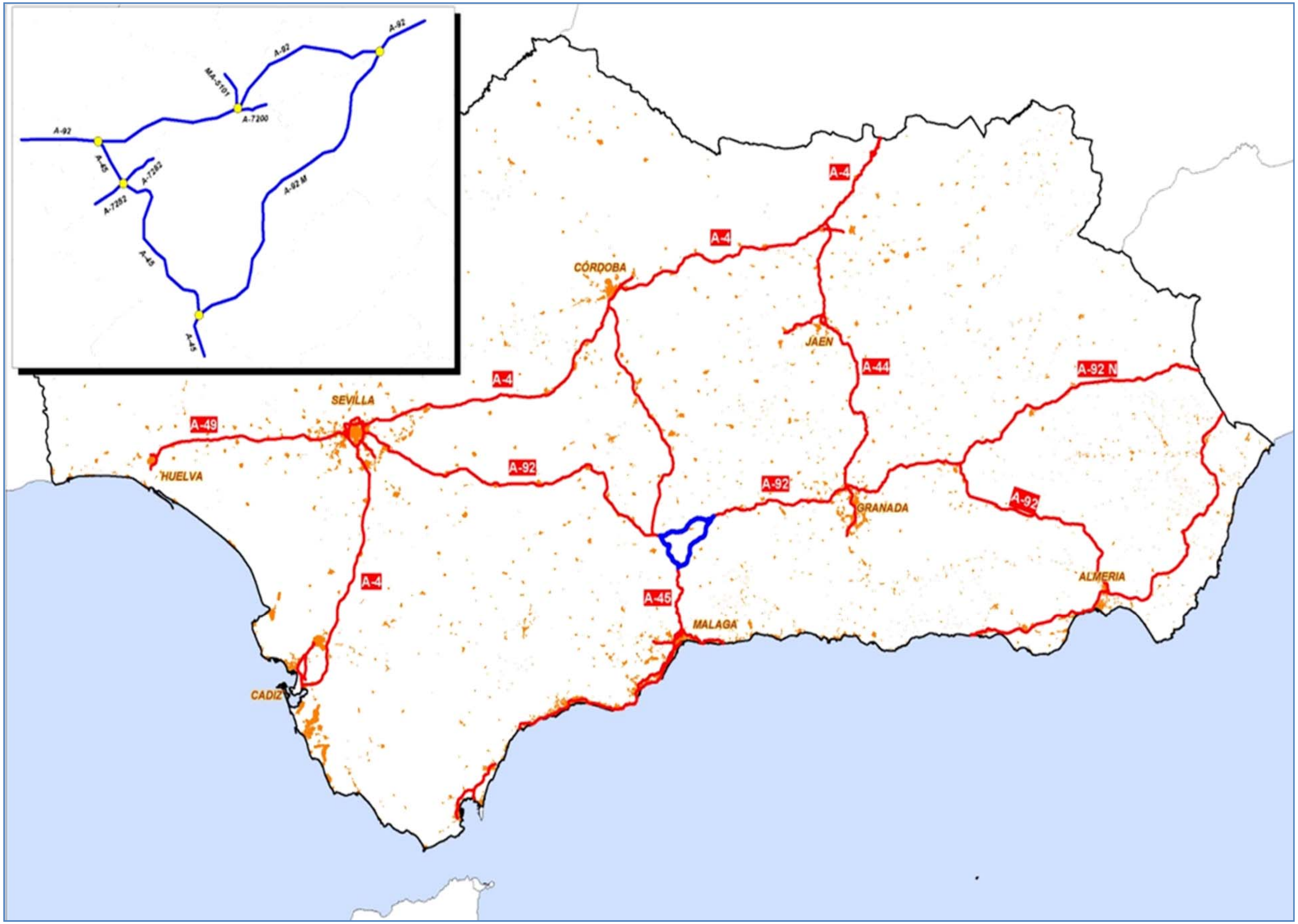
- A **third indicator** should be taken into account in those examples: the **average difference between the true consistent values and the adjusted values.**

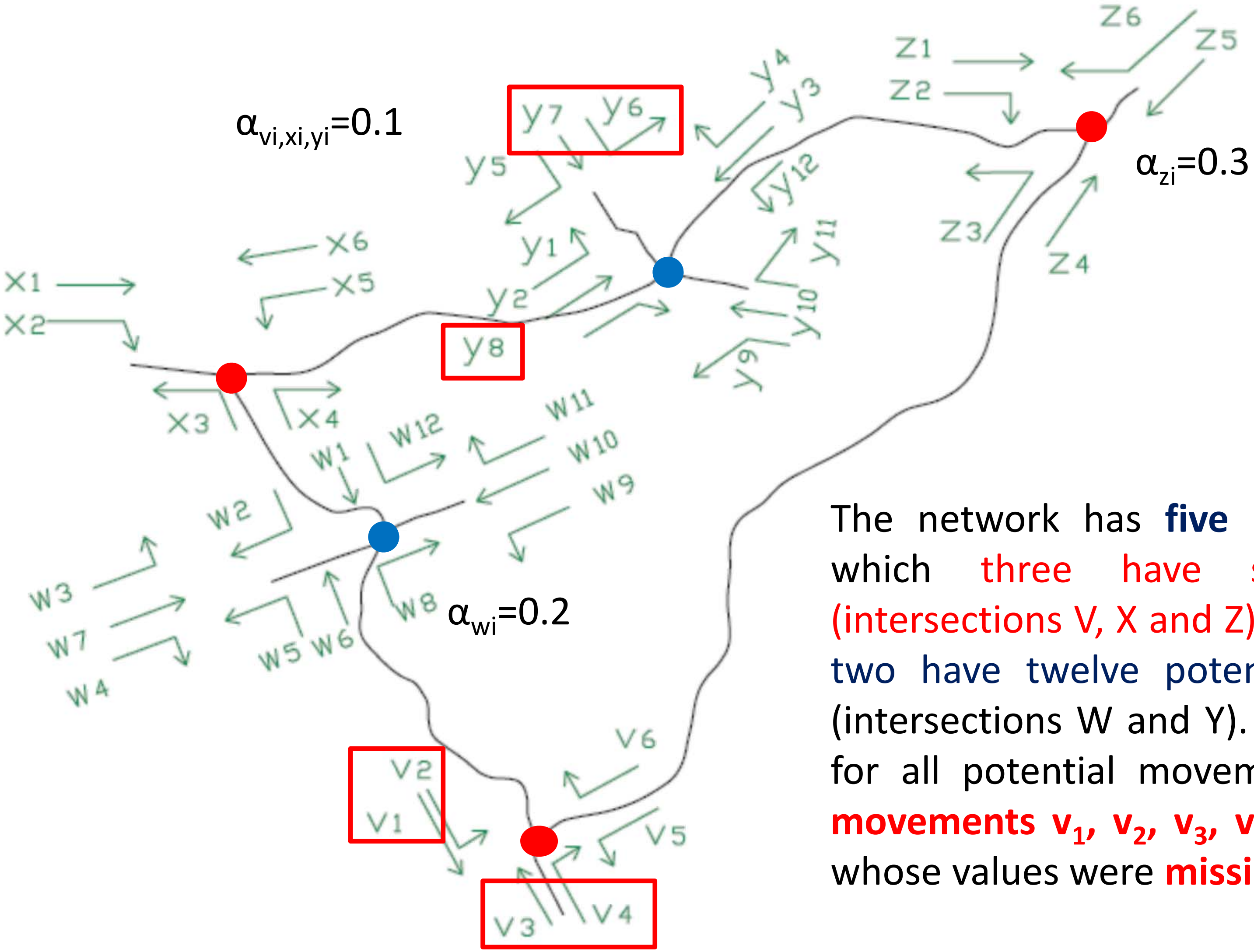
OPTIMIZATION PROCESS: NEW METHOD

It is a two-step or bilevel optimization method

x_i	Observed Value	max-min(h)		max Σh_i		BO	
		Adjusted value	h_i	Adjusted value	h_i	Adjusted value	h_i
x_1	1,170	1,195	0.7863	1,285	0.0171	**1,217**	**0.5983**
x_2	750	772	0.7067	750	1.0000	**750**	**1.0000**
x_3	1,850	1,967	0.3676	2,035	0.0000	**1,967**	**0.3676**
x_4	700	656	0.3714	635	0.0714	**656**	**0.3714**
x_5	1,400	1,311	0.3643	1,400	1.0000	**1,311**	**0.3643**
x_6	800	797	0.9625	800	1.0000	**847**	**0.4125**
x_7	2,200	2,108	0.5818	2,200	1.0000	**2,158**	**0.8091**
x_8	1,450	1,358	0.3655	1,400	0.6552	**1,358**	**0.3655**
x_9	800	750	0.3750	800	1.0000	**800**	**1.0000**
	min h_i	0.3643		0.0000		**0.3643**	
	sum h_i	4.8811		5.7437		**5.2887**	
	Average Δ	NA		NA		**NA**	




A-92
MA-5101
A-7200
A-45
A-7282
A-92 M
A-4
CORDOBA
JAEN
A-44
A-92 N
SEVILLA
A-49
HUELVA
GRANADA
A-45
MALAGA
ALMERIA
CADIZ



The network has **five intersections**, of which three have six movements (intersections V, X and Z), while the other two have twelve potential movements (intersections W and Y). Data is available for all potential movements except for **movements v_1, v_2, v_3, v_4, y_6, y_7, and y_8**, whose values were **missing.**

INDICATOR	MM α=0.1	MM α*	MS α=0.1	MS α*	BO α=0.1	BO α*
min h	0.29	0.54	~~0.00~~	~~0.49~~	**0.29**	**0.54**
sum h_i	26.16	36.50	40.14	40.85	$\boxed{35.97}$	$\boxed{38.61}$
Average Δ	NA	NA	NA	NA	**NA**	**NA**

* α=0.2 for w_i; α=0.3 for z_i and α=0.1 for rest of cases

Method to Detect Malfunctioning Traffic Count Stations, (2012). IET Intelligent Transport Systems, 6(4):364.

