Dissecting traffic flows in congestion areas using GPS data

Feng Liu, Muhammad Adnan, Ansar yasar, Davy Janssens

Institute for MOBility (*IMOB*), Hasselt University, Belgium

Track & Know

Presentation Structure

- Problem statement and the goal of this study
- GPS data
- Methodology
- Preliminary results

Track & Know Introduction

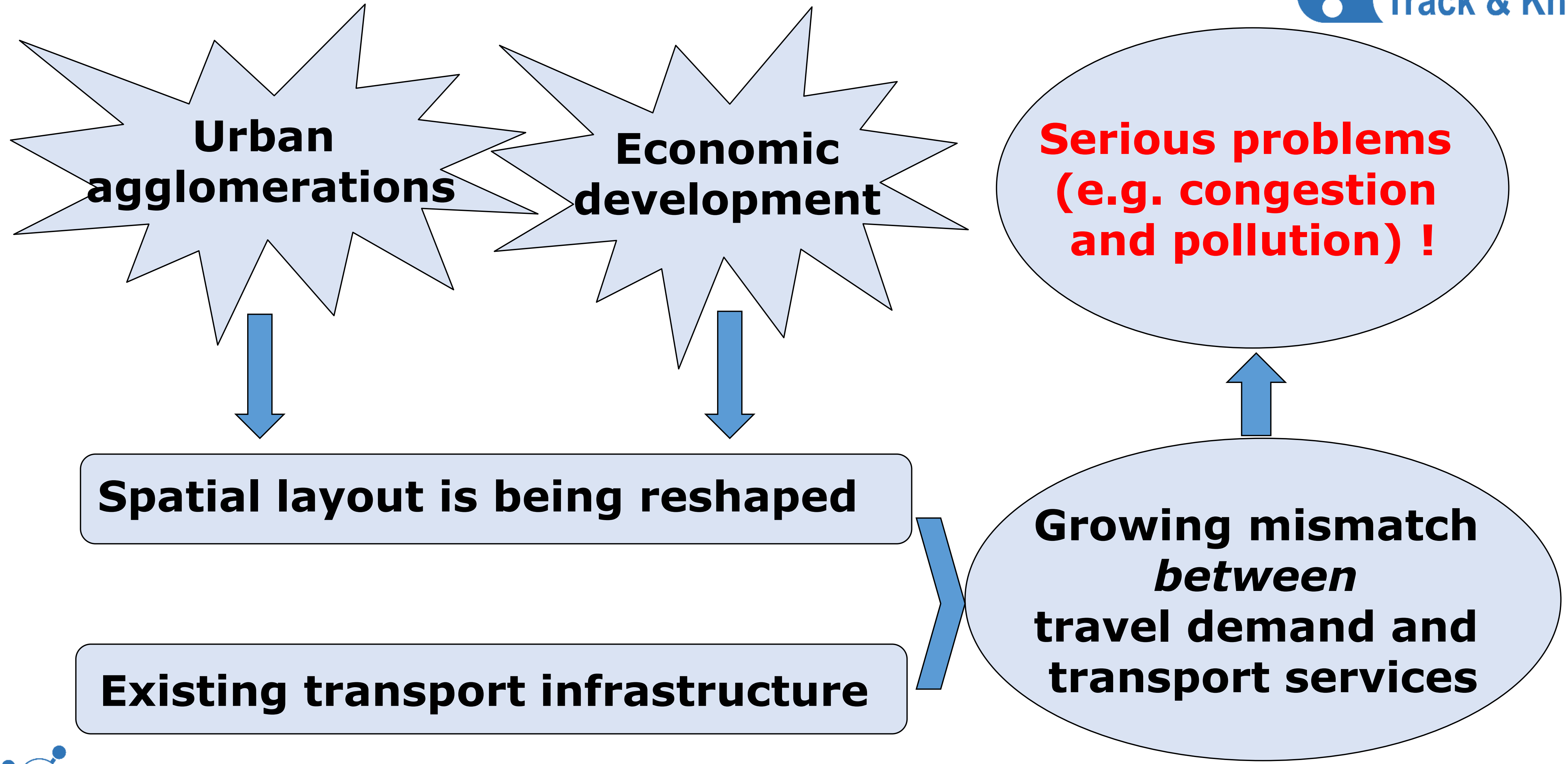
The Track & Know platform is designed to provide a variety of tools that can operate stand alone or in conjunction with each other.

- **Track & Know Big data Platform**
- **Track & Know Tool boxes**
  - **Big data Processing**
  - **Big data Analytics**
  - **Big data Visual Analytics**
- **Big data Use Cases**
  - **Three Pilots use cases; Mobility, Insurance and Health**

The Mobility use case (Vodafone Innovus (VFI)) can be summed into 3 categories:

- **Error detection and correction (GPS location, speed etc.)**
- **Route features aggregation and prediction (trajectory statistics, hot spots, prediction of a route and its components like average speed, mileage, fuel consumption)**
- **Driver behavior (categorize, correlate road condition with fuel and driver behavior)**

Problem statement



The Goal

A method *to*

- ***Analyze*** **mobility demand and underlying transport systems**
- ***Identify*** **areas with serious mismatch problems (congestion)**
- ***Dissect*** **traffic flows in the detected congestion areas**
- ***Find*** **alternative routes and** ***Suggest*** **policy measures**

Presentation Structure

- Problem statement and the goal of this study
- GPS data
- Methodology
- Preliminary results

GPS data:

- **Data: by Vodafone Innovus**

❑ **Location based data**
❑ **7500 vehicles**
❑ **216GB**
❑ **Period: 01/06/2017 – 01/06/2018**
❑ **Vehicle types (15): Passenger Car, Bus, Van,…**

- **Data used for the preliminary results:**

❑ **12,213 Files under 55 Folders**
❑ **533 vehicles (7.1% of the total)**
❑ **3.46 GB**
❑ **Period: 08/10/2017 – 12/10/2018 (5 weekdays)**

Presentation Structure

- Problem statement and the goal of this study
- GPS data
- Methodology
- Preliminary results

Methodology:

(5 steps!)

Raw GPS data

↓

Trip extraction

↓

City-wide travel pattern modeling

↓

Congestion zone identification

↓

Traffic flow dissecting

↓

Alternative route searching

Track & Know

Track & Know

Step 1: Trip extraction

- **Each GPS point: $p_i(lat_i, lon_i, t_i, v_i, engineStatus_i, vehicle_i, vehicleT_i)$**

Stop points ($engineStatus_i$=0 and v_i=0)

Non-stop points

Trips ($Distance > TH_{trip}$)

$Distance > TH_{trip}$?

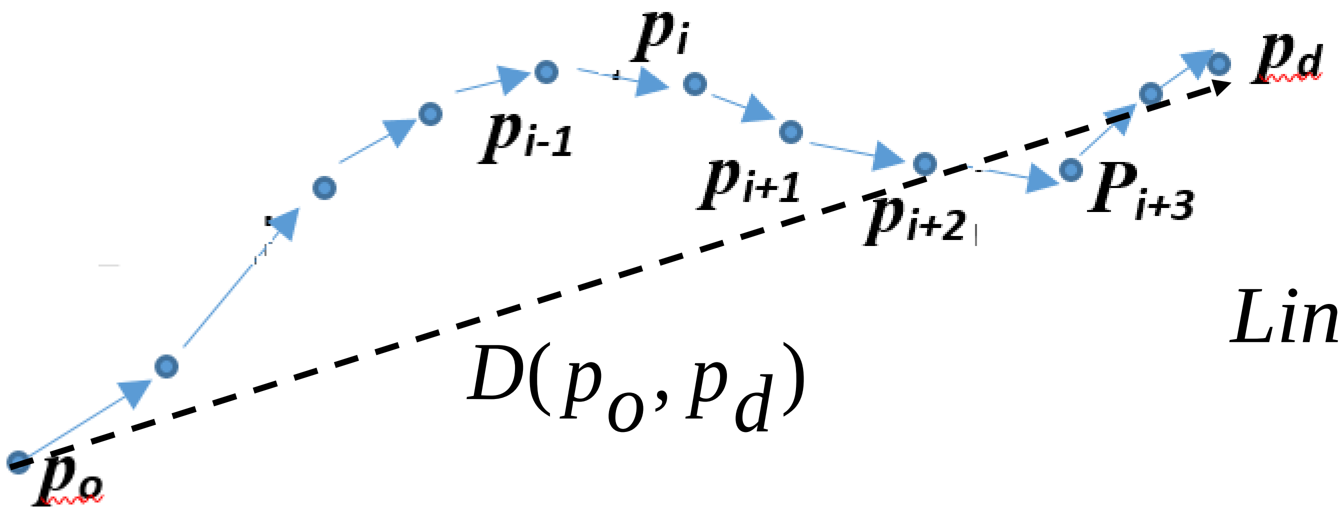
$Distance > TH_{trip}$?

$$p_1 \text{->} \; p_2 \text{->} \ldots \; p_i \text{->} \; p_{i+1} \ldots \text{->} \; p_j \text{->} p_{j+1} \; \ldots \text{->} p_n$$

....

