URBAN-GEO BIG DATA

Analysis of the Floating Car Data of Turin Public Transportation system: first results

Roberta Ravanelli, Mattia Crespi

University of Rome "La Sapienza"
Department of Civil, Constructional and Environmental Engineering
Geodesy and Geomatics Division

Turin, 18th July 2018

PRIN PROJECT: *URBAN GEOmatics for Bulk Information Generation, Data Assessment and Technology Awareness*

MINISTERO DELL'ISTRUZIONE DELL'UNIVERSITA' E DELLA RICERCA

Floating Car Data (FCD)

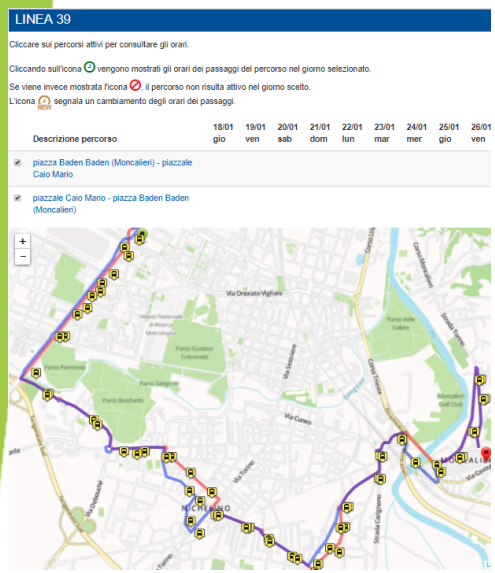
- The **Floating Car Data** (**FCD**) of **Turin Public Transportation** system were acquired by every vehicle of the fleet through its On Board Unit (OBU) in the **month** of **April 2017**, with a **variable time interval** (difference of several seconds)
- The **FCD** are provided in the **CSV** format and include the **geographical coordinates** along with a **set of attributes** (vehicle code, line code, turn, timestamp, ecc.)
- The original file is very **heavy** (2.19 GB, **30·000·000**) and it was converted in a **database** through a Python script based on the `sqlite3` and `pandas` libraries