ERROR TYPES

ADMISSIBLE ERRORS: Errors that are within the measuring device's tolerance and, consequently, they depend on the precision defined for each device by the manufacturer

INADMISSIBLE ERRORS : These are errors that not only give erroneous information, but also invalidate the work done

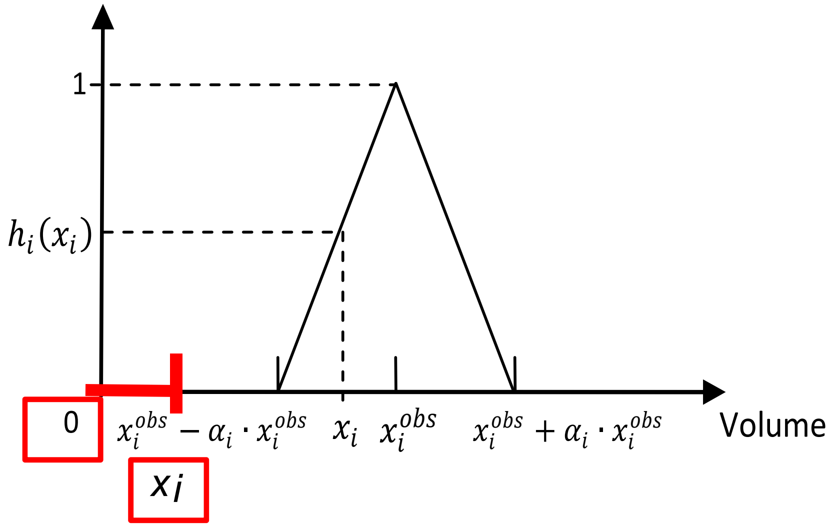
Analysis of recorded historical data from several cities in Unites States (Turner et al. (2000)) and found:

-**25% of missing data or under suspicion** to be erroneous.

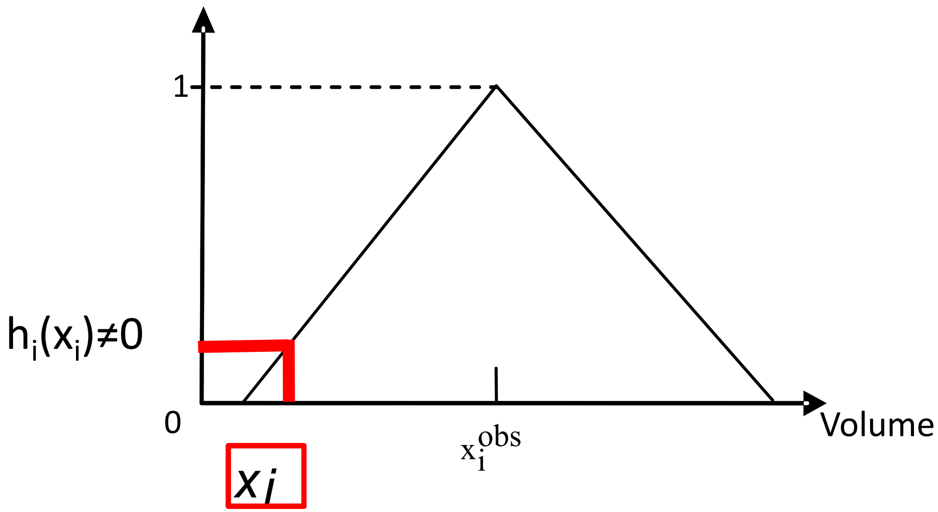
- At least the **25% of the detectors have any kind of failure** at any given time (Nguyen and Scherer (2003))and,

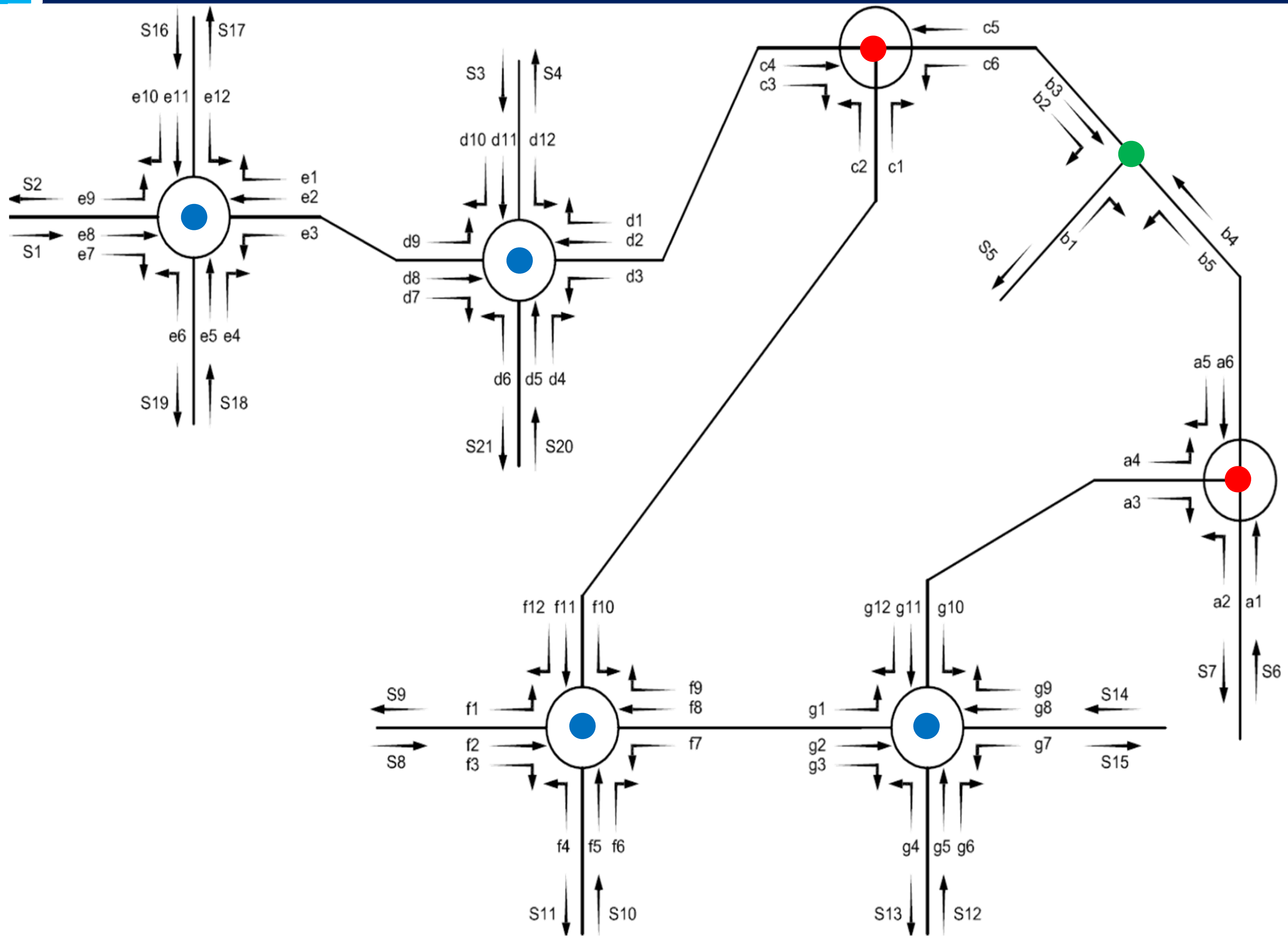
- **Even properly working detectors** often have until **a 5% of missing data**, (Kwon et al. (2004)).

✓ **First step**, with **α= device's tolerance** check if the program gets any $h_i(x_i)=0$.



✓ If so, **second step**, spread α until it is big enough to get $h_i(x_i) \neq 0$

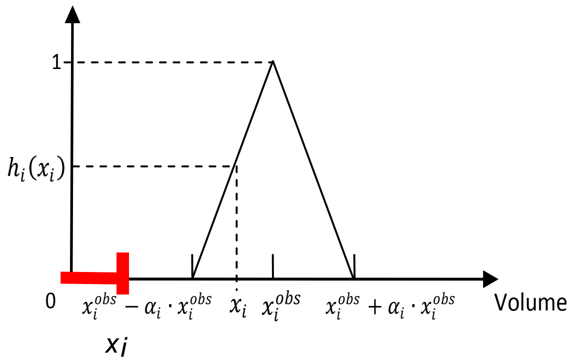


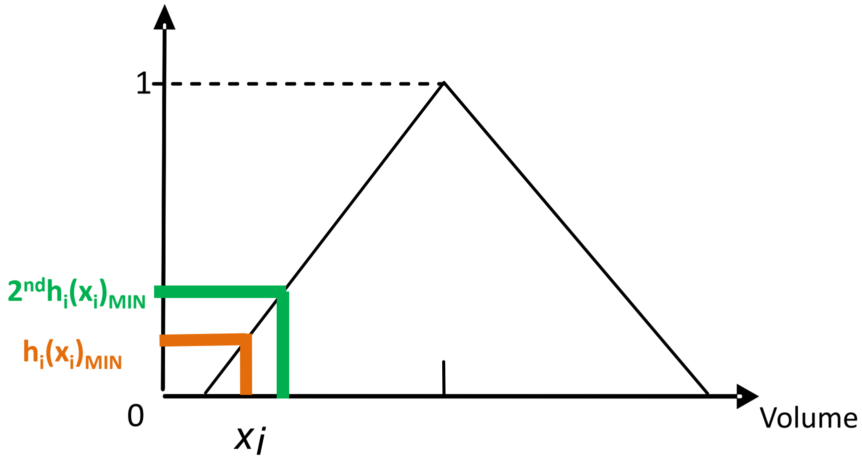

S16
S17
e10 e11
e12
S2
e9
e1
e2
e3
S1
e8
e7
e6
e5 e4
S19
S18
S3
S4
d10 d11
d12
d9
d1
d2
d3
d8
d7
d6
d5 d4
S21
S20
c5
c6
c4
c3
c2
c1
b3
b2
b1
b4
b5
S5
a5 a6
a4
a3
a2
a1
S7
S6
f12 f11
f10
S9
f1
f9
f8
f7
f2
f3
S8
f4
f5
f6
S11
S10
g12 g11
g10
g1
g9
g8
S14
g2
g3
g7
S15
g4
g5 g6
S13
S12

Consistent true data are available (Theoretical Values – **TV**), **randomly deform values** with the device's tolerance, ±3% $\Rightarrow$ (Observed Values – **OV**).