non-stop points **stop points** **non-stop points**

Track & Know



$$Link(p_o, p_d) = \sum_{i=o}^{d-1}[D(p_i, p_{i+1})$$

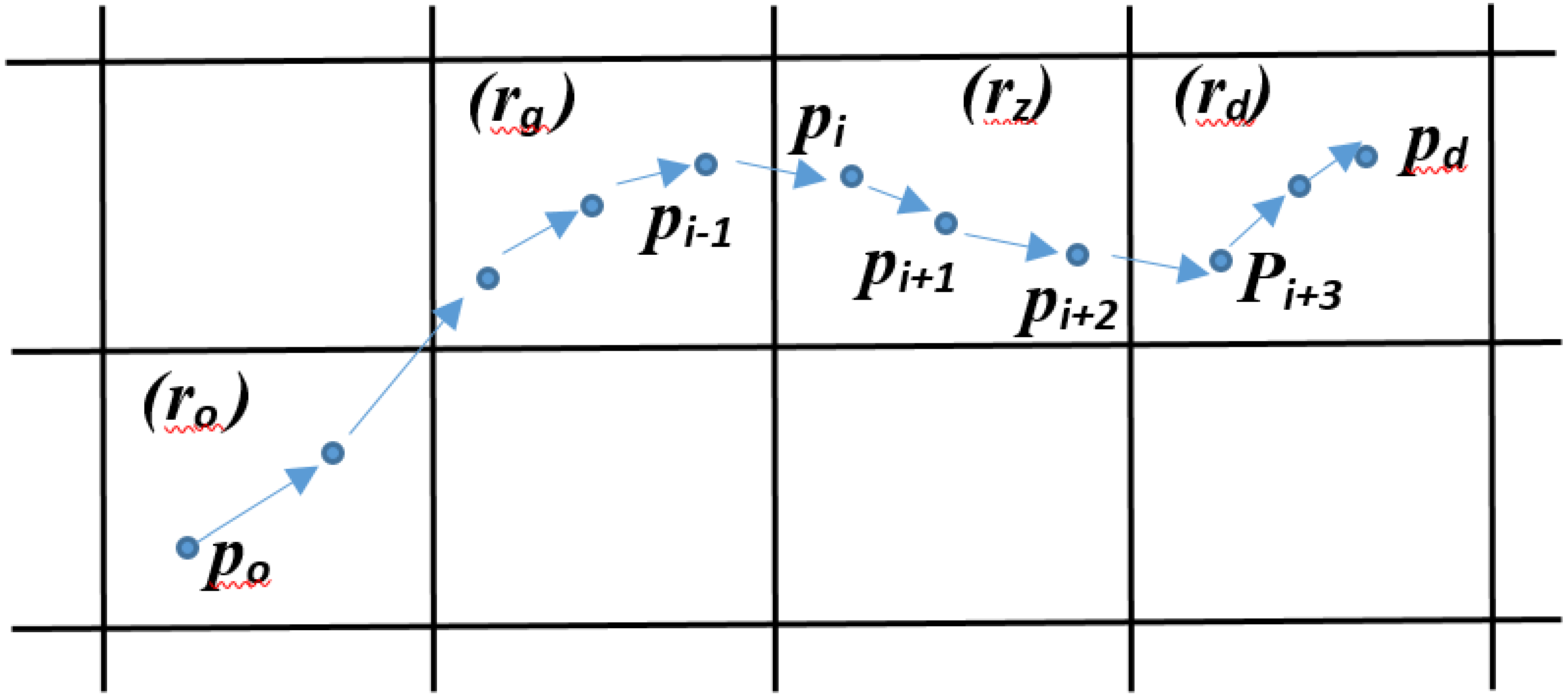
Travel time: $$TTrip = t_d - t_o$$

Track & Know

Step 2: City-wide travel pattern modeling

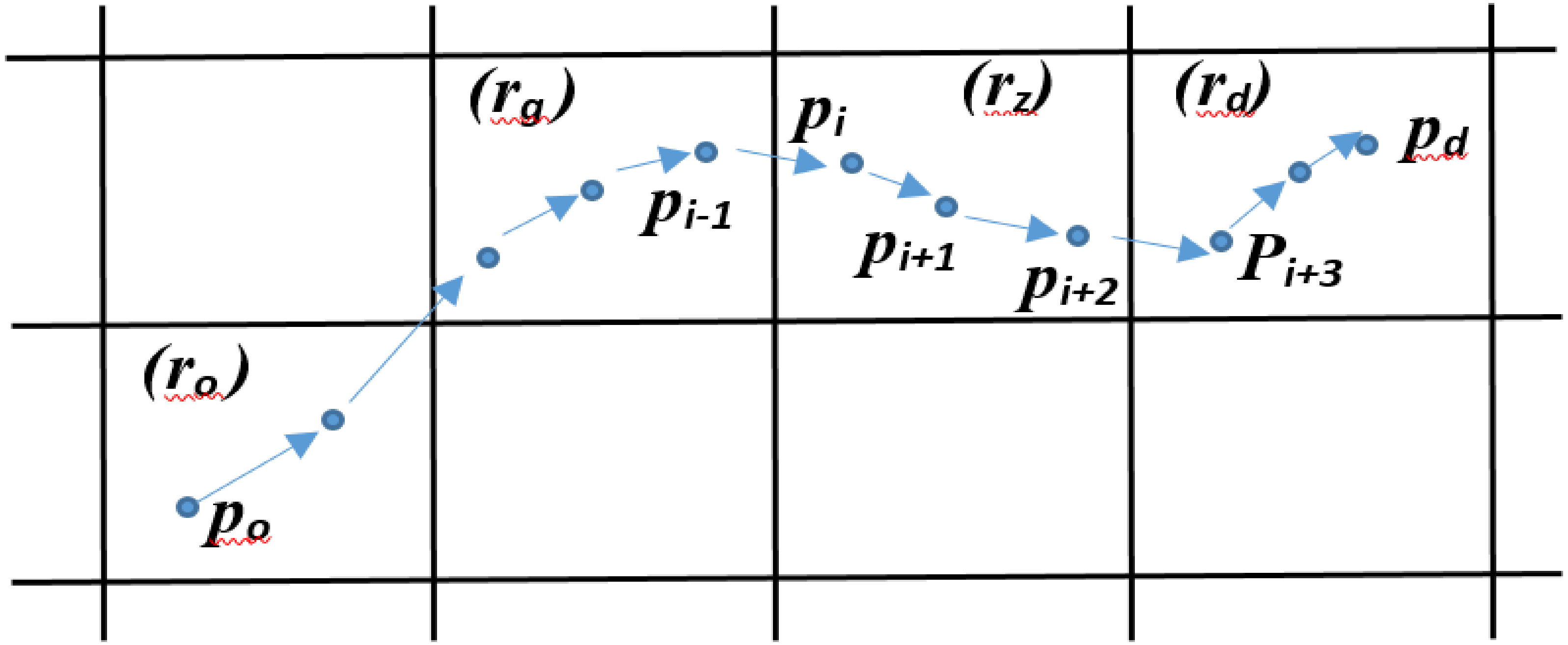
- **Spatial partition**
  - **zones: *Grid_X* × *Grid_Y***
  - r_z, $z=1 \ldots Grid\_X \times Grid\_Y$

- **Temporal partition**
  - ***TimeP:* different time periods of a day**
  - ***DayT:* different types of the day**



- **Zone-traffic-condition-matrix:** ***Z(r$_z$, TimeP, Day, DayT)***
- **OD-travel-pattern-matrix:** ***OD(r$_o$, r$_d$, TimeP, Day, DayT)***

Zone-traffic-condition-matrix *Z(r_z, TimeP, Day, DayT)*: *average driving speed of each zone*

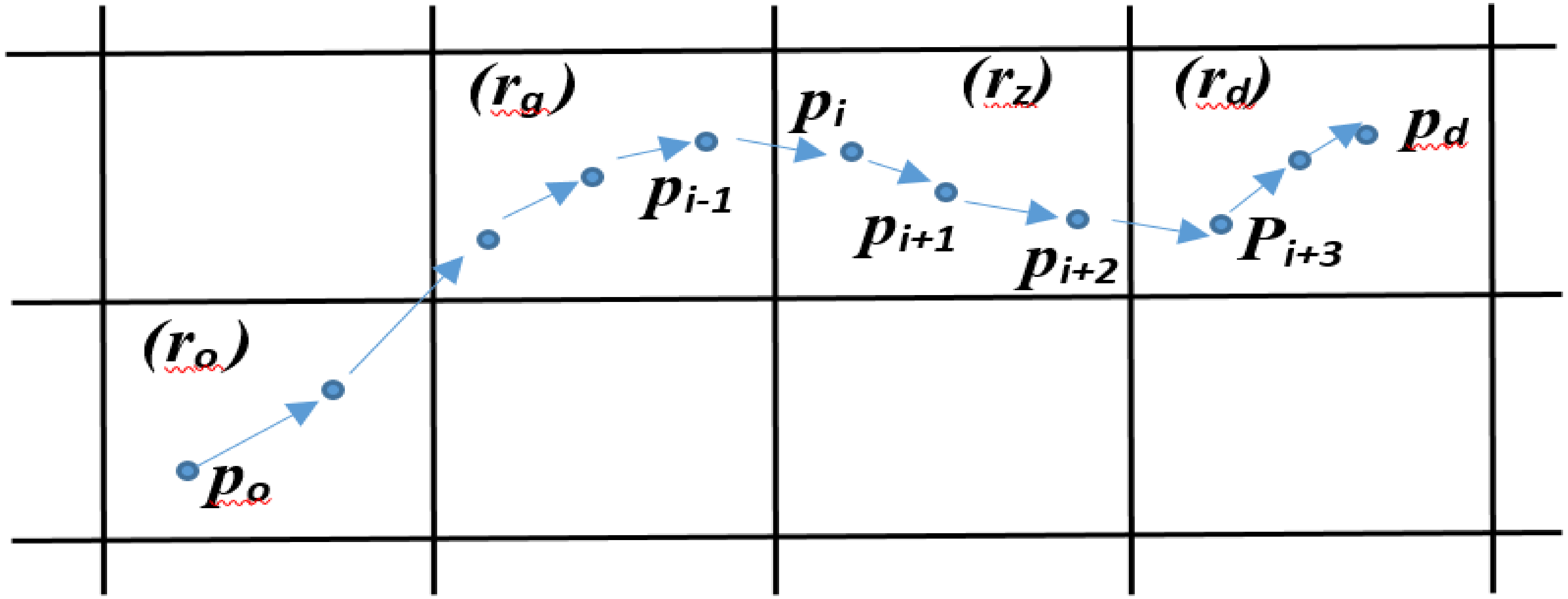


- **Each matrix element**

❑ **Total number of points M_z in r_z**

❑ **Average speed of the points** $V_Z \widehat{=} V_Z(r_Z, TimeP, Day, DayT) = \dfrac{\sum_{k=1}^{M_Z} v_Z^{(k)}}{M_Z}$

OD-travel-pattern-matrix *OD(r_o, r_d, TimeP, Day, DayT)*: travel demand between origin and destination zones



- **Each matrix element**

❑ **Total number of trips between r_o and r_d**

❑ **Average travel time, speed and route directness over all the trips**

Average travel time:

$$U_{o,d} \mathrel{\widehat{=}} U_{o,d}(r_o, r_d, TimeP, Day, DayT) = \frac{\sum_{k=1}^{M_{o,d}} U_k}{M_{o,d}}$$

Average travel speed:

$$V_{o,d} \mathrel{\widehat{=}} V_{o,d}(r_o, r_d, TimeP, Day, DayT) = \frac{\sum_{k=1}^{M_{o,d}} V_k}{M_{o,d}}$$