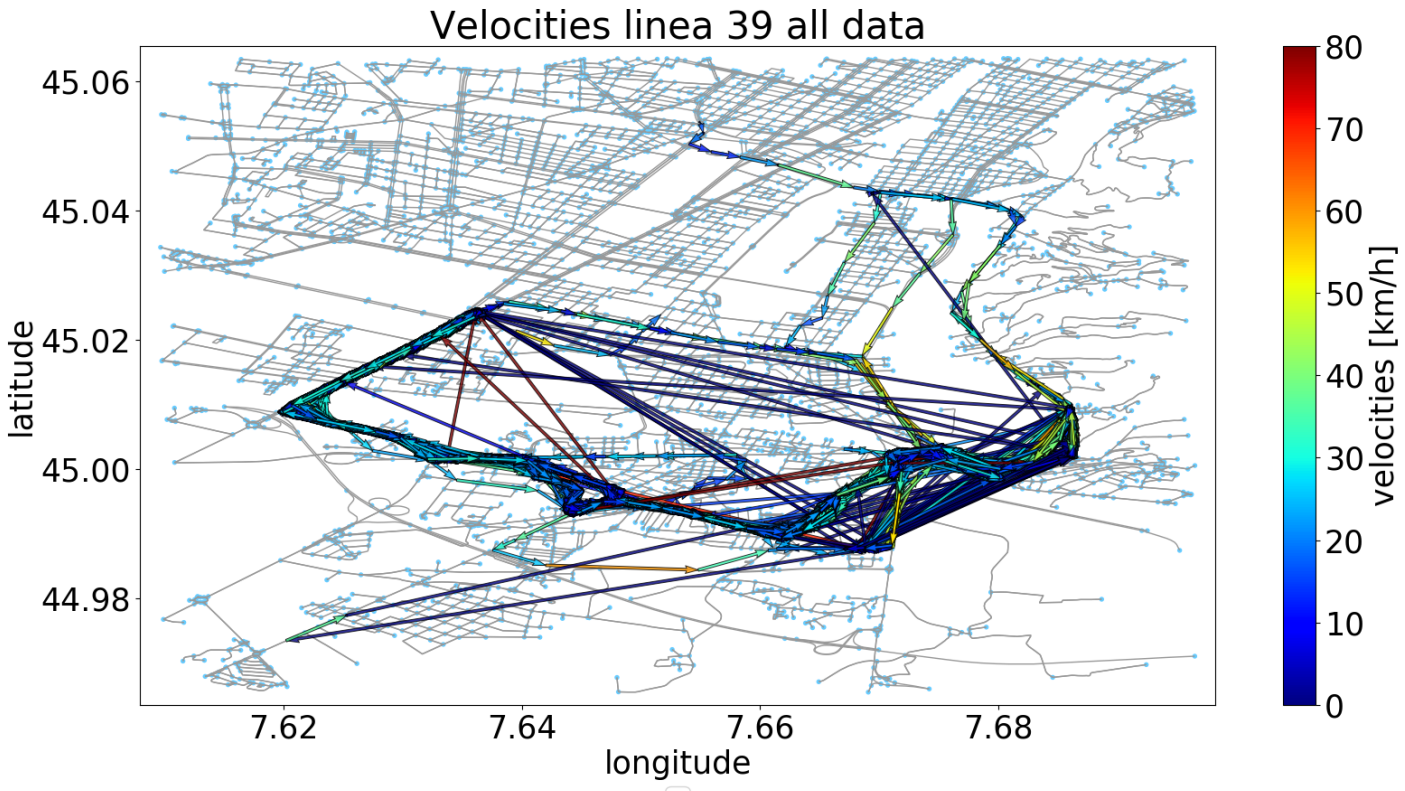
Velocity analysis

- The **FCD** were **analysed** for **lines**, then for **vehicles** and finally they were **chronologically ordered**
- For every line of the transportation network:
  - the *Vincenty* formula was used to compute the **planimetric displacement** Δs between **two positions** of the specific vehicle in two **consecutive time moments**
  - the **velocities** were computed as $v = \frac{\Delta s}{\Delta t}$
- The computed **velocities** were represented as **arrows** and plotted on top of the **Turin drive network graph**, automatically downloaded from Open Street Map
- Before proceeding with the time analysis, the **outliers** were **removed** by eliminating all the records not statistically significant ($\Delta t > 99.5^{th}$ percentile & $\Delta t < 0.5^{th}$ percentile & $v > v_{mean}$)

Line 39: velocities

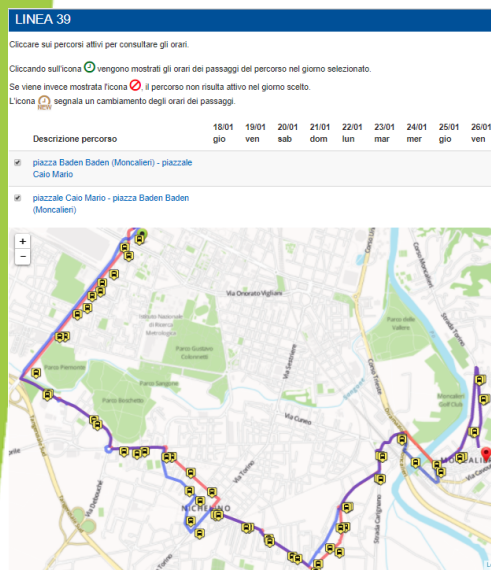
After the **outlier removal**, the reconstructed path follows more closely the actual line route: the **longest arrows**, probably due to the bus routes from and to the depot, are **eliminated**