Next, **a datum is randomly selected and distorted to simulate a detector error** that exceeds the error specified by the manufacturer. Specifically, deviation of 75%, 500 examples generated:

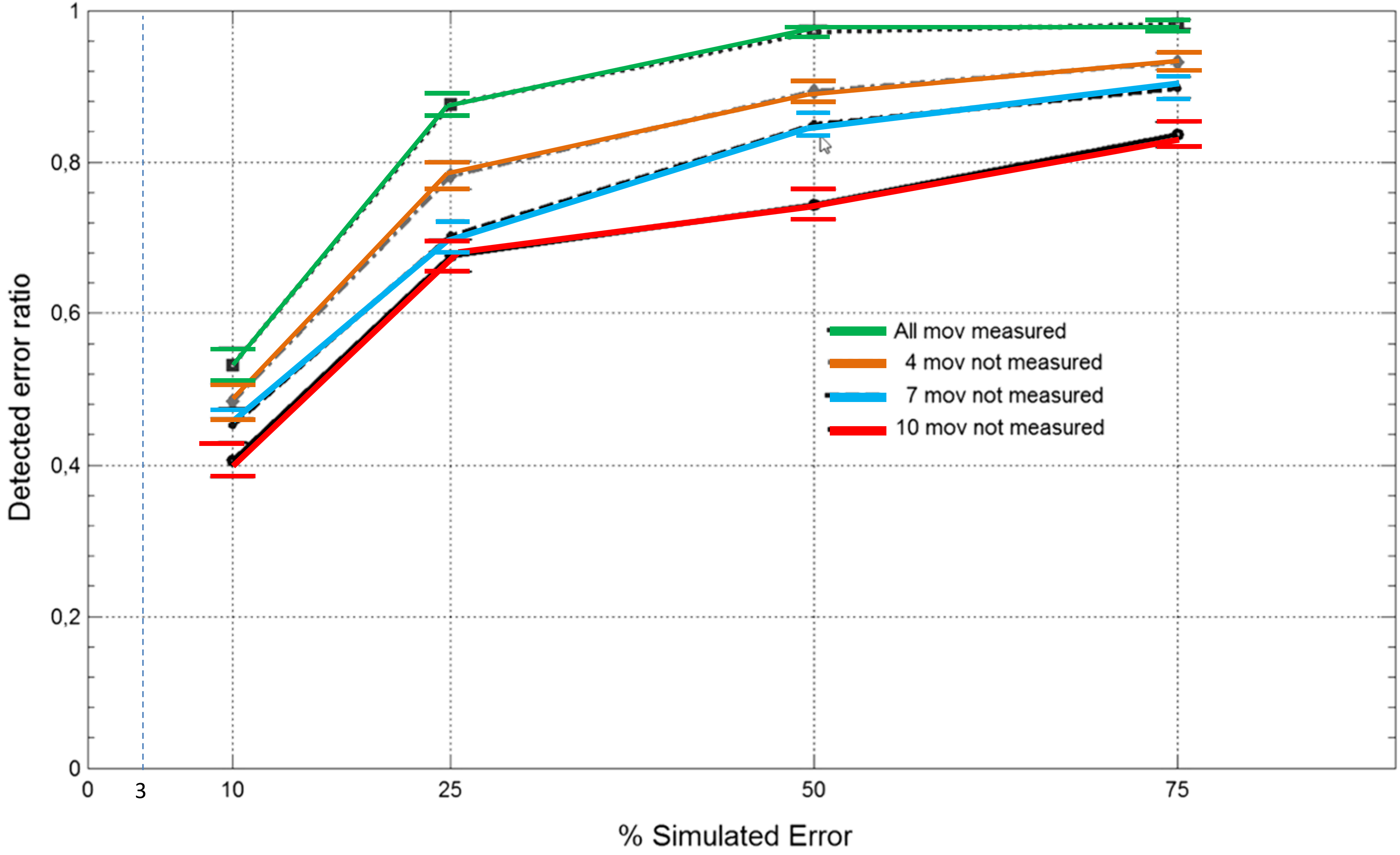
Percentage simulated error	Error is detected		Error is pointed out and gets by h_{min}		Error is pointed out and gets by 2nd h_{min}		Error is pointed out in total
	Number of times (A)	Proportion A/500	Number of times (B)	Proportion B/500	Number of times (C)	Proportion (C)/500	Success proportion (B+C)/500
75	491	0.982	443	0.886	23	0.046	0.932





In the same way a 50%, 25% and 10% error are simulated:

Percentage simulated error	Error is detected		Error is pointed out and gets by h_{min}		Error is pointed out and gets by 2^{nd} h_{min}		Error is pointed out in total
	Number of times (A)	Proportion A/500	Number of times (B)	Proportion B/500	Number of times (C)	Proportion (C)/500	Success proportion (B+C)/500
75	491	0.982	443	0.886	23	0.046	0.932
50	486	0.972	386	0.772	37	0.074	0.846
25	438	0.876	269	0.538	27	0.054	0.592
10	266	0.532	137	0.274	24	0.048	0.322


Detected error ratio
% Simulated Error
0
0,2
0,4
0,6
0,8
1
0
3
10
25
50
75
All mov measured
4 mov not measured
7 mov not measured
10 mov not measured

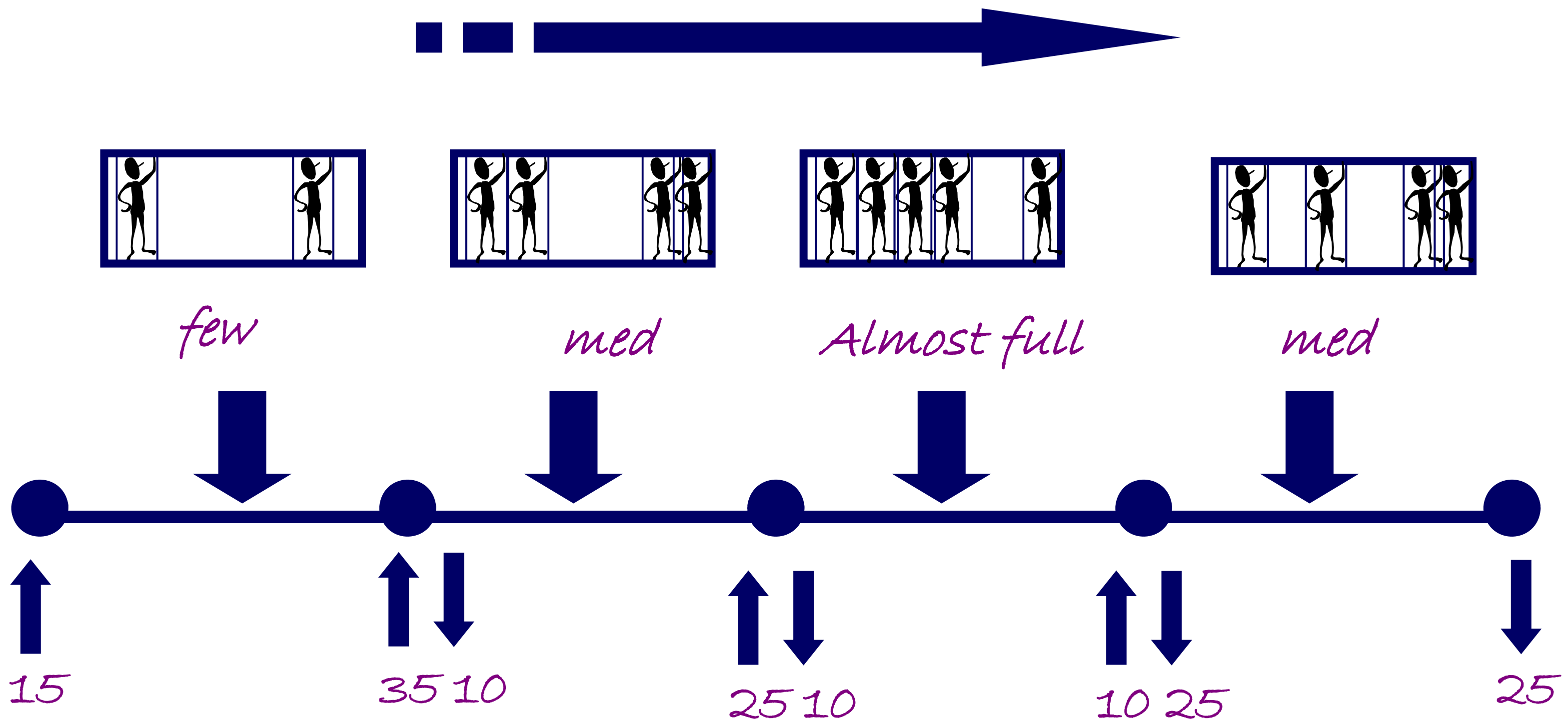
Adjustment boarding and alighting passengers on a bus transit line using qualitative information, (2013). Applied Mathematical Modelling, 38 (3), pp.1147-1158

STATE OF THE ART: THE PROBLEM OF ADJUSTING BOARDING AND ALIGHTING PASSENGERS IN A TRANSIT LINE

Collecting **data on passenger boardings is easy thanks to ticketing systems as payment option.**

However, the systems cannot be used to obtain data on the number of alightings

So far, to get an acceptable estimation of alighting several surveyors may be needed.