Average route directness:

$$R_{o,d} \mathrel{\widehat{=}} R_{o,d}(r_o, r_d, TimeP, Day, DayT) = \frac{\sum_{k=1}^{M_{o,d}} R_k}{M_{o,d}}$$

Step 3: Congestion zone identification

- **Zone-traffic-condition-matrix: $Z(r_z, TimeP, Day, DayT)$**

- **Zones with congestion on a day**

❑ **Total number of points: $M_z > TH_{MZ}$**

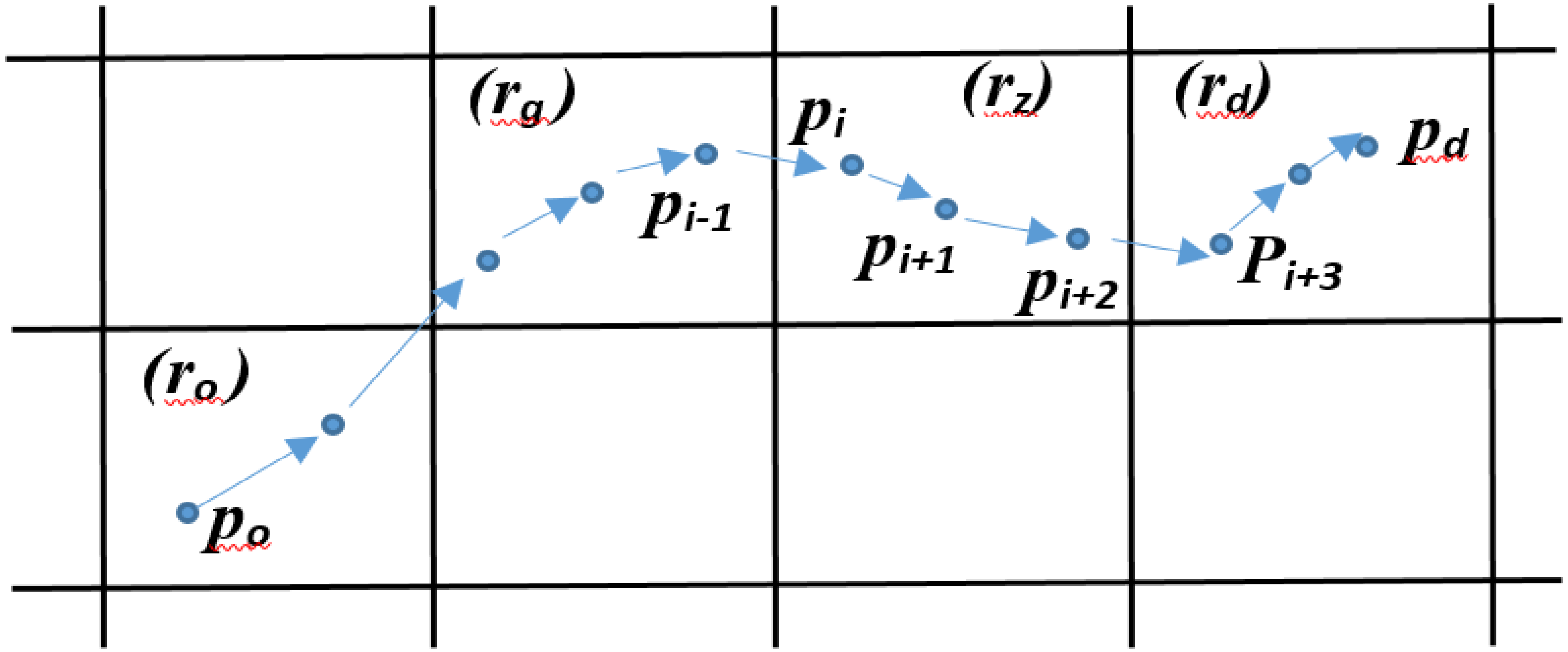
❑ **Average speed: $V_z < TH_{VZ}$**

- **The probability of congestion on multiple days**

❑ **$P > TH_P$**

Step 4: Traffic flow dissecting in a congestion area

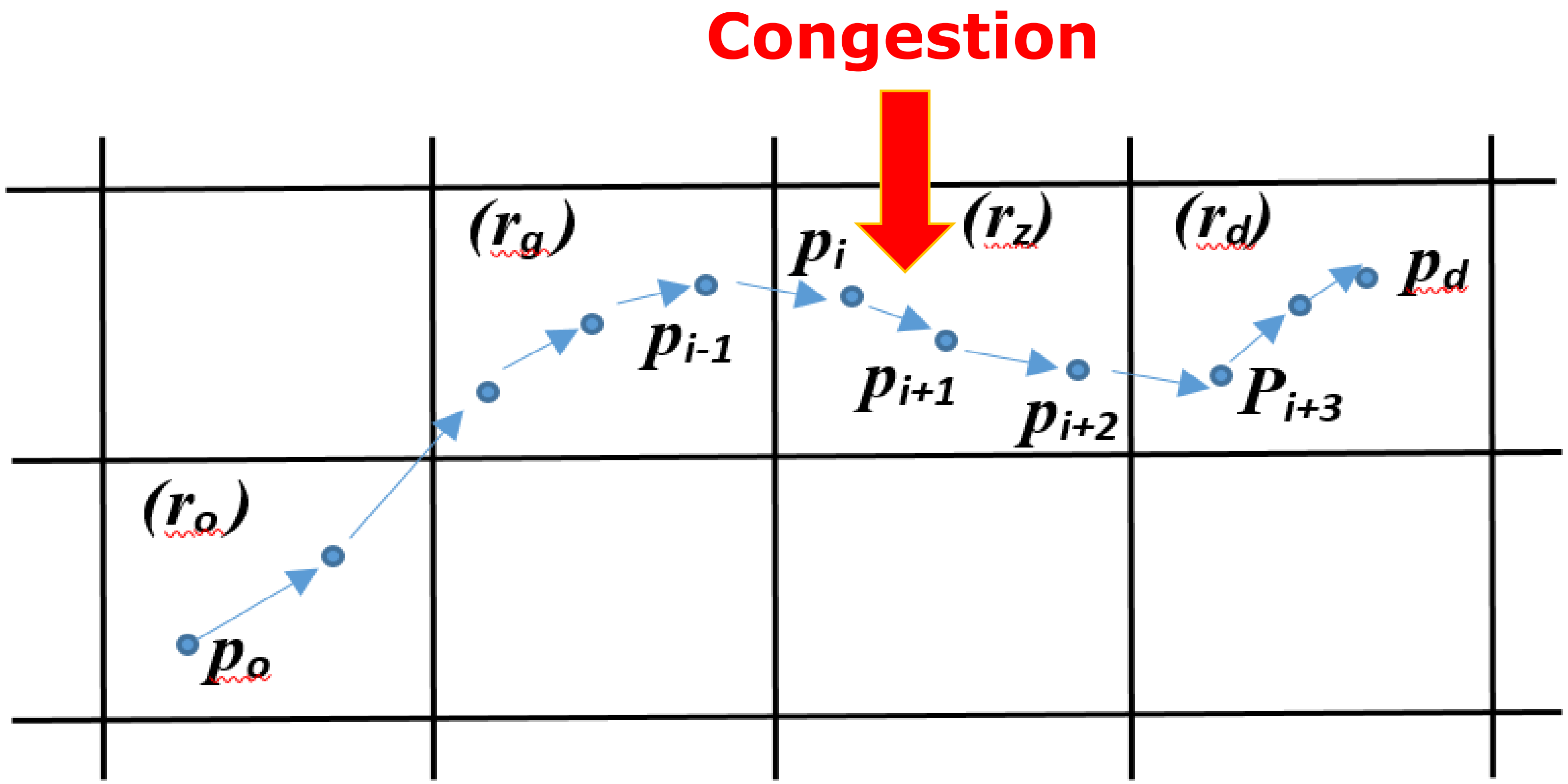
- OD-travel-pattern-matrix: *OD(r_o, r_d, TimeP, Day, DayT)*
- Project each OD trips into zones



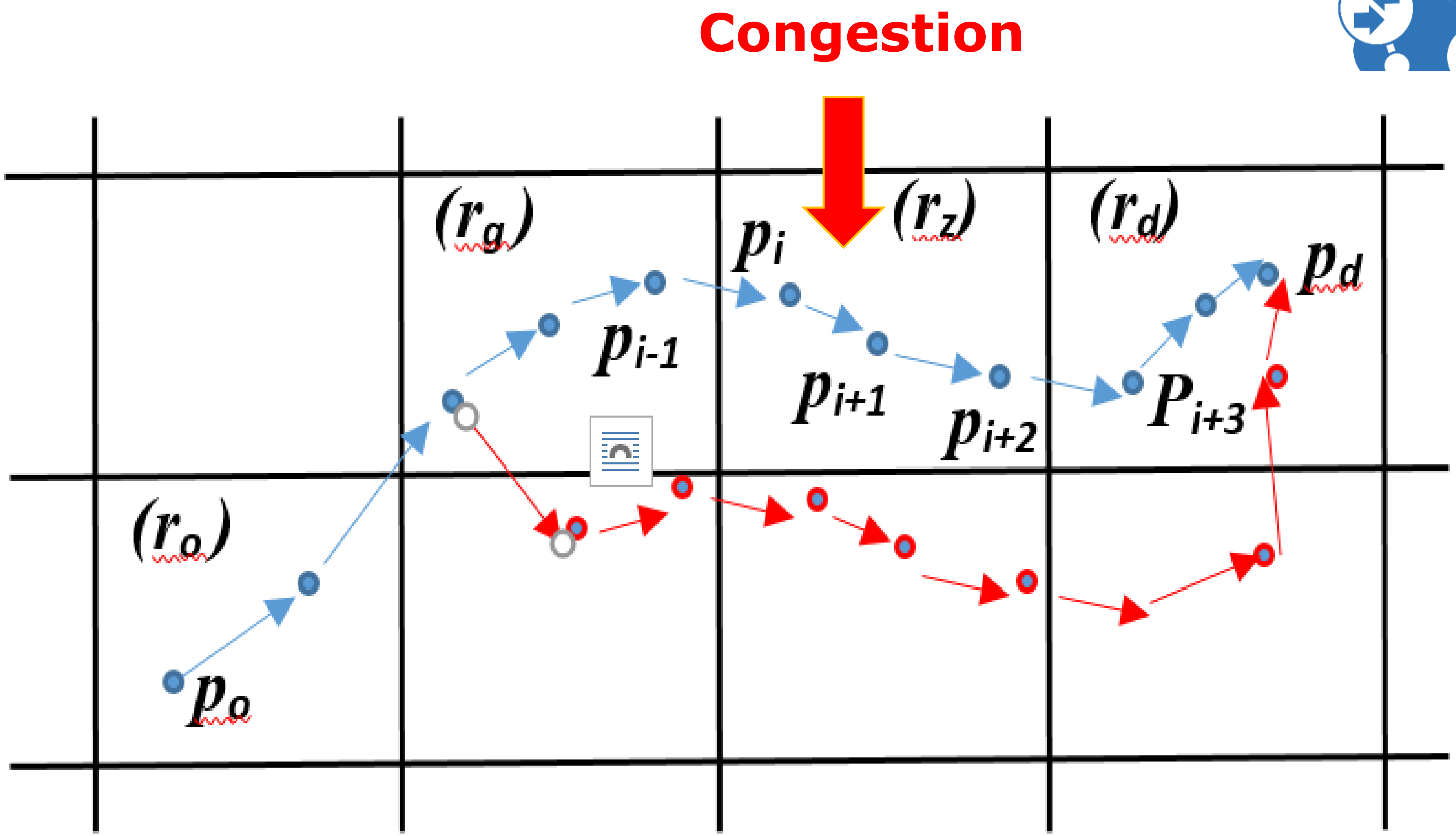
p_o-> ...p_i->p_{i+1}->...->p_d ⇒ r_o-> r_g->r_z-->r_d