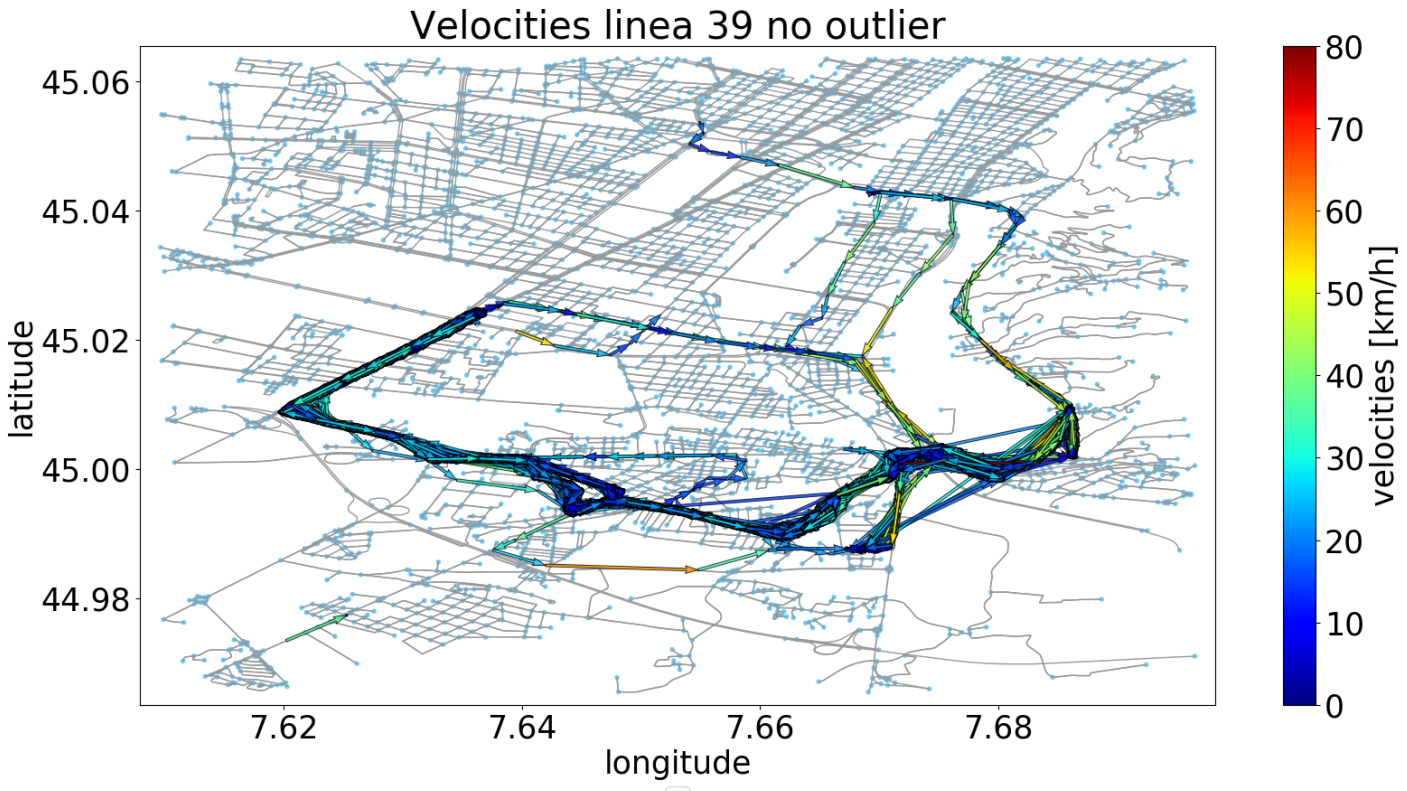




Line 39: velocities

After the **outlier removal**, the reconstructed path follows more closely the actual line route: the **longest arrows**, probably due to the bus routes from and to the depot, are **eliminated**

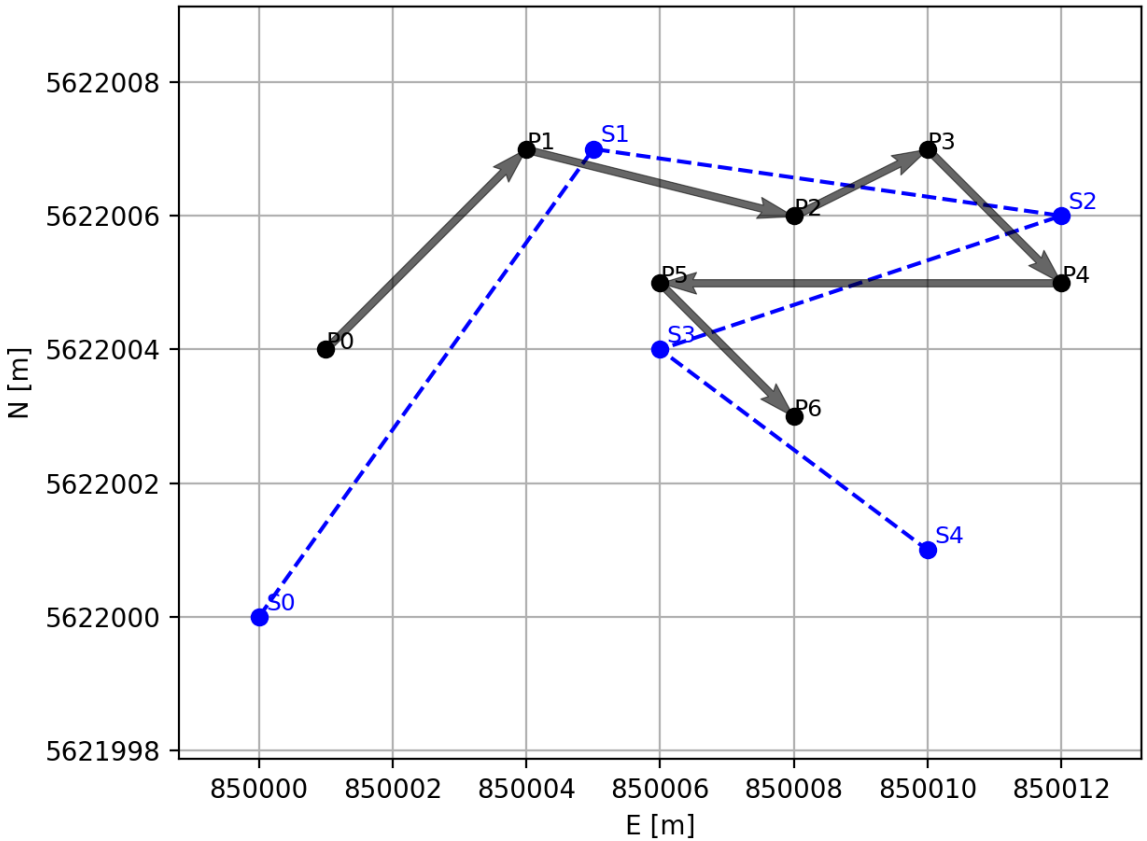




FCD projection to line networks

A preliminary strategy was implemented to **assign** the **velocities to** the **line network topology**:

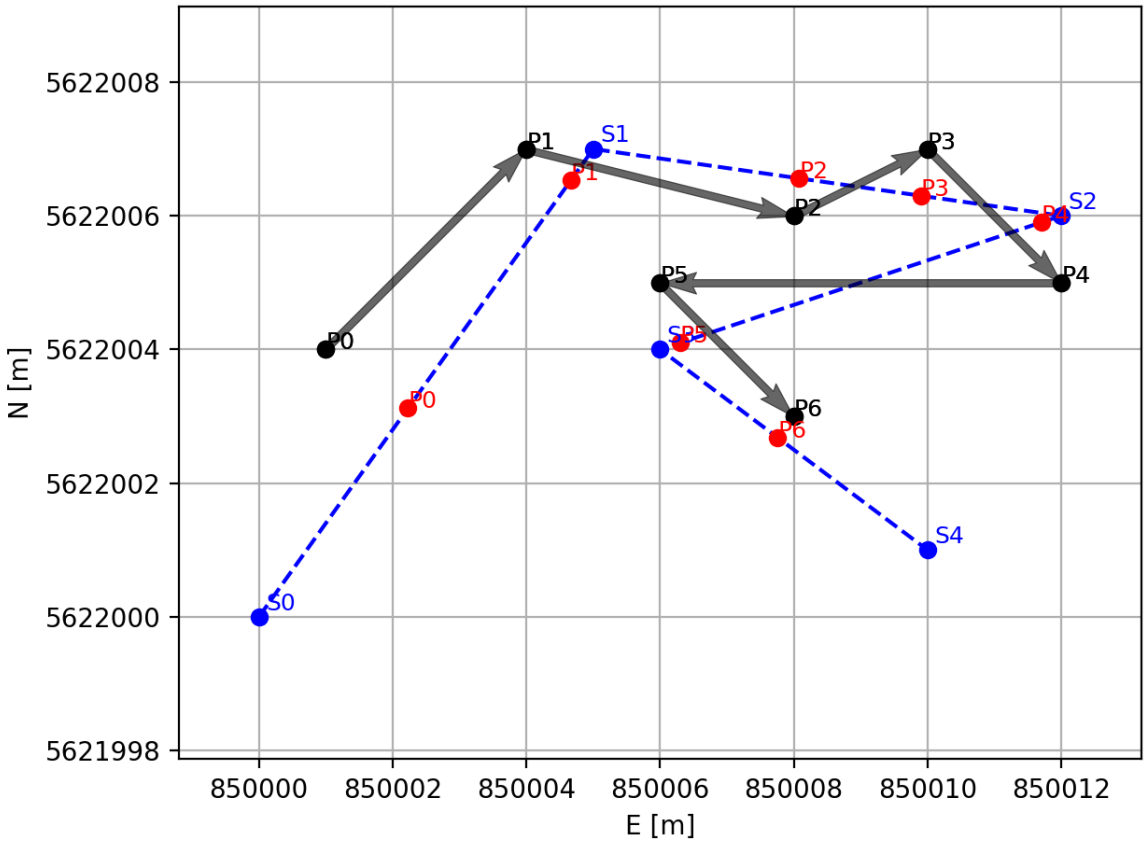
- for **every FCD point**, the **closest tree** of the specific line network is selected



FCD projection to line networks

A preliminary strategy was implemented to **assign** the **velocities to** the **line network topology**:

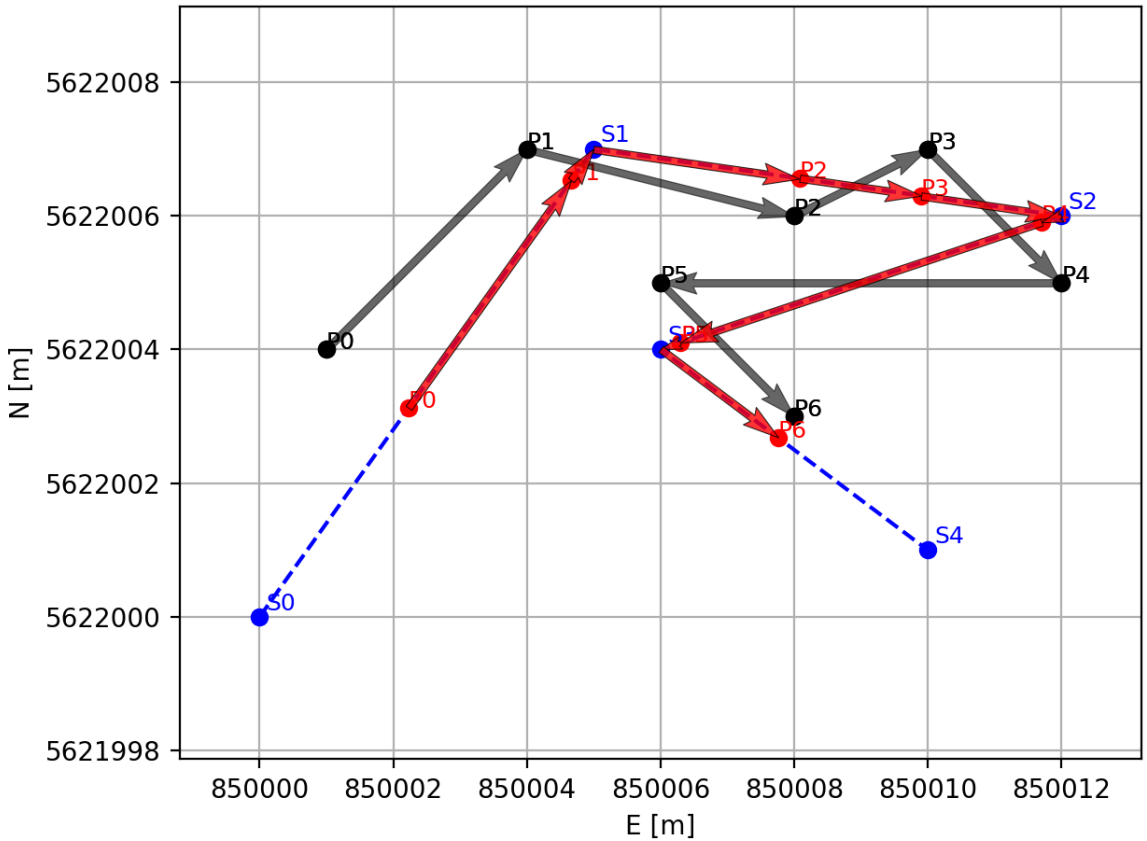
- for **every FCD point**, the **closest tree** of the specific line network is selected