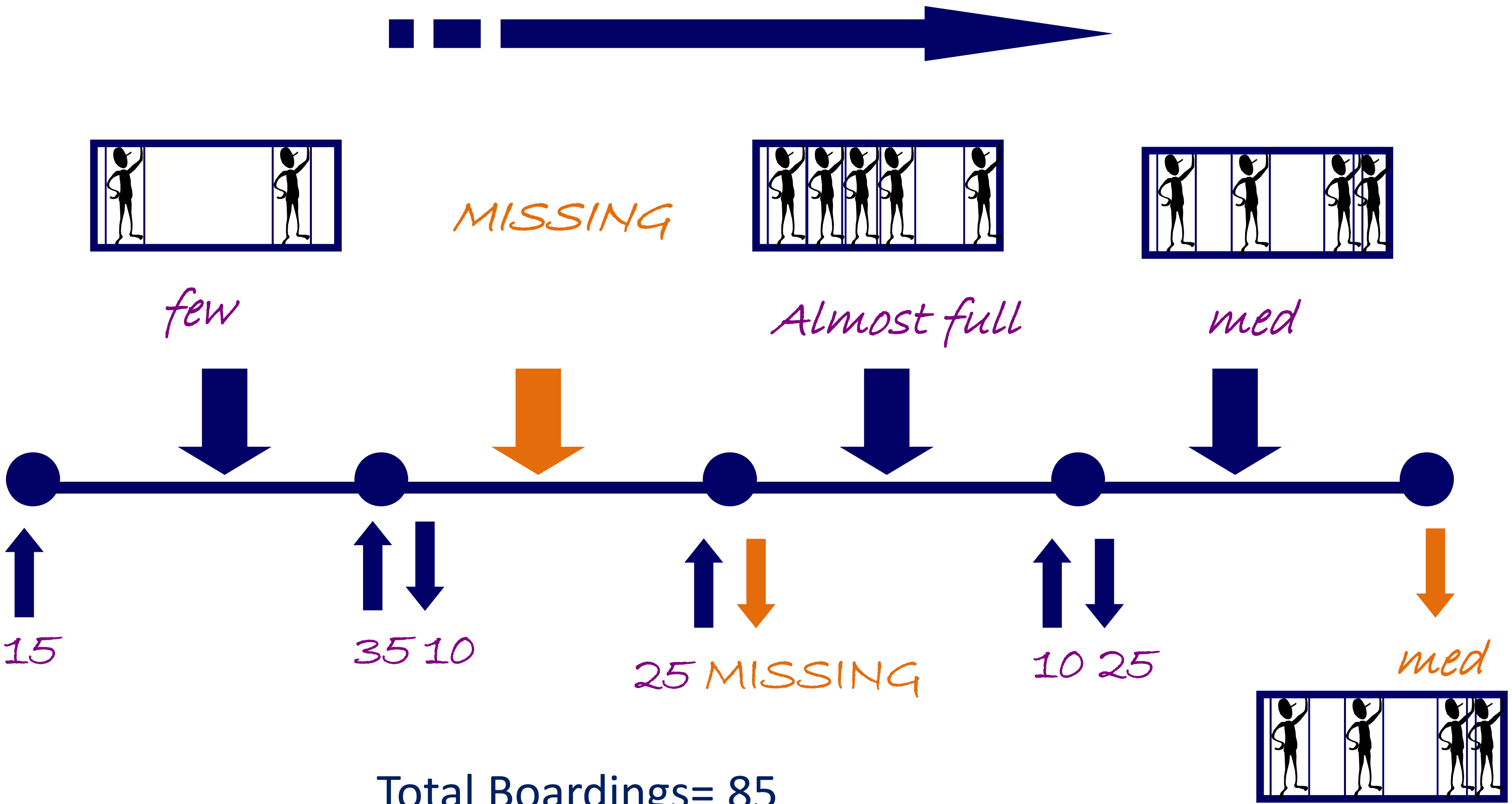


Total Boardings= 85

Total Alightings= 70

Total Boardings= Total Alightings

Adjusted (B_i, A_i , L_{ij}) ≈ Observed (b_i , a_i, l_{ij})



Total Boardings= 85

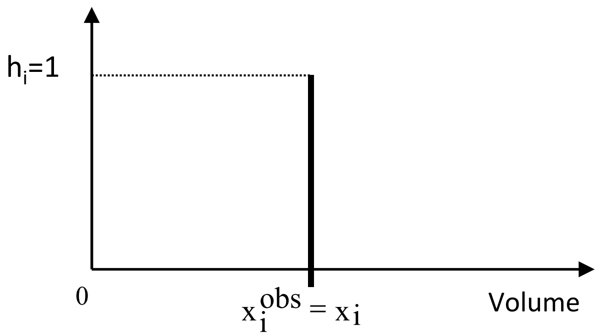
Total Alightings= 70

Total Boardings= Total Alightings

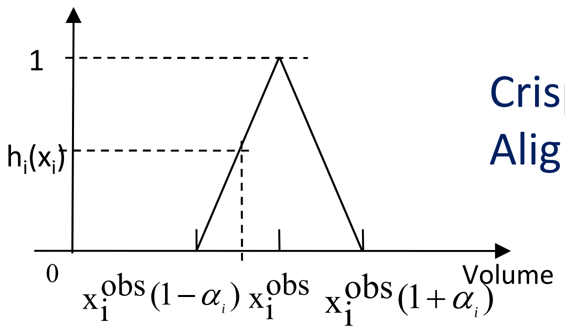
Adjusted (B_i, A_i , L_{ij}) ≈ Observed (b_i , a_i, l_{ij})

DATA: TYPES OF DATA

Fix
Boardings



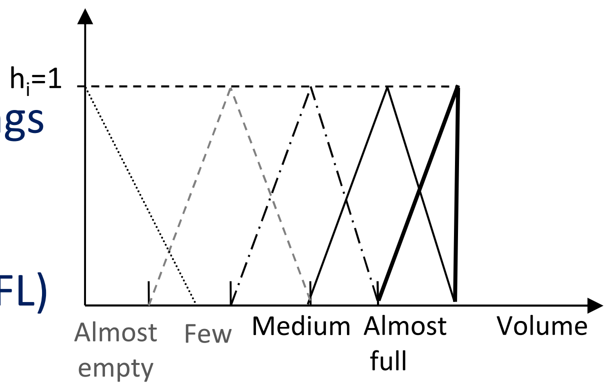
Fix Number



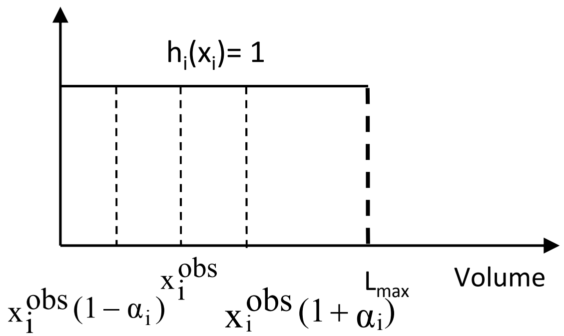
Crisp Number

Crisp
Alightings

Fuzzy Alightings
(FA)
&
Fuzzy Loads (FL)



Fuzzy Value

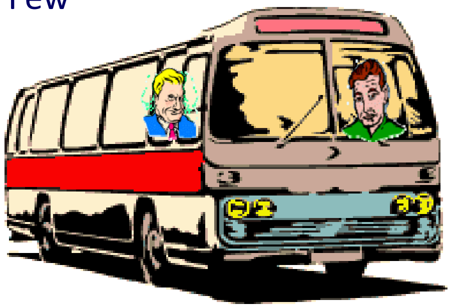


Missing Value

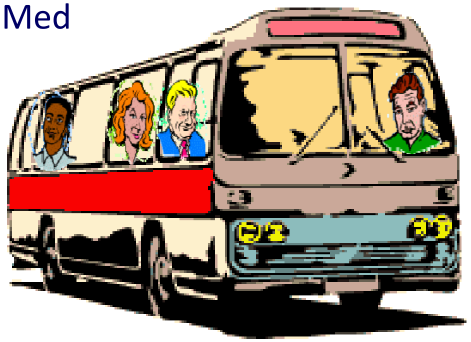
Missing Alightings
(MA)
&
Missing Loads
(ML)