Step 5: Alternative route searching

- **Search for alternative routes that avoid the congestion zones**



Track & Know


Congestion
(r_g)
p_i
(r_z)
(r_d)
p_d
p_{i-1}
p_{i+1}
p_{i+2}
P_{i+3}
(r_o)
p_o

Presentation Structure

- **Problem statement and the goal of this study**
- **GPS data**
- **Methodology**
- **Preliminary results**

Result 1: Trips

- 23934 trips over 533 vehicles and 5 weekdays
- 8.98 trips/vehicle and day

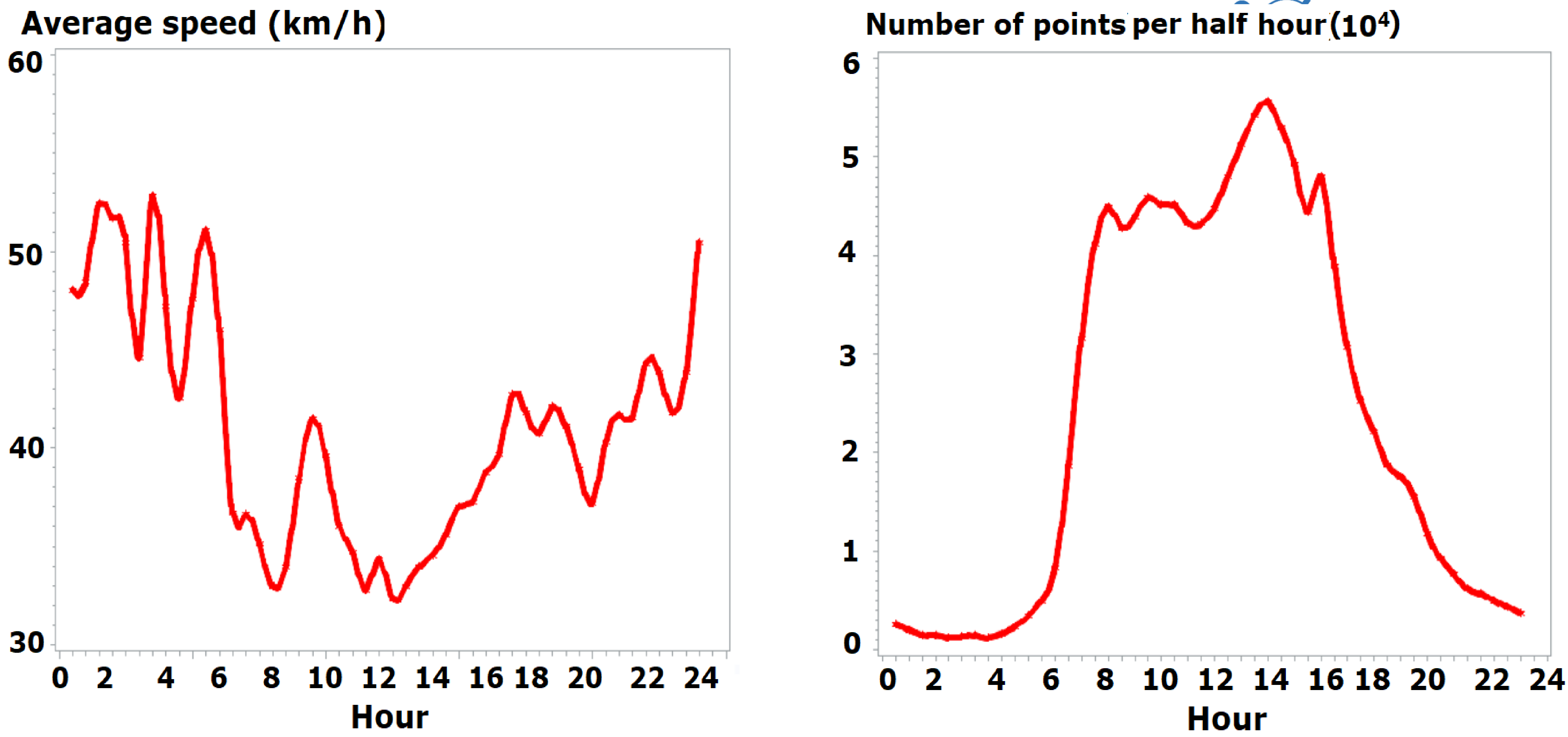


Fig. Distribution of average speeds and number of trips per half an hour

Track & Know

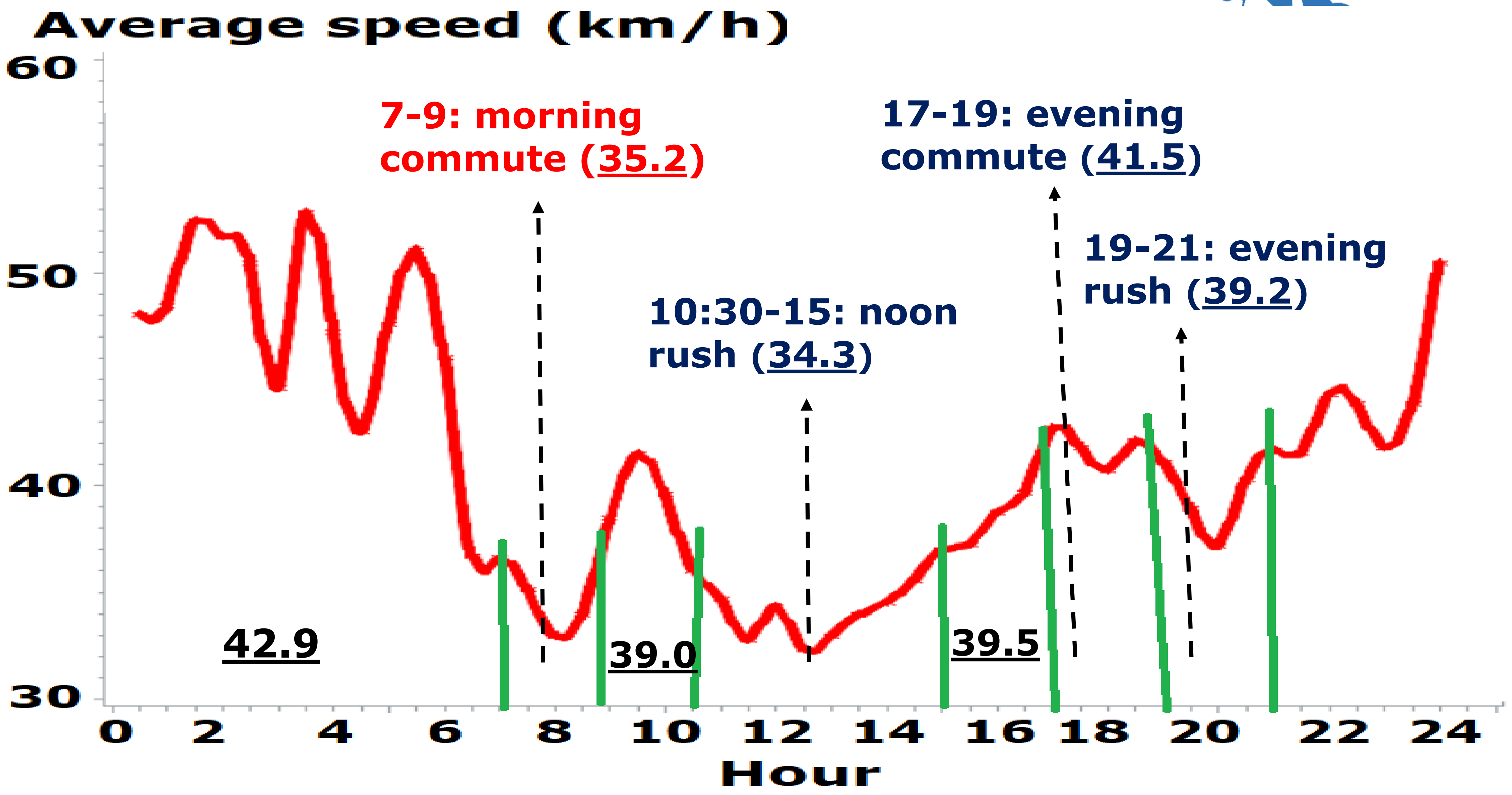


Fig. Seven Time periods

Track & Know

Result 2: City-wide travel pattern modelling

- **Zone-traffic-condition-matrix:** $Z(r_z, TimeP, Day, DayT)$
- **OD-travel-pattern-matrix:** $OD(r_o, r_d, TimeP, Day, DayT)$

- ❑ **z=1,…,400x400, Each zone: 1.62 km^2**
- ❑ **TimeP=1…7**
- ❑ **Day=1…5**
- ❑ **DayT=weekdays, weekends and holidays**

In this case study, morning commute (7-9am) on weekdays!