FCD projection to line networks

A preliminary strategy was implemented to **assign** the **velocities to** the **line network topology**:

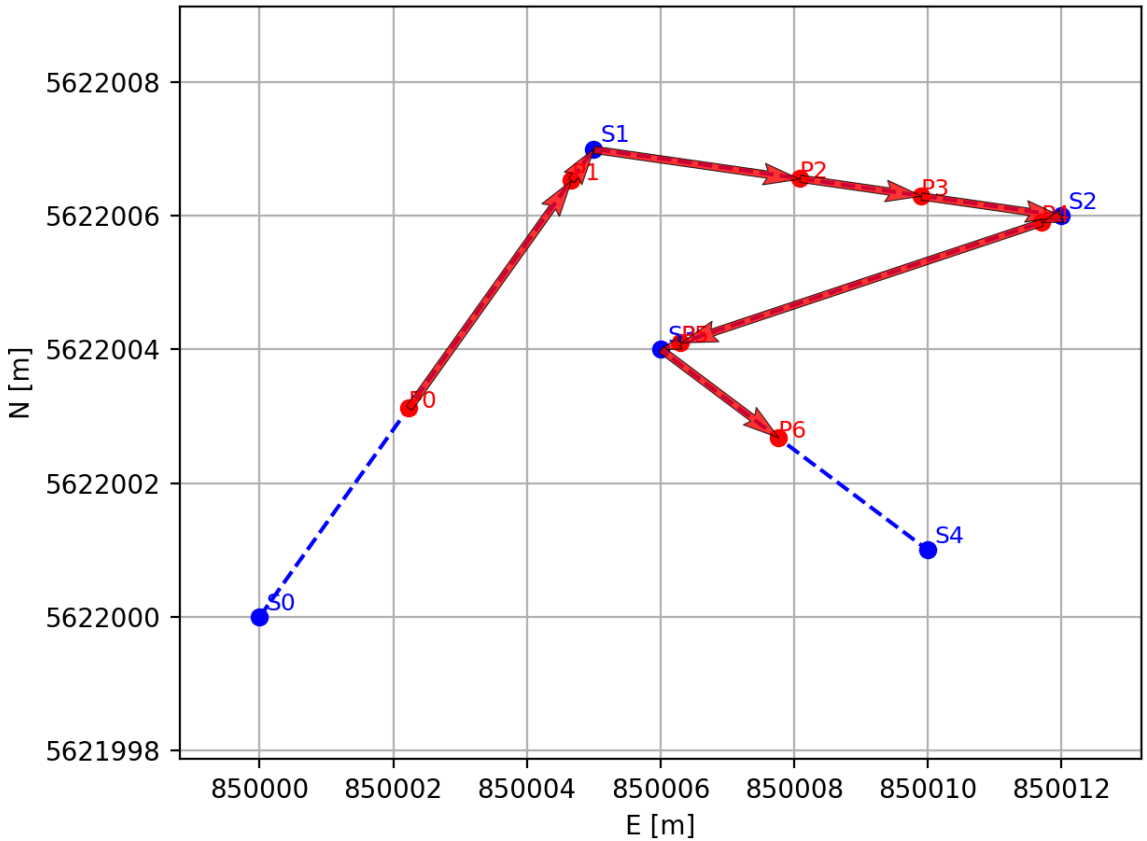
- for **every FCD point**, the **closest tree** of the specific line network is selected