FUZZY VALUE: LOADS AND ALIGHTINGS

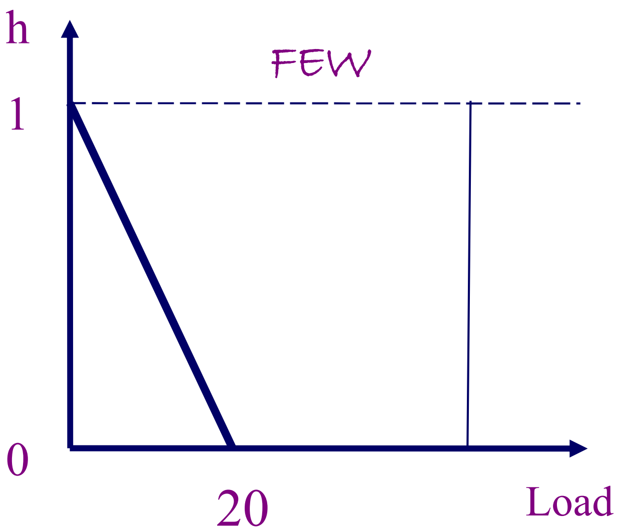
Few

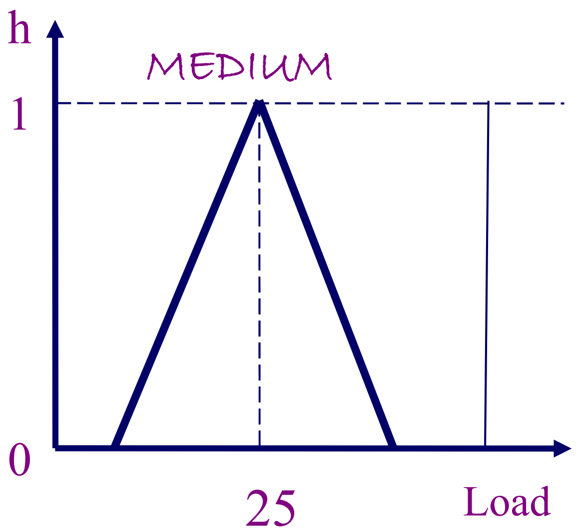


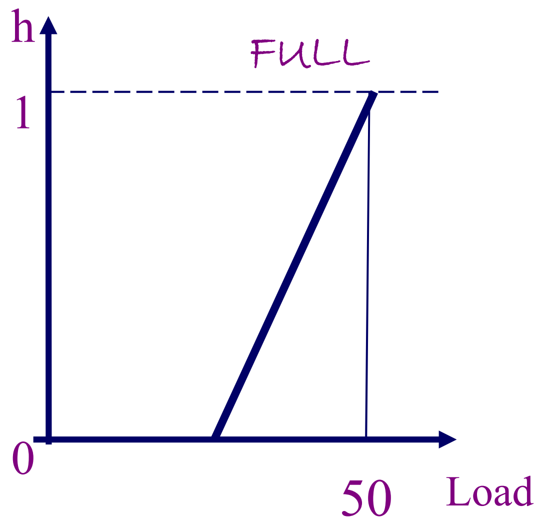
Med

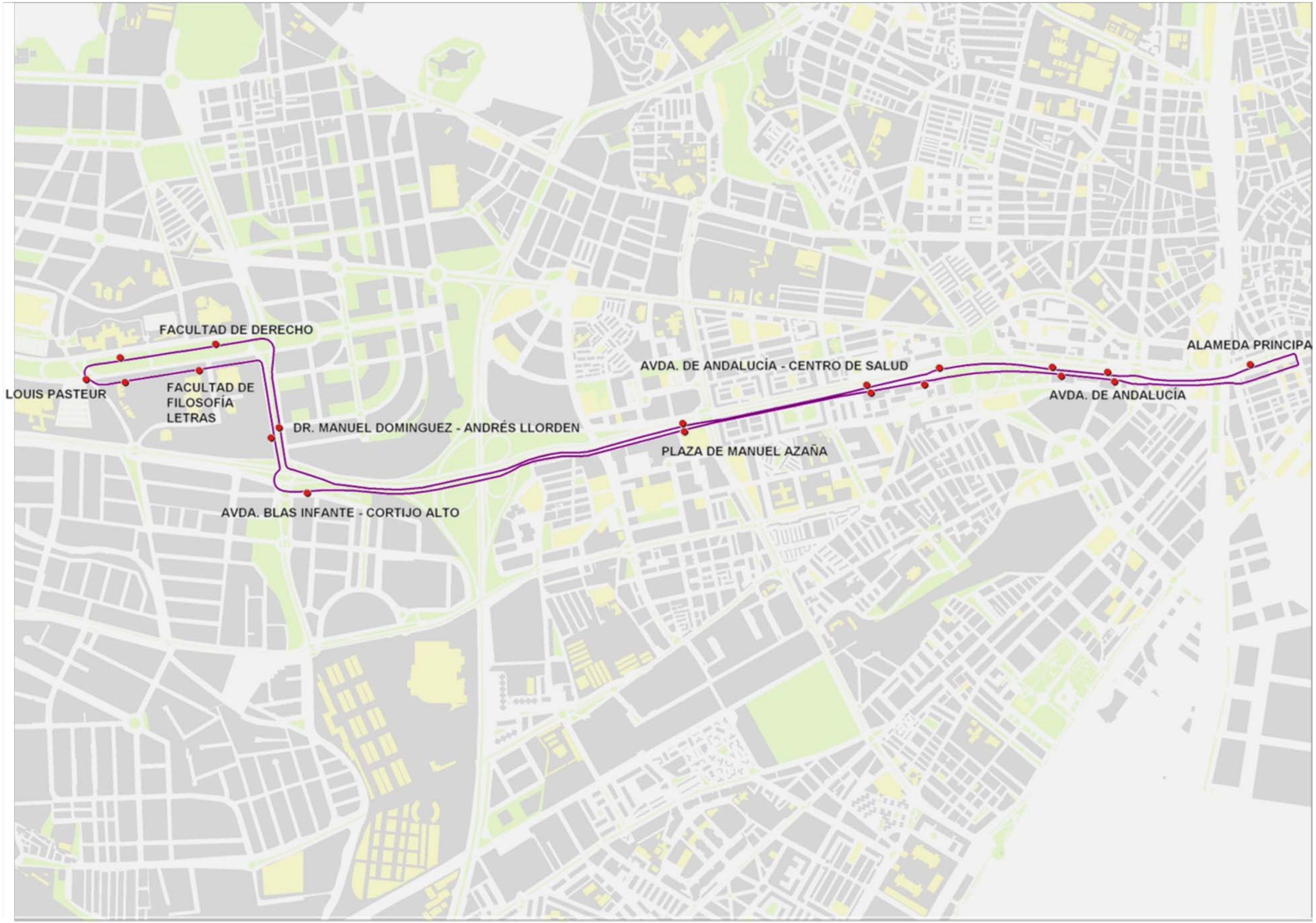


Full








FACULTAD DE DERECHO
LOUIS PASTEUR
FACULTAD DE FILOSOFÍA LETRAS
DR. MANUEL DOMINGUEZ - ANDRÉS LLORDEN
AVDA. BLAS INFANTE - CORTIJO ALTO
AVDA. DE ANDALUCÍA - CENTRO DE SALUD
PLAZA DE MANUEL AZAÑA
AVDA. DE ANDALUCÍA
ALAMEDA PRINCIPA

As the true values are known, 1000 database of observed values have been created randomizing the alightings and the load:

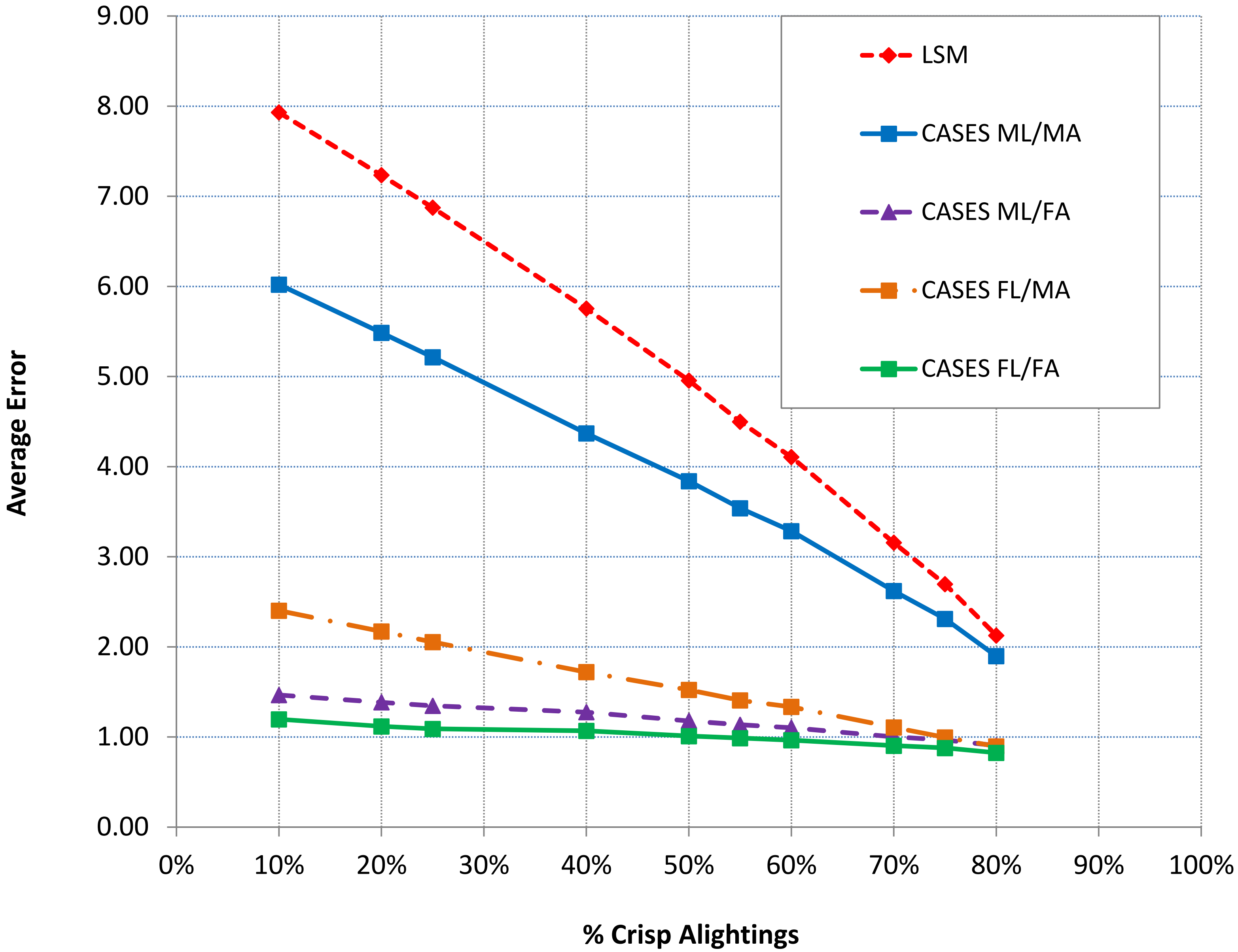
Cases ML/MA	Cases FL/MA	Cases ML/FA	Cases FL/FA
ML/20MA	FL/20MA	ML/20FA	FL/20FA
ML/25MA	FL/25MA	ML/25FA	FL/25FA
ML/30MA	**FL/30MA**	**ML/30FA**	**FL/30FA**
ML/40MA	FL/40MA	ML/40FA	FL/40FA
ML/45MA	FL/45MA	ML/45FA	FL/45FA
ML/50MA	FL/50MA	ML/50FA	FL/50FA
ML/60MA	**FL/60MA**	**ML/60FA**	**FL/60FA**
ML/75MA	FL/75MA	ML/75FA	FL/75FA
ML/80MA	FL/80MA	ML/80FA	FL/80FA
ML/90MA	**FL/90MA**	**ML/90FA**	**FL/90FA**
Note: ML: missing load; FL: fuzzy load; MA: missing alightings; FA: fuzzy alightings; xxMA: xx% of missing alightings, (100-xx)% of alightings crisp; xxFA: xx% of alightings fuzzy, (100-xx)% of alightings crisp			

The true boarding and alighting data x_i^{true} are used as reference to calculate the third indicator.

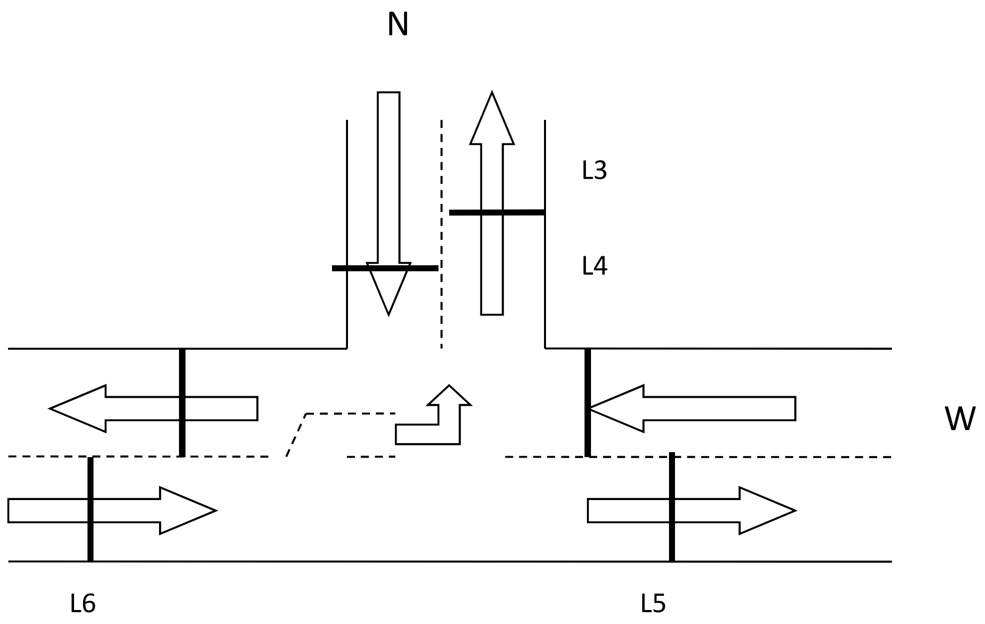
$$\varepsilon = \frac{\sum_{i=0}^{n} \left| x_i^{adj} - x_i^{true} \right|}{n}$$

For comparison purposes, we use the Least Square Method (LSM) as benchmark. LSM uses only quantitative data, so it is applied and only compared with the 10 Cases ML/MA.