Result 3: Congestion zones (M_z >50/day, V_z < 20, and $P \geqslant 0.8$)

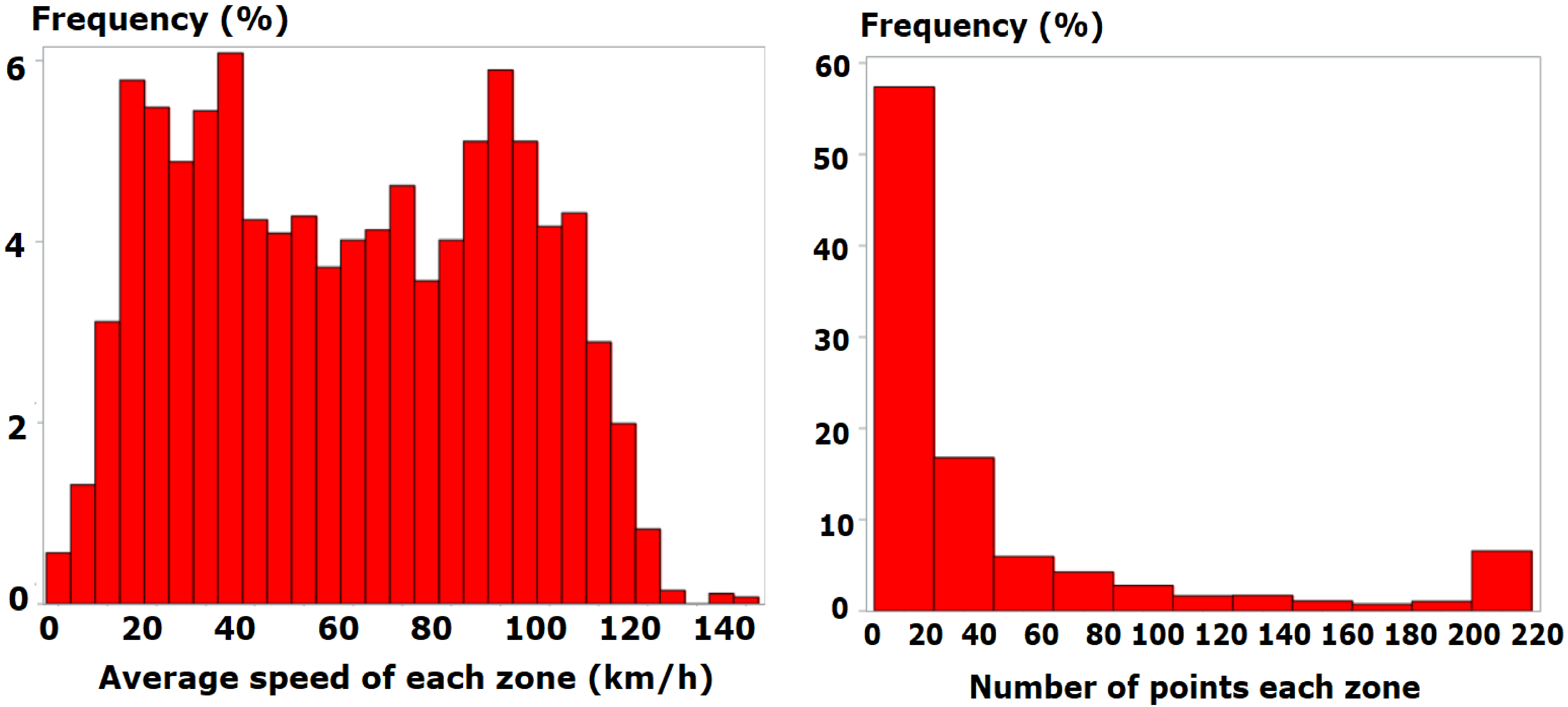


Fig. The average speed and number of points each zone in the morning

Track & Know

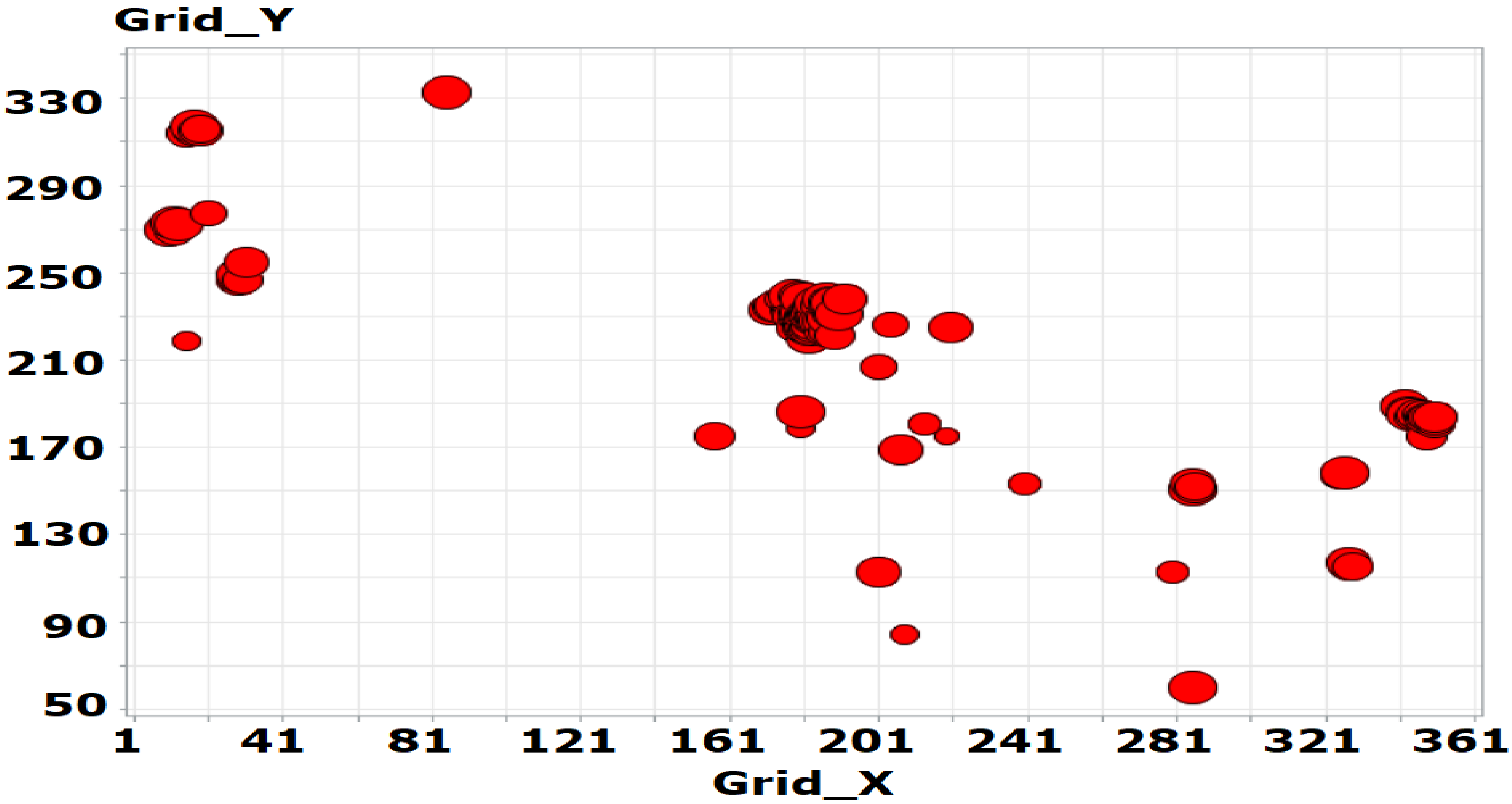
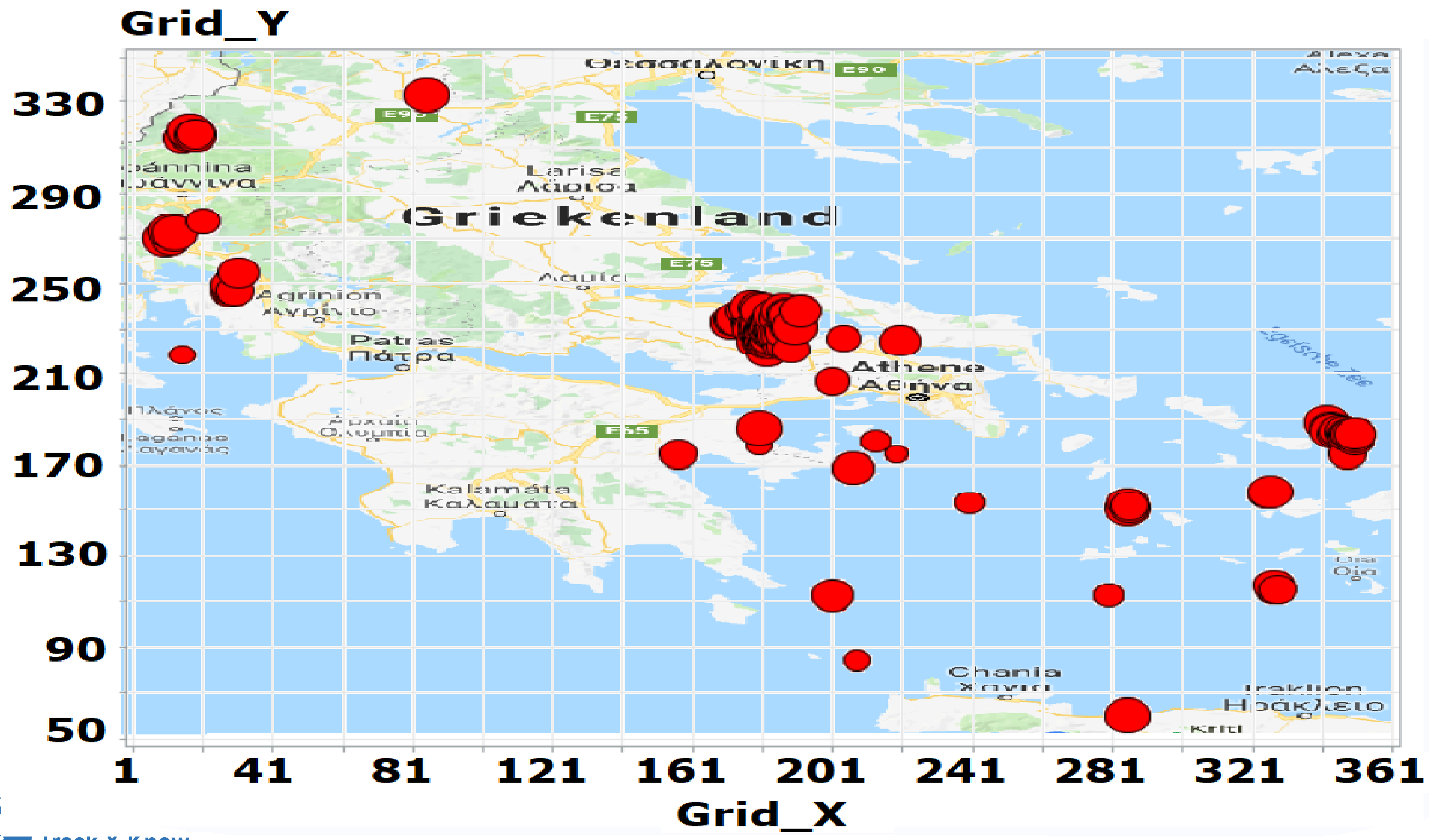