FCD projection to line networks

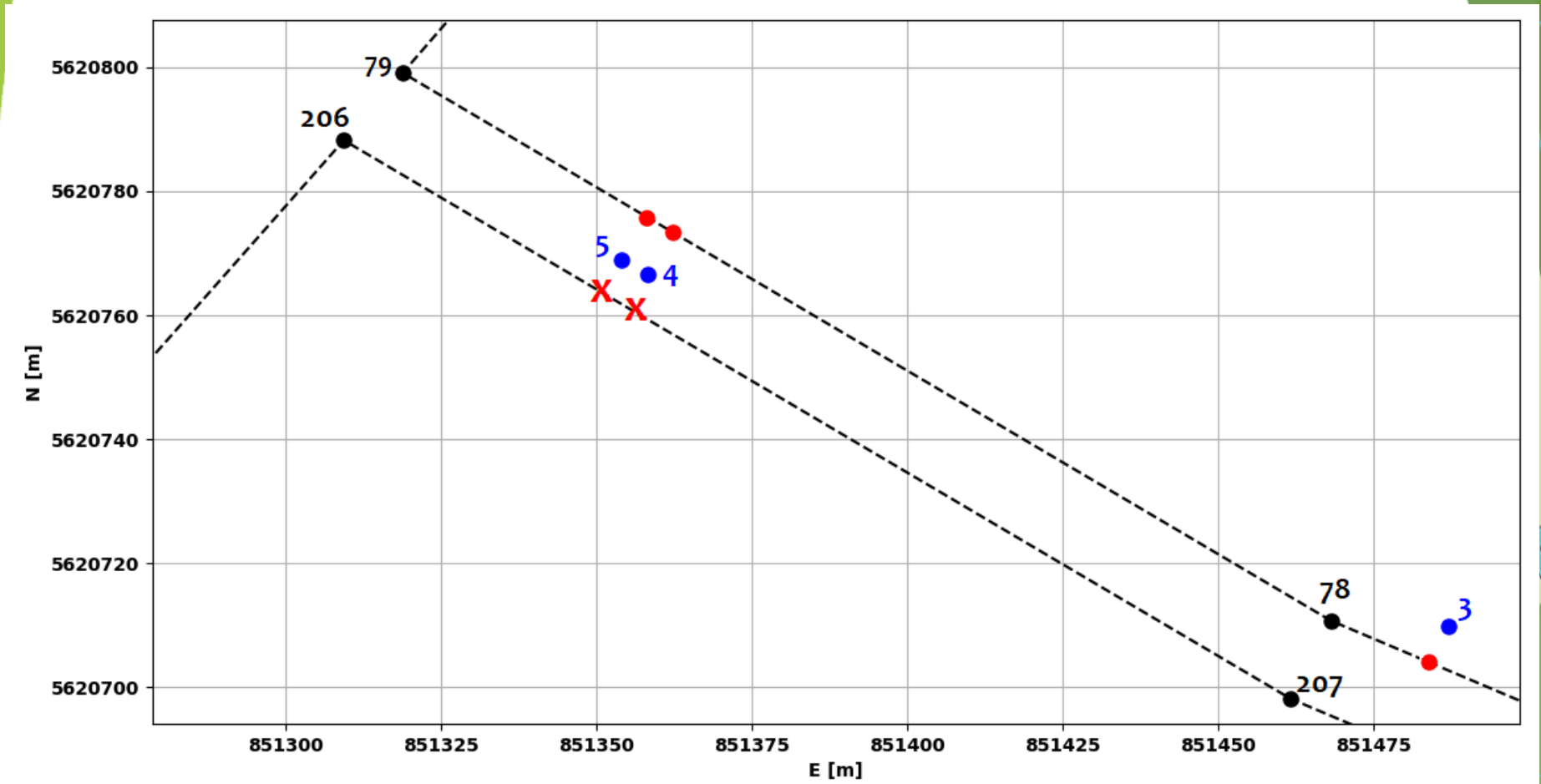
A preliminary strategy was implemented to **assign** the **velocities to** the **line network topology**:

- for **every FCD point**, the **closest tree** of the specific line network is selected



Topological issues

Topological issues occur when the FCD point is located in a segment in which the **distance between two** (or more) **trees** is **comparable** to the **GNSS measurement errors**



Topological issues

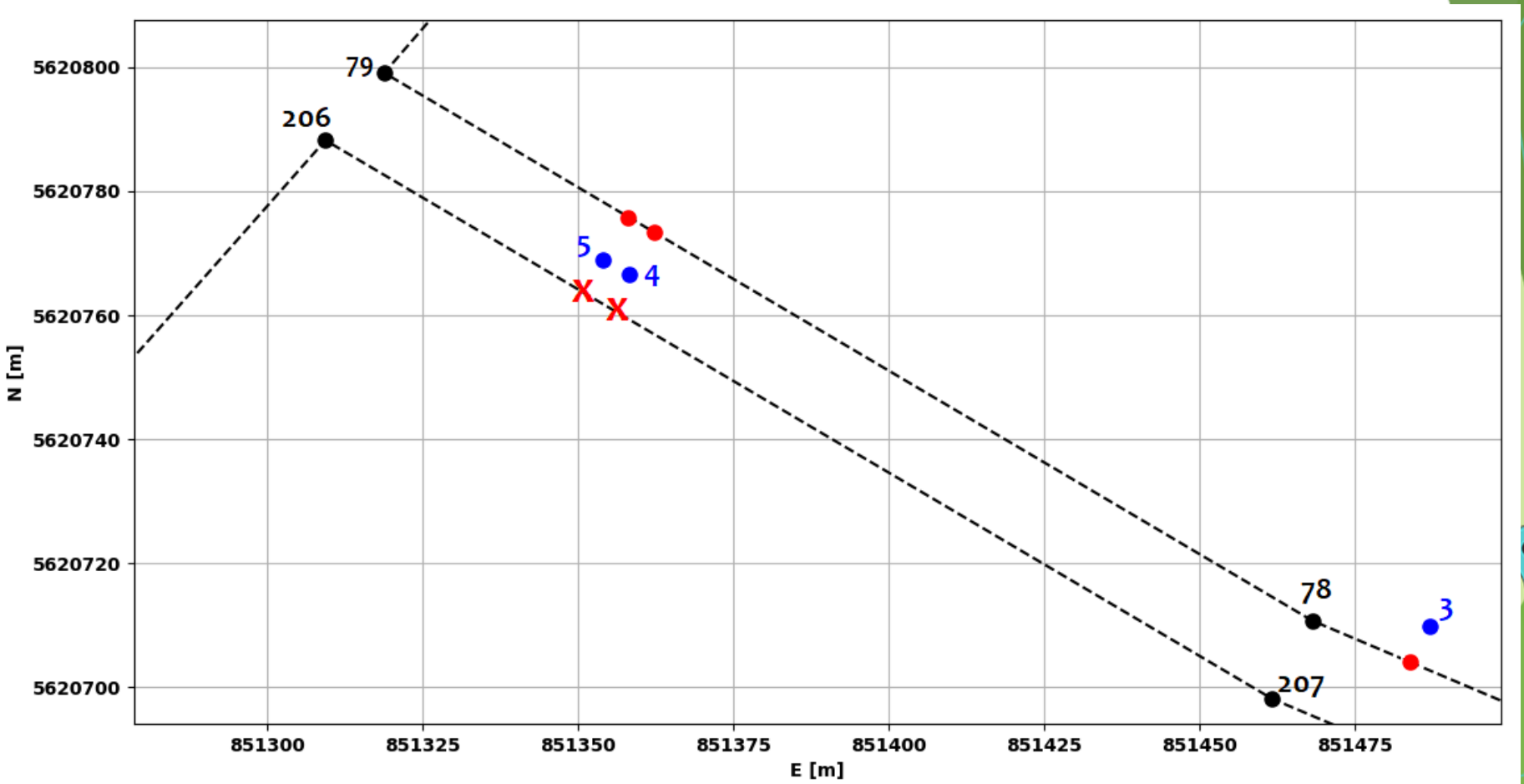
A possible solution is to consider:

- the **cardinality information** contained in the **line network**
- the **temporal information** contained in the **FCD**

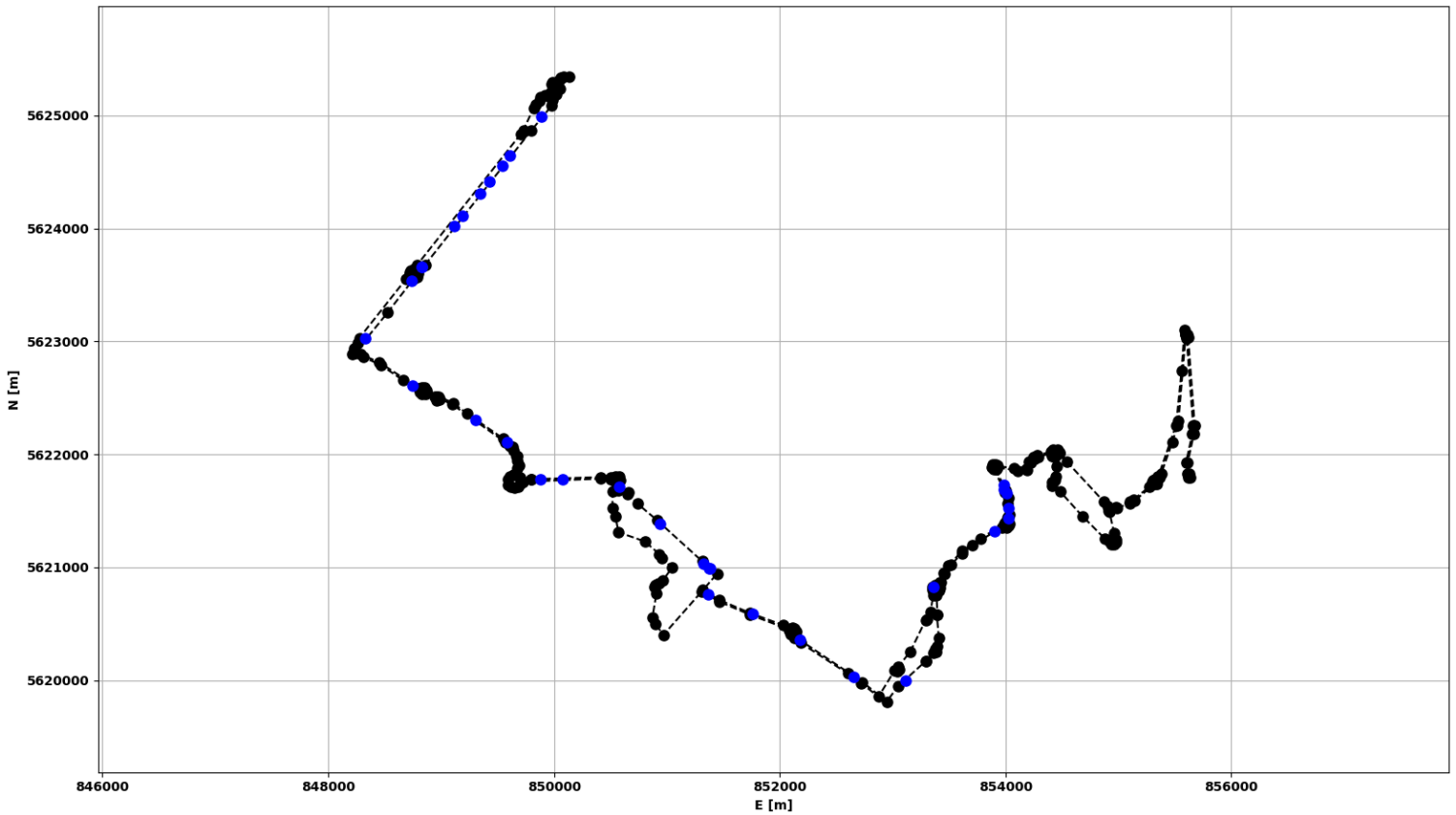
Select the segment closest to the previous selected tree