LSM		Cases ML/MA		Cases FL/MA		Cases ML/FA		Cases FL/FA	
CASE	ε	CASE	ε	CASE	ε	CASE	ε	CASE	ε
ML/20MA	2.13	ML/20MA	1.90	FL/20MA	0.90	ML/20FA	0.91	FL/20FA	0.82
ML/25MA	2.70	ML/25MA	2.31	FL/25MA	1.00	ML/25FA	0.97	FL/25FA	0.88
ML/30MA	3.16	ML/30MA	2.62	FL/30MA	1.11	ML/30FA	1.00	FL/30FA	0.90
ML/40MA	4.11	ML/40MA	3.28	FL/40MA	1.33	ML/40FA	1.10	FL/40FA	0.96
ML/45MA	4.50	ML/45MA	3.54	FL/45MA	1.41	ML/45FA	1.14	FL/45FA	0.99
ML/50MA	4.96	ML/50MA	3.84	FL/50MA	1.52	ML/50FA	1.18	FL/50FA	1.01
ML/60MA	5.75	ML/60MA	4.37	FL/60MA	1.72	ML/60FA	1.28	FL/60FA	1.07
ML/75MA	6.87	ML/75MA	5.21	FL/75MA	2.05	ML/75FA	1.35	FL/75FA	1.09
ML/80MA	7.24	ML/80MA	5.49	FL/80MA	2.17	ML/80FA	1.39	FL/80FA	1.12
ML/90MA	7.93	ML/90MA	6.02	FL/90MA	2.40	ML/90FA	1.47	FL/90FA	1.20


9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
Average Error
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
% Crisp Alightings
LSM
CASES ML/MA
CASES ML/FA
CASES FL/MA
CASES FL/FA

Evaluation of Trade-Offs between two data sources for the accurate Estimation of OD Matrices, (2015). Transportmetrica B Transport Dynamics, 3 (3), pp.222-245



O/D	E	N	W
E	-	x_1	x_2
N	x_3	-	-
W	x_4	x_5	-

The specific objectives pursued are:

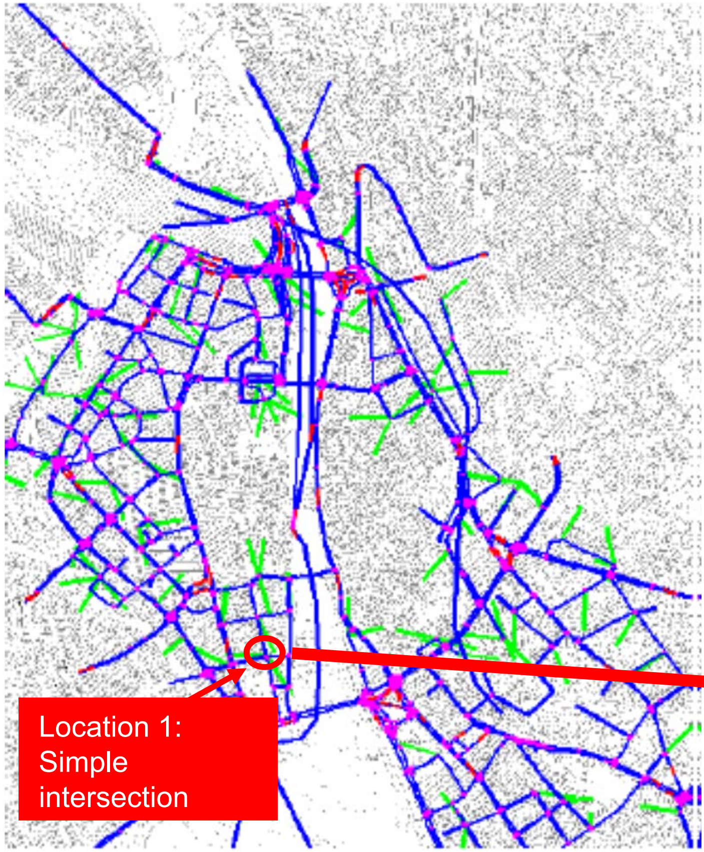
1) The introduction of a new BO method for the estimation of **O-D matrices** that has clear advantages over other existing methods.

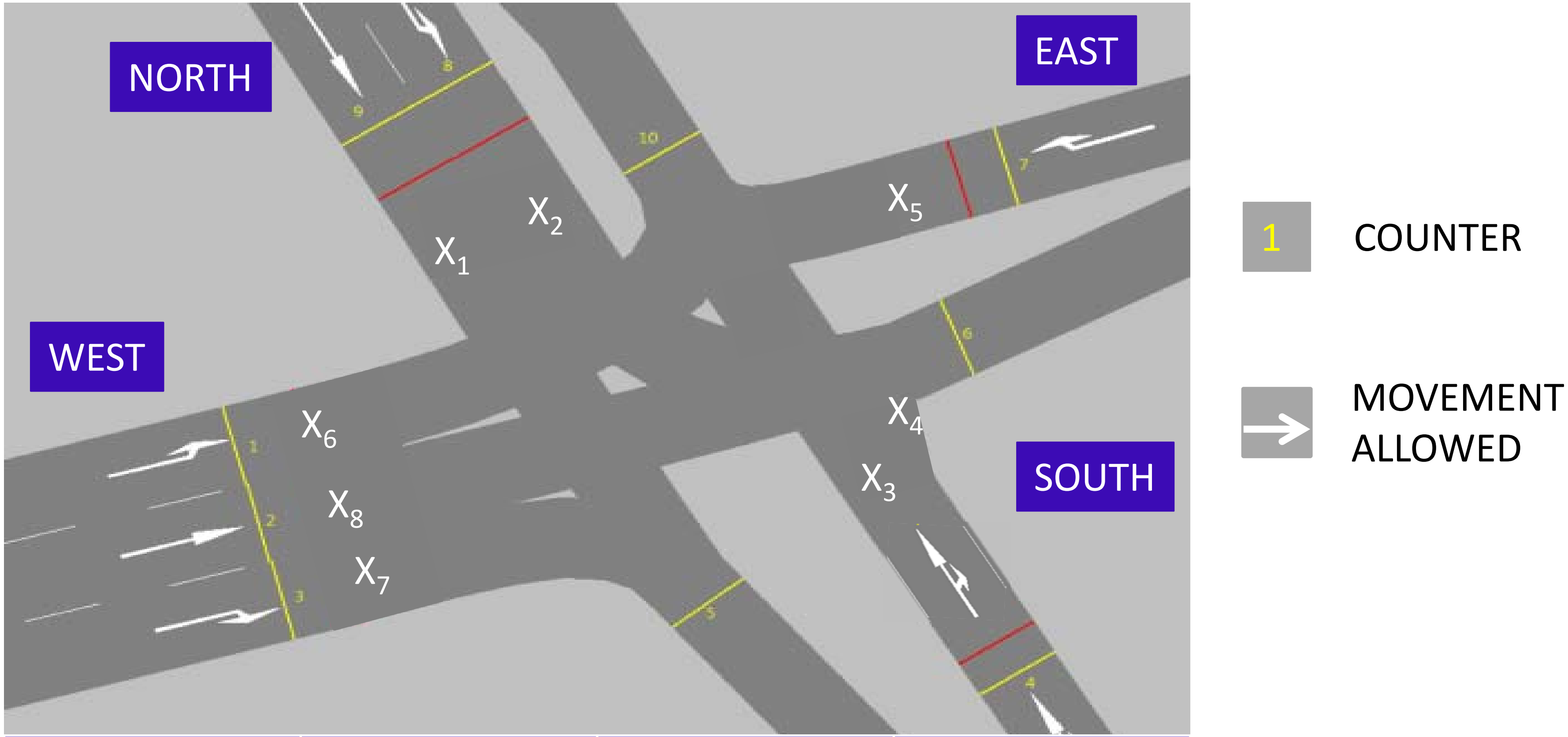
2) The evaluation of trade-offs between Floating Car Data-Penetration Rate (FCD-PR) and Loop Detector Data (LDC) for different accuracy levels on the estimation of OD matrices.