Fig. Congestion zones (125)

Track & Know

now



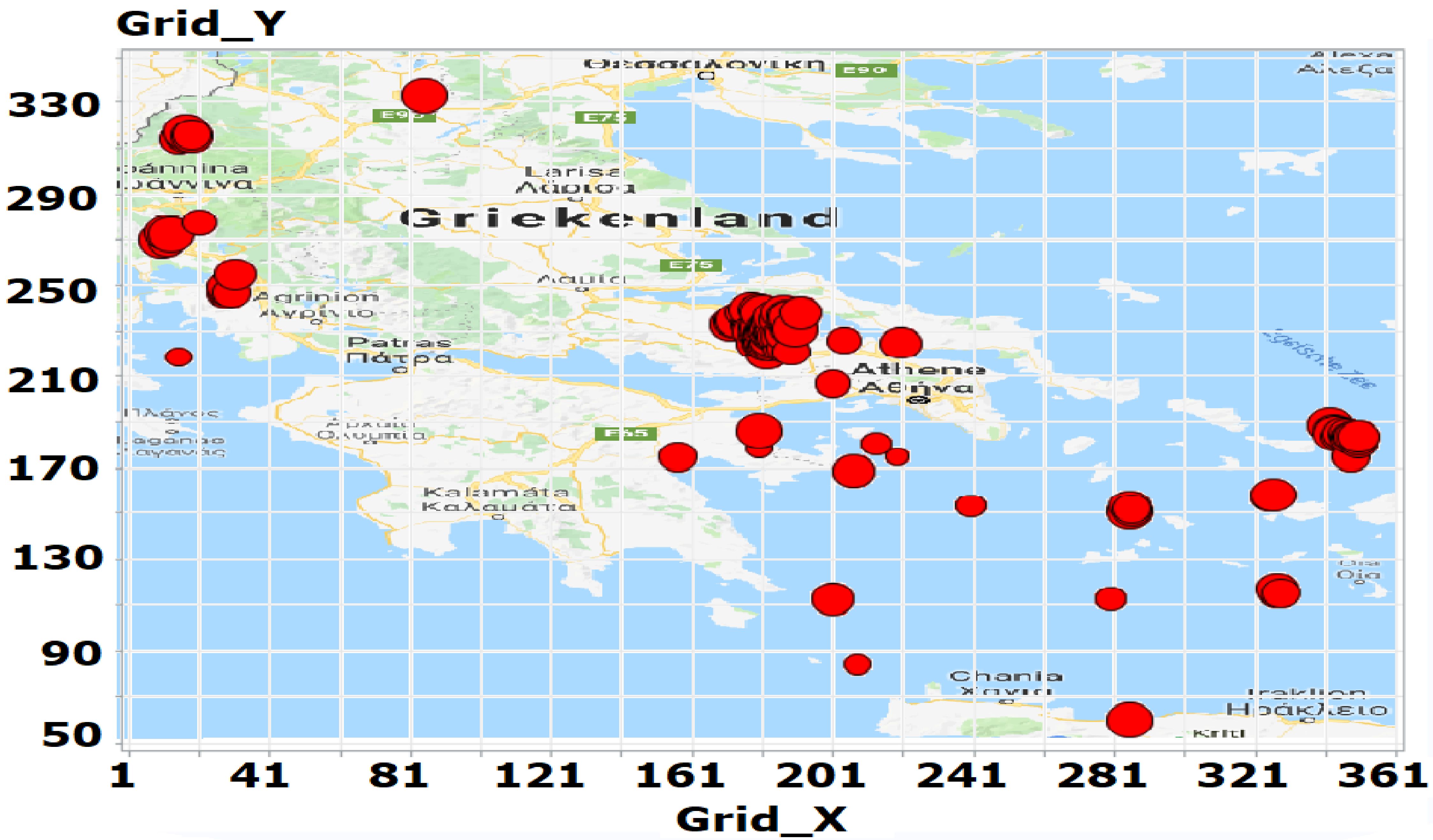
Track & Know

Result 4: Traffic flow dissecting

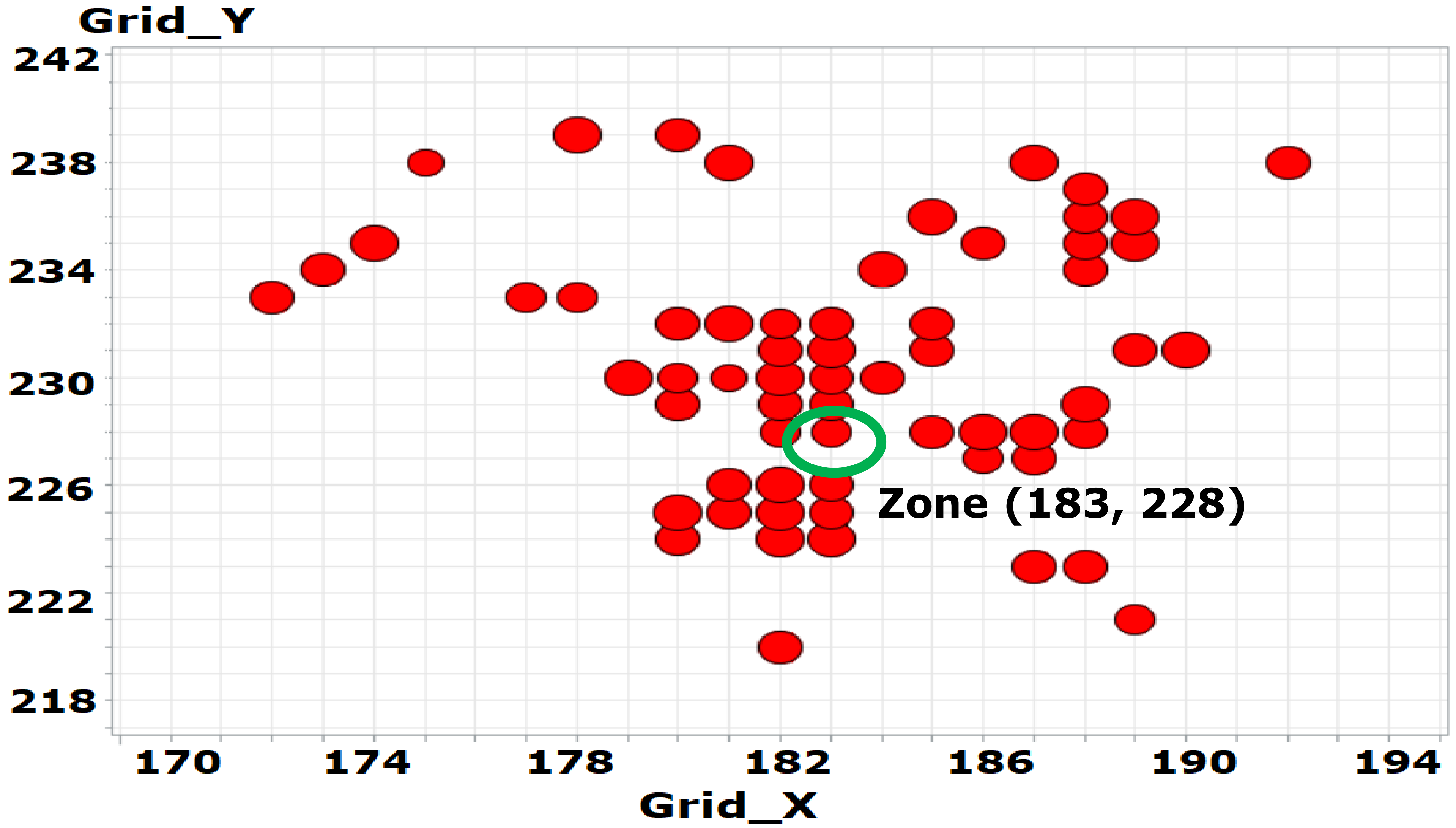
	Region_x_id_old	Region_y_id_old	Average_Speed_Zone	Number_Of_Points_Zone
1	219	175	5.7049180328	183
2	15	219	6.2388059701	67
3	328	114	6.6515151515	66
4	180	179	6.8055555556	72
5	208	84	6.8709677419	62
6	213	181	7.8961038961	77
7	280	113	8.1206896552	116
8	240	153	8.84375	64
9	377	217	9.74	50
10	204	226	10.245283019	53
11	201	207	10.324324324	74
12	175	238	10.431034483	58
13	21	277	10.825396825	63
14	369	284	11.053571429	56
15	181	230	11.205479452	219
16	157	175	12.696202532	79
17	178	233	12.722627737	137
18	328	115	12.737179487	156
19	182	232	12.90376569	239
20	348	175	12.98265896	173
21	350	182	13.130434783	115
22	348	184	13.226315789	190
23	286	152	13.4	55
24	346	184	13.436363636	55
25	286	151	13.592592593	81
26	327	115	13.755102041	735
27	177	233	13.837719298	228
28	19	316	14.207792208	77
29	345	184	14.23655914	186
30	30	247	14.381578947	76
31	375	307	14.407407407	81
32	325	157	14.465753425	73
33	15	314	14.510869565	92
34	347	[illegible]	[illegible]	135
35	183	228	14.568668407	1915
36	182	[illegible]	[illegible]	471