Topological issues

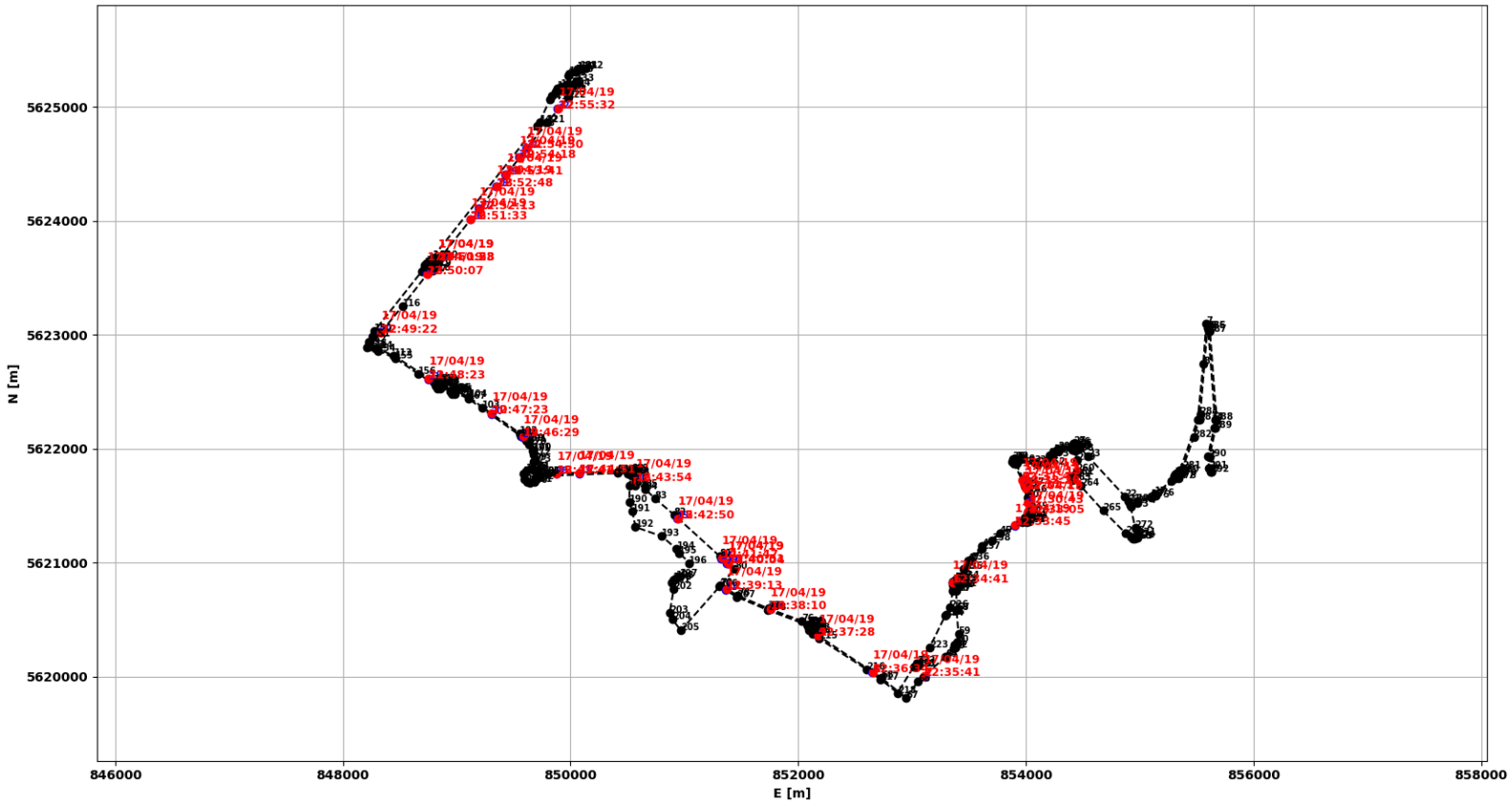
It is rather improbable that the FCD **point 4** and **point 5** may be assigned to the **tree 206-207** of the network, since the vehicle was located in the **tree 77-78** few moments before



FCD projection to line networks



FCD projection to line networks



Conclusions and further developments

A **general methodology** able to **analyse** the huge amount of information contained in **Transport Big Data** has been developed

Necessary further developments:

- refine the outlier removal process in order to all the **velocities not referable** to the **actual path** of the **lines**
- to **test** the developed **topological procedure** on **all** the **velocity data**, by checking the **effective reliability** and **real-time feasibility** of the designed methodology
- to **compute** the **impedence maps** and deliver the corresponding **metadata**
- extend the developed methodology to the other cities of the PRIN

Thank you for your kind attention!