2.1. REQUIRED INPUTS

- ✓ Traffic micro-simulation model of the central area of Zurich (around 2.6 km^2) implemented in VISSIM

- ✓ Data from loop detectors and FCD

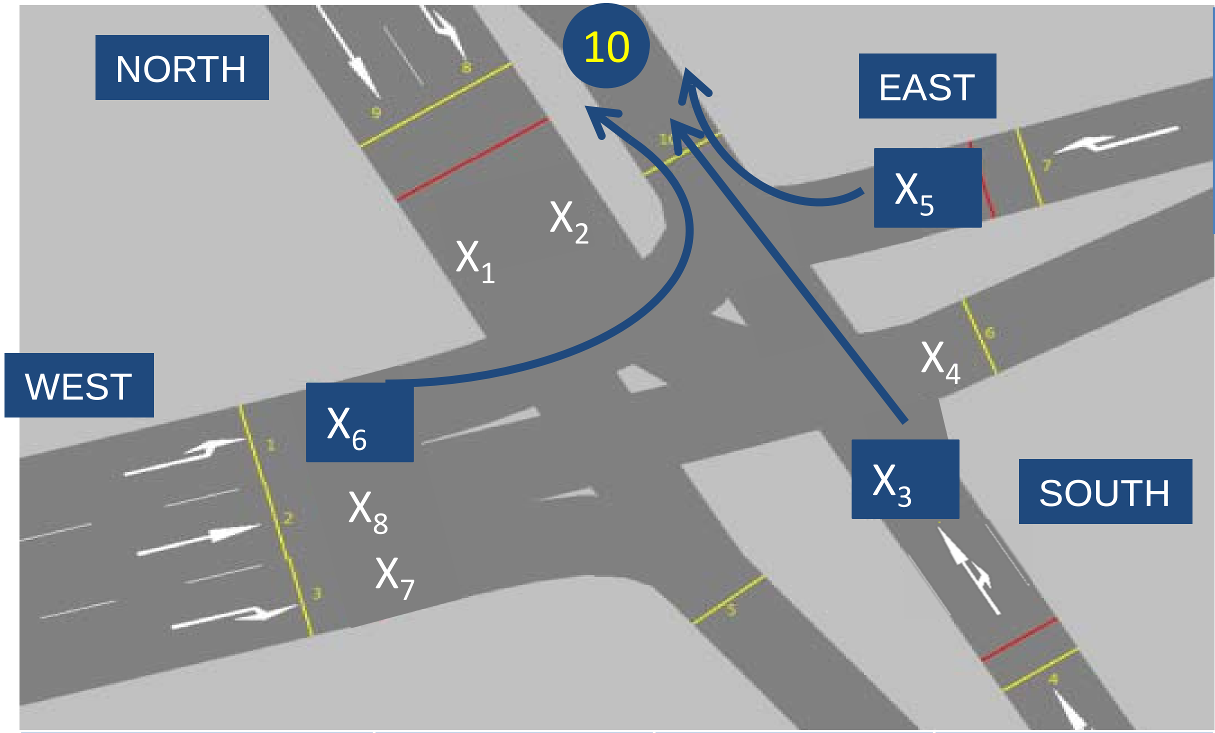




O/D	N	S	E
N	-	x_1	x_2
S	x_3	-	x_4
E	x_5	-	-
W	x_6	x_7	x_8

2.2. BILEVEL OPTIMIZATION PROCESS (BO)

•2.2.4. RELATIONSHIPS BETWEEN THE LOOP DETECTORS AND MOVEMENTS



Loop 1= x_6	Loop 6=x_2+x_4+x_8
Loop 2=x_8	Loop 7=x_5
Loop 3=x_7	Loop 8=x_2
Loop 4=x_3+x_4	Loop 9=x_1
Loop 5=x_1+x_7	**Loop 10=x_3+x_5+x_6**

Loop detectors/movement	N	S	E
N	-	loop9, loop5	loop8, loop6
S	loop4, loop10	-	loop4, loop6
E	loop7, loop10	-	-
W	loop1, loop10	loop3, loop5	loop2, loop6

3.1. ANALYSIS OF RESULTS

To validate the method for different data availability levels, and to generalize the results, the example covers different scenarios:

(i) **different FCD-PR**, from 5% to 100%

(ii) **different LDC levels,** from 100% to 0%, corresponding to realistic situations:

- incomplete loop detector data (malfunctioning loop detectors, or no loop detectors installed).

(iii) For each scenario, **500 random iterations** are performed.

(iv) Within each iteration, random vehicles / loop detectors are selected based on the hypothesis about data availability corresponding to that specific scenario.

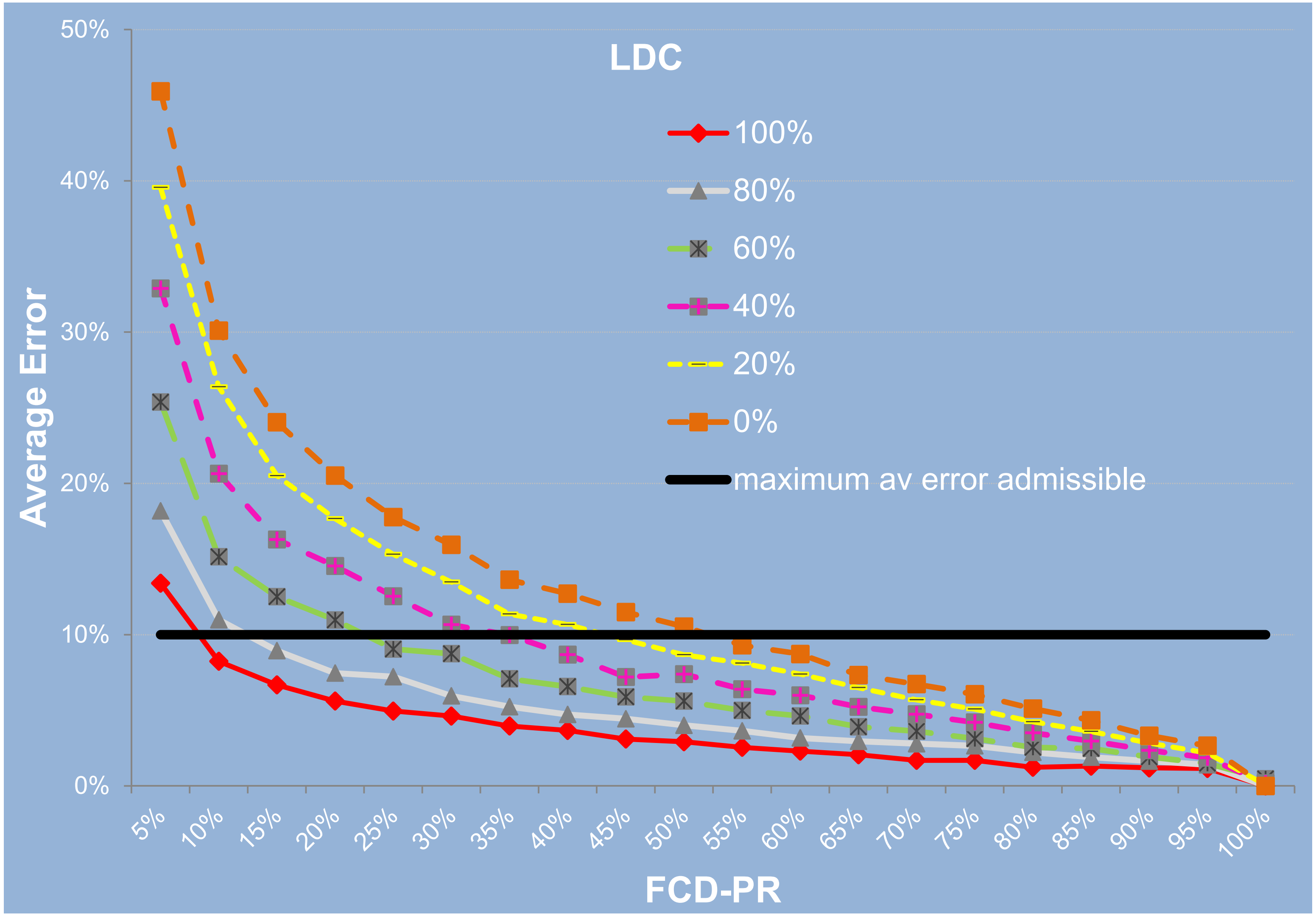
These variations in FCD-PR and LDC yield **220 different scenarios**

3.2. EXAMPLE RESULTS FOR ALL SCENARIOS

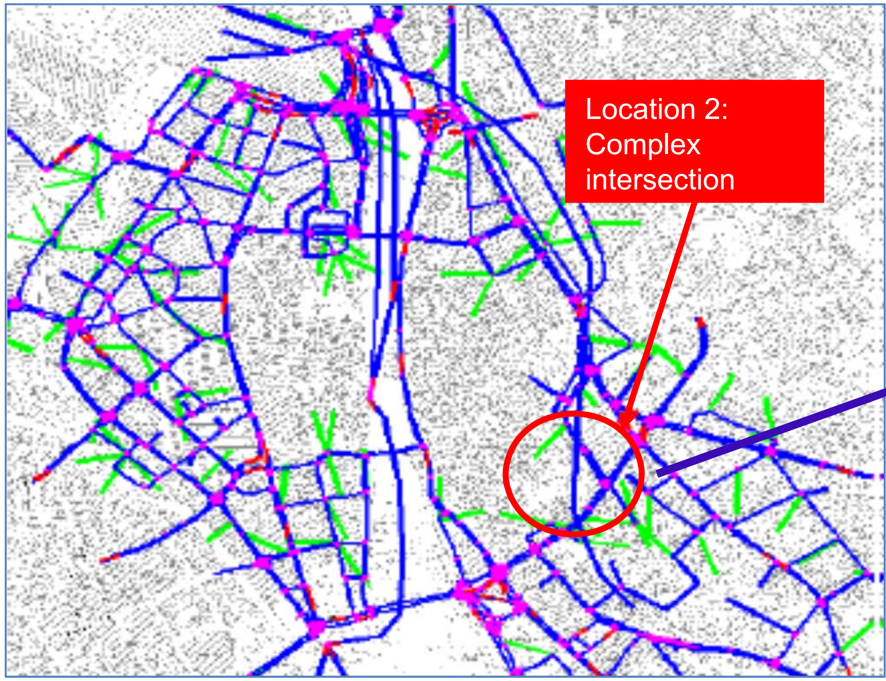
AVERAGE RELATIVE ERROR		LOOP DETECTOR AVAILABILITY										
		100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	0%
FCD PENETRATION RATE	5%	13%	15%	18%	21%	25%	30%	33%	36%	40%	42%	46%
	10%	8%	9%	11%	13%	15%	19%	21%	23%	26%	28%	30%
	15%	7%	8%	9%	11%	13%	14%	16%	19%	21%	23%	24%
	20%	6%	7%	7%	9%	11%	12%	15%	15%	18%	19%	21%
	25%	5%	6%	7%	8%	9%	10%	13%	14%	15%	16%	18%
	30%	5%	5%	6%	7%	9%	9%	11%	12%	13%	15%	16%
	35%	4%	5%	5%	6%	7%	8%	10%	10%	11%	12%	14%
	40%	4%	4%	5%	6%	7%	8%	9%	9%	11%	12%	13%
	45%	3%	4%	4%	5%	6%	7%	7%	9%	10%	10%	11%
	50%	3%	3%	4%	4%	6%	6%	7%	8%	9%	9%	11%
	55%	3%	3%	4%	4%	5%	6%	6%	7%	8%	9%	9%
	60%	2%	3%	3%	4%	5%	5%	6%	6%	7%	8%	9%
	65%	2%	3%	3%	3%	4%	4%	5%	6%	7%	7%	7%
	70%	2%	2%	3%	3%	4%	4%	5%	5%	6%	6%	7%
	75%	2%	2%	3%	3%	3%	4%	4%	5%	5%	6%	6%
	80%	1%	2%	2%	2%	3%	3%	4%	4%	4%	5%	5%
	85%	1%	2%	2%	2%	2%	3%	3%	3%	4%	4%	4%
	90%	1%	1%	2%	2%	2%	2%	2%	3%	3%	3%	3%
	95%	1%	1%	1%	1%	1%	2%	2%	2%	2%	2%	3%
	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