35, (183, 228), 14.6km/h, 383/day

now

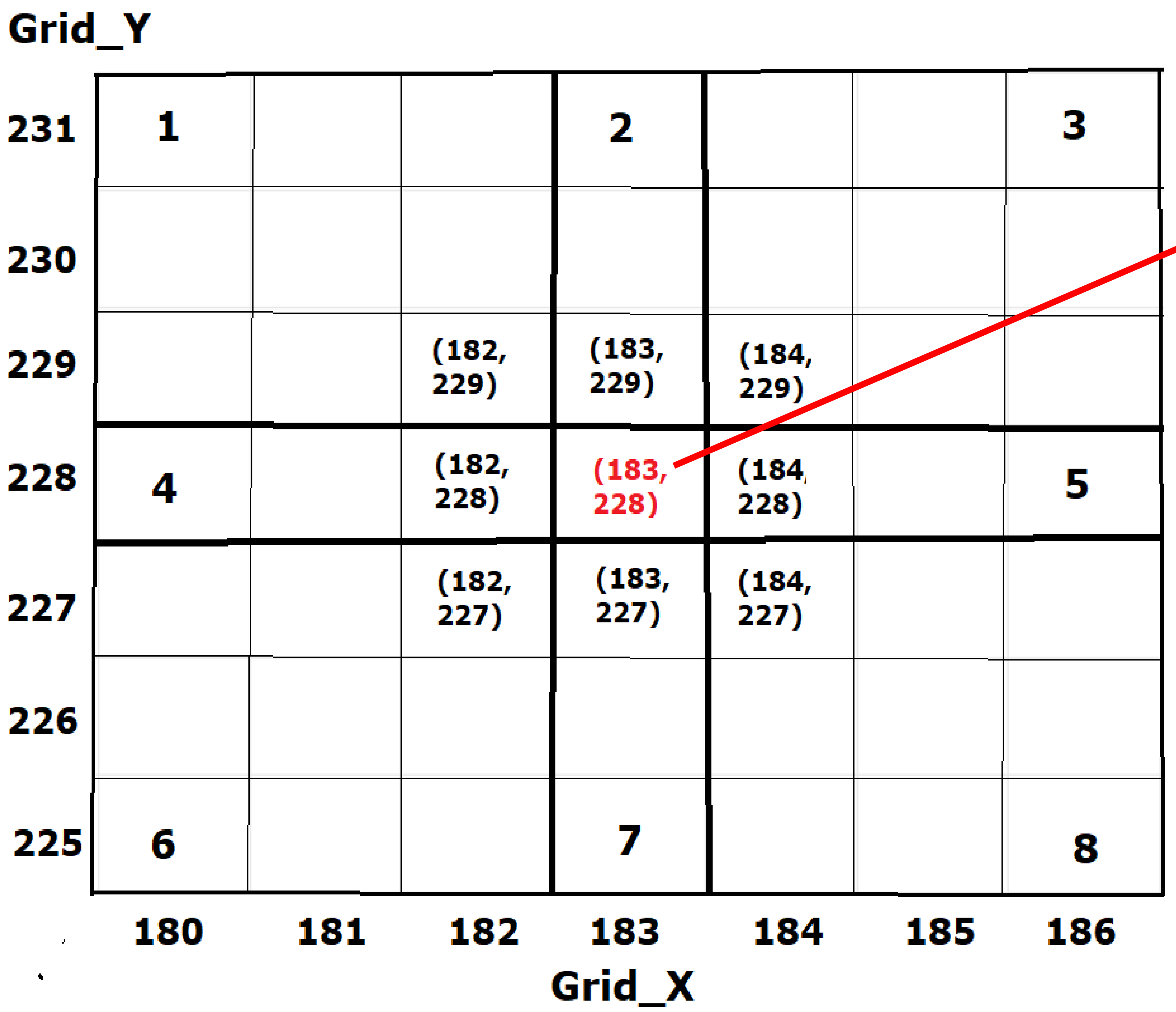


Track & Know



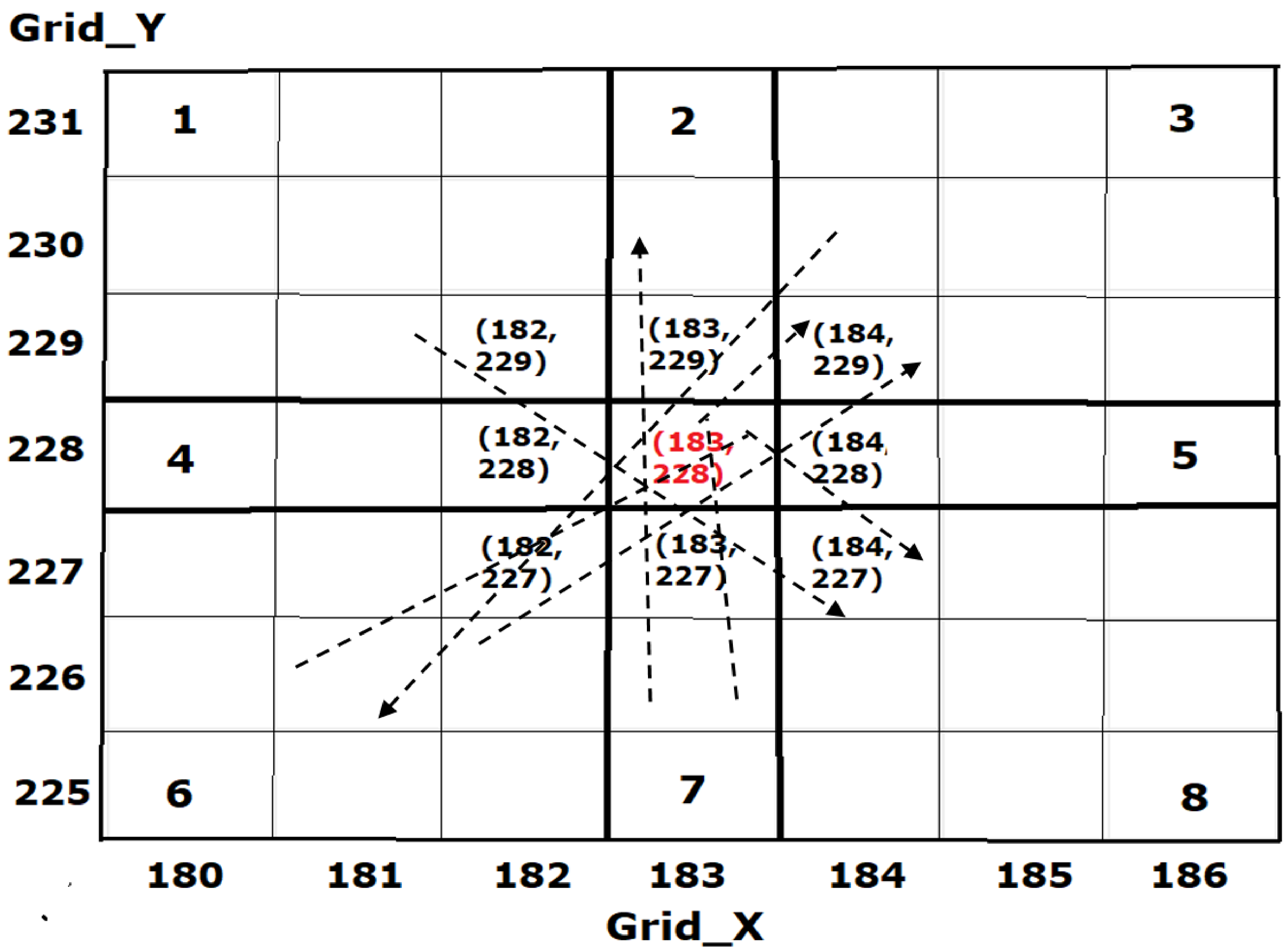
Grid_Y
242
238
234
230
226
222
218
170
174
178
182
186
190
194
Grid_X
Zone (183, 228)

Track & Know



(183,228): 217 trips

- **Start and/or end: 41.9%**
  - ❑ Start: 16.1%
  - ❑ End: 4.1%
  - ❑ Inside: 21.7%
- **Passing: 58.1%**
  - ❑ 6->3: 17.1%
  - ❑ 1->8: 6.9%
  - ❑ 6->8: 4.1%
  - ❑ 7->3: 4.1%
  - ❑ 3->6: 3.7%
  - ❑ 7->2: 2.7%



- **Passing: 58.1%**
  - 6->3: 17.1%
  - 1->8: 6.9%
  - 6->8: 4.1%
  - 7->3: 4.1%
  - 3->6: 3.7%
  - 7->2: 2.7%