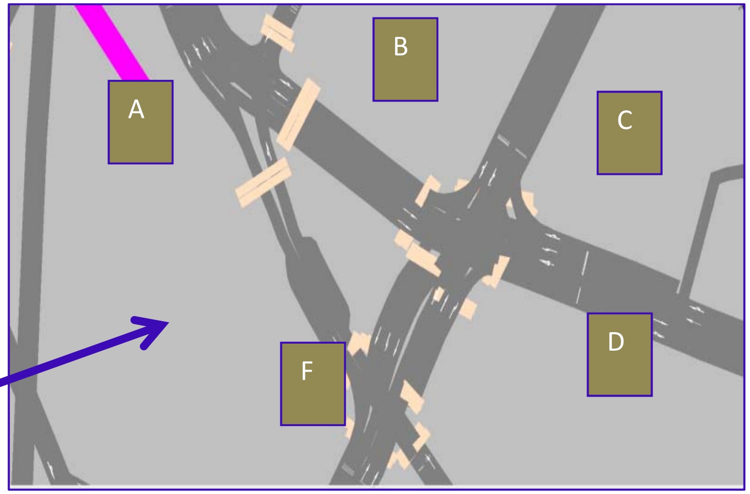
On average, if loop detectors cover 30% of the lanes, a **40% FCD-PR should be enough to provide an OD matrix** with an admissible error

3.2. EXAMPLE RESULTS FOR ALL SCENARIOS

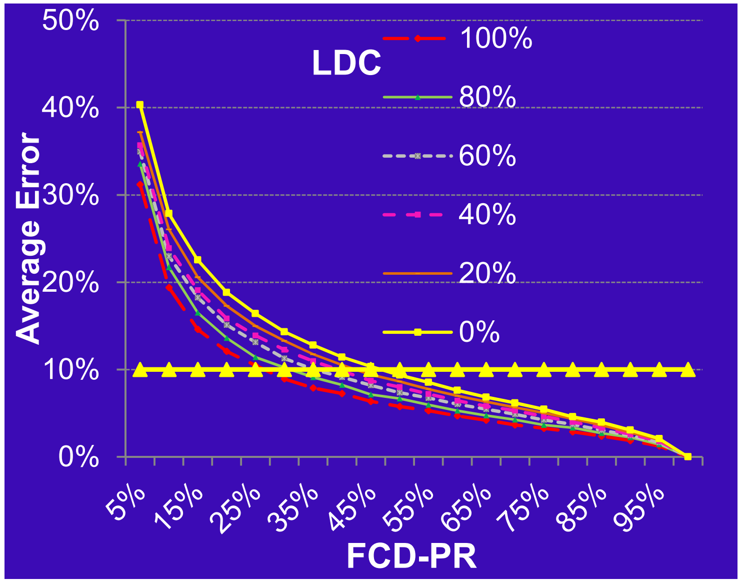


3.3. TO VALIDATE THE RESULTS





O/D	A	B	C	D	E	F
A	-	x_1	-	x_2	x_3	x_4
B	x_5	-	-	-	-	-
C	x_6	x_7	-	-	x_8	x_9
D	x_{10}	x_{11}	x_{12}	-	x_{13}	x_{14}
E	x_{15}	x_{16}	-	-	-	-
F	-	-	x_{17}	x_{18}	-	-



3. MAIN CONCLUSIONS

CONCLUSIONS

The method is able to:

✓Preserve the **integrity** of the observed data as much as possible.

✓**Ensure flow consistency** at any point in the network.

✓Handle data **reliability** (alpha parameter).

✓Be solved in a **short computation time**.

✓The **method is able to detect inadmissible errors in TCS** and **identifying which device is more likely to be failing**.

✓The **only incoming data required** are the observed data, with **no need** for pre-processing using **historical data,** although they could be complementary.

✓ The model is applied in **adjusting boardings and alightings** in a transit line.

✓ Novelty: **Subjective information is used** as an input.

CONCLUSIONS

GENERAL CONCLUSIONS:

- Another contribution is in the field of Data Imputation and Base Data Integrity.
- The method is based on a double linear optimization process that can easily be performed with existing software on the market, and which we consider highly useful for practitioners.
- The proposed approach is robust enough to deal with other common data discrepancies in any engineering problem.

4. RESEARCH GOING ON

PRIMARY CONTROL OF FREQUENCY IN A GRID

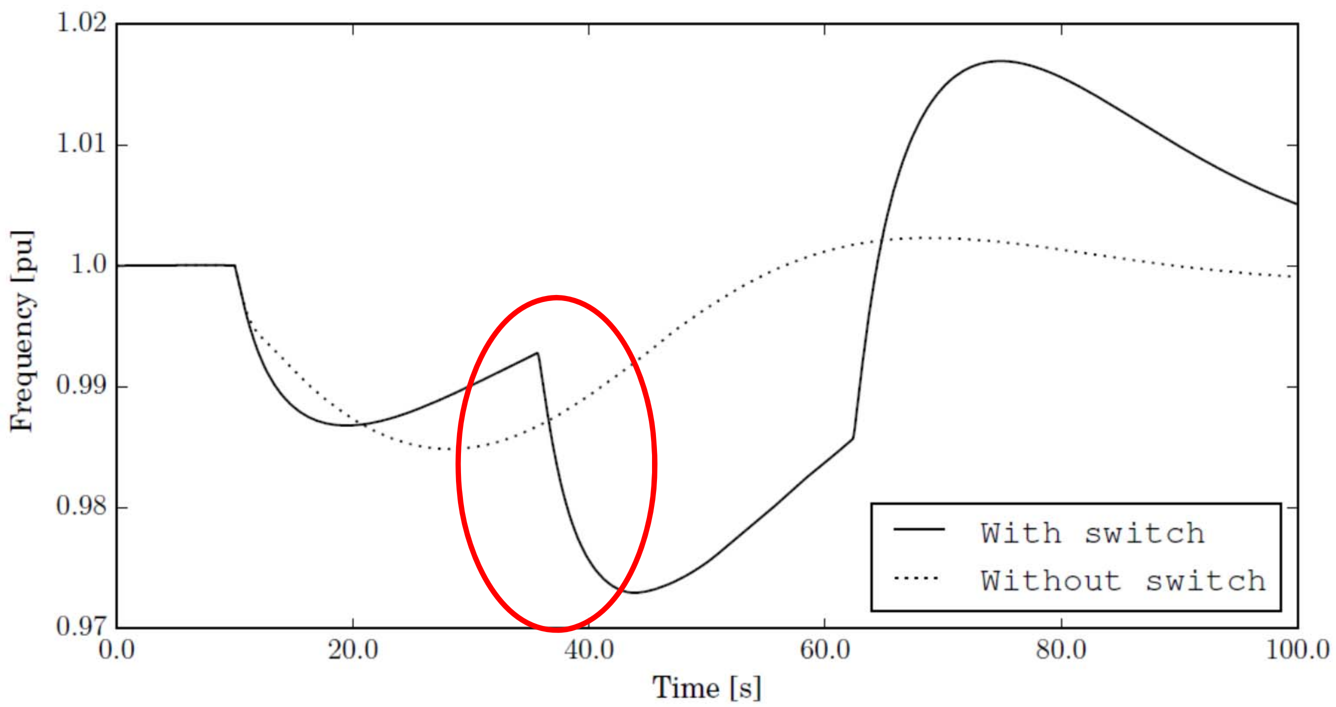
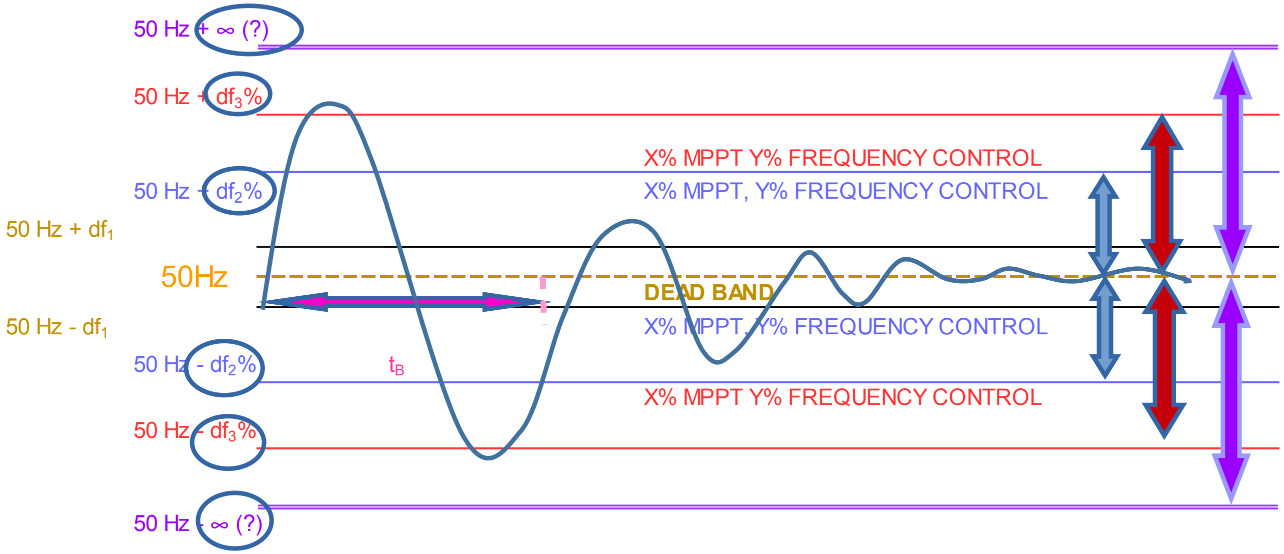


Fig. 4. Frequency of the COI of the WSCC system considering the effect of the MPPT switch coupled with the droop & ROCOF controls.

PRIMARY CONTROL OF FREQUENCY IN A GRID



Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

THANK YOU FOR YOUR ATTENTION

For more info:
penelopegj@uma.es

Dublin, March the 24th 2017