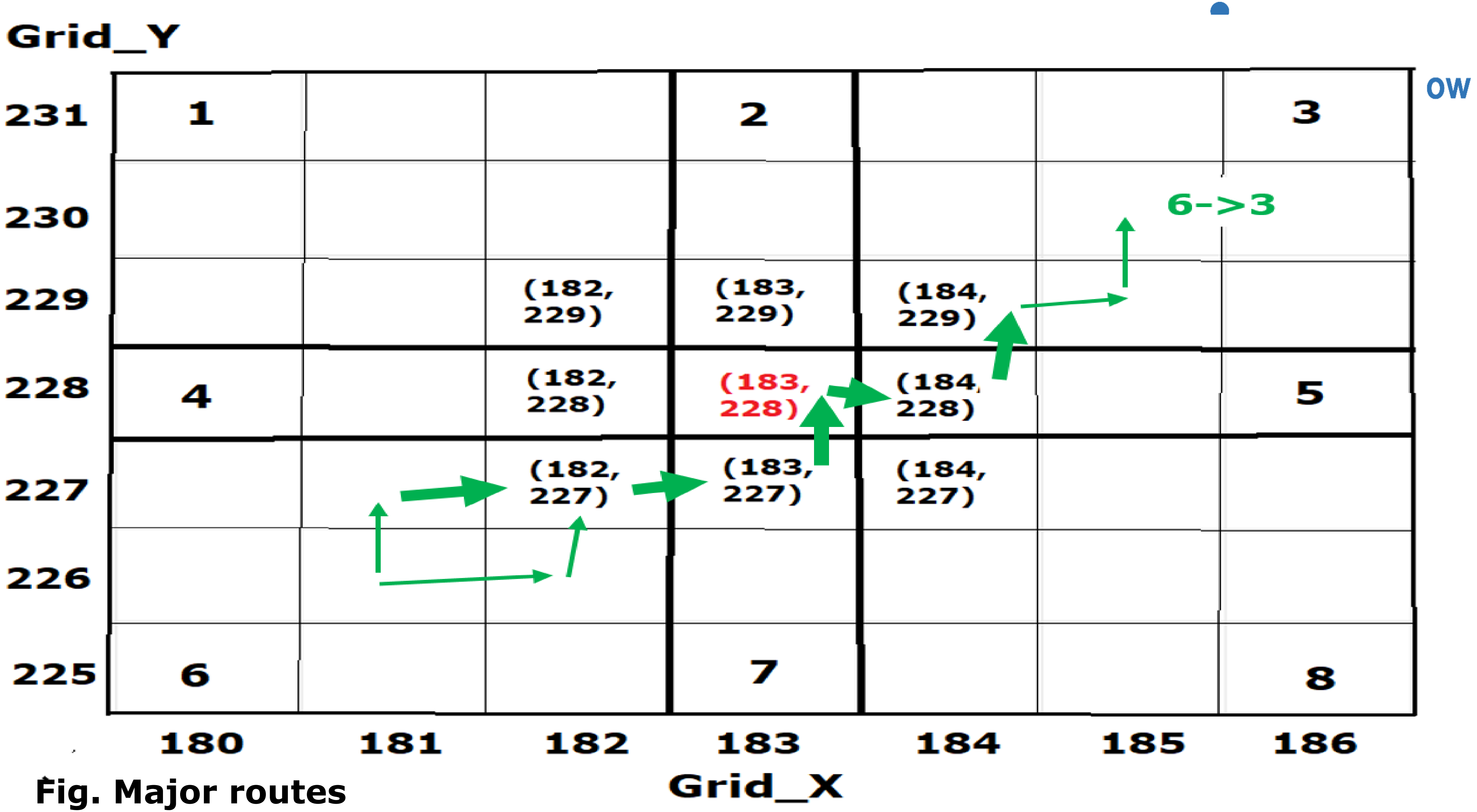
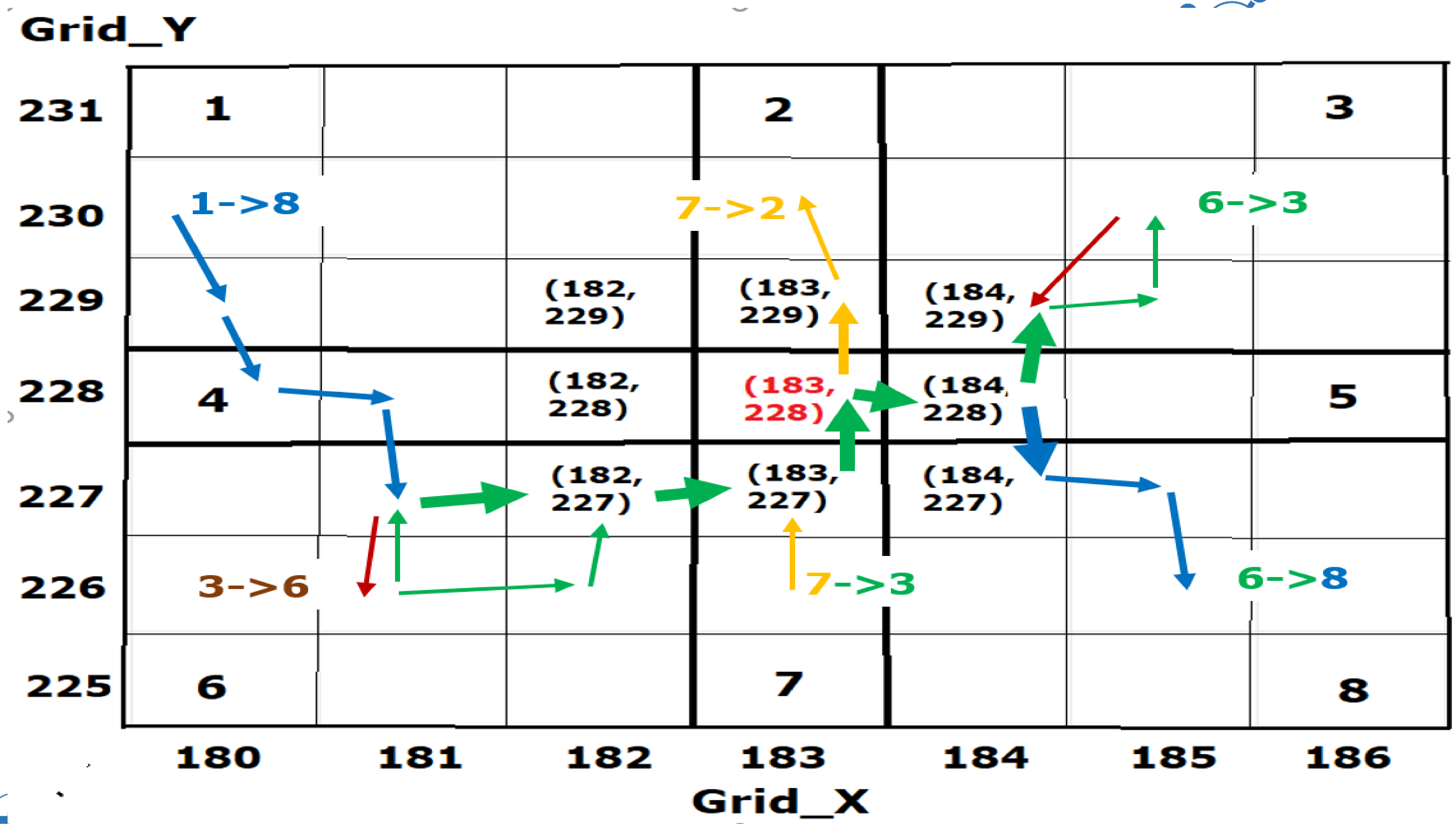
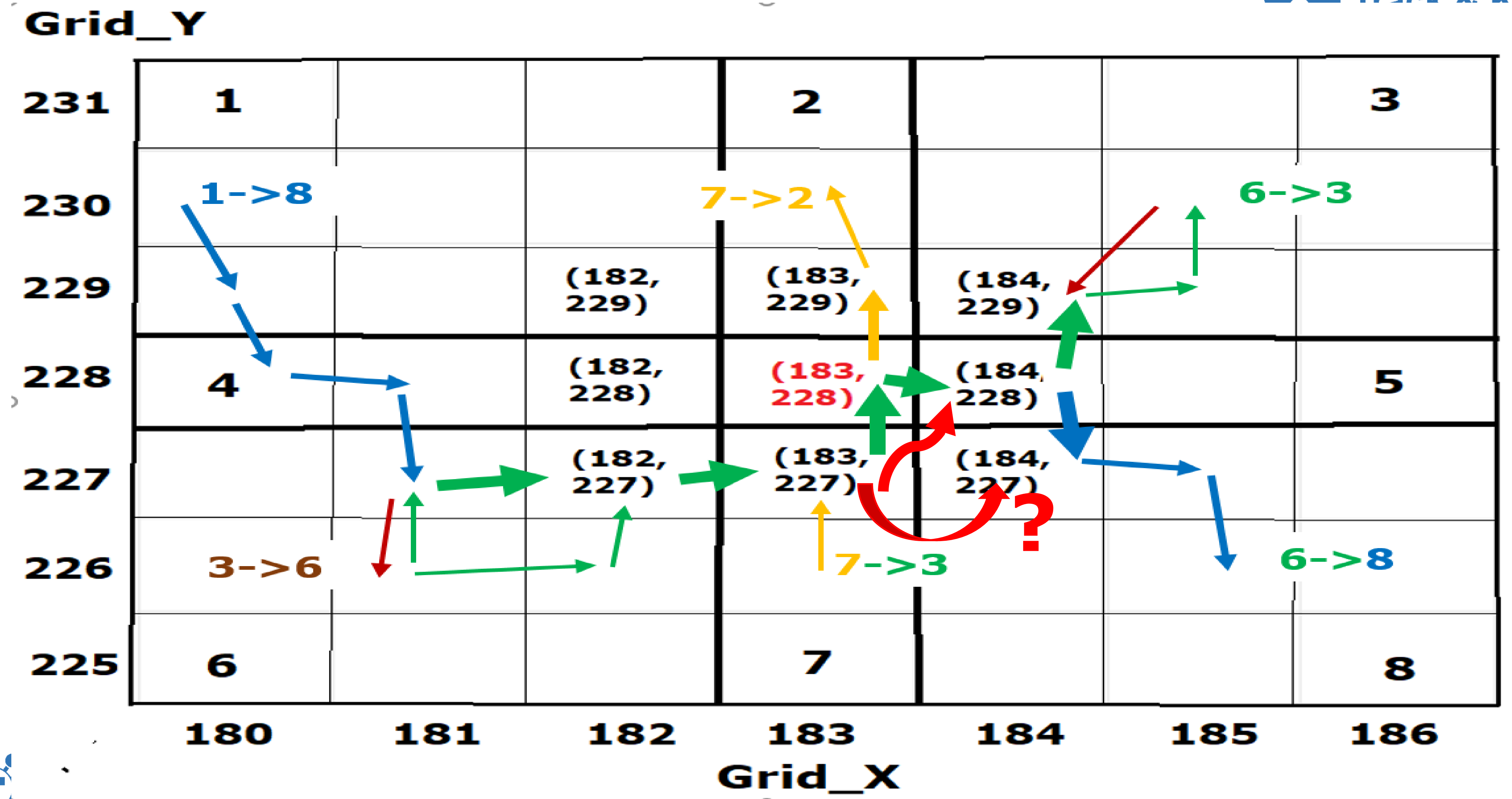


Fig. Major routes


Grid_Y
231
230
229
228
227
226
225
1
2
3
4
5
6
7
8
1->8
7->2
6->3
3->6
7->3
6->8
(182, 229)
(183, 229)
(184, 229)
(182, 228)
(183, 228)
(184, 228)
(182, 227)
(183, 227)
(184, 227)
180
181
182
183
184
185
186
Grid_X

Result 5: Alternative route searching



Future work

- ***Extend*** **the size of the data to be used**
- ***Find*** **alternative routes and** ***compare*** **these routes with the currently observed ones**
- ***Visualize*** **the results using visualization tools**
- ***Suggest*** **policy measures**
- ***Change*** **regular grids into spatial compartment?**

Thank you!

(feng.liu@uhasselt.be)