



**SAPIENZA**  
UNIVERSITÀ DI ROMA

# Analisi comportamentale degli utenti stradali sulla base di Big Data in mobilità

**Facoltà di Ingegneria**  
**Dipartimento di Ingegneria Civile Edile e Ambientale**  
**Dottorato in Infrastrutture e Trasporti**

**Ing. Agnese Bracci**  
**Matricola 1401712**

Relatore  
Prof. Ing. Gaetano Fusco

A.A. 2019-2020

# Indice

<b>1 Introduzione.....</b>	<b>2</b>
1.1 Obiettivo della tesi .....	3
<b>2 Stato dell'arte .....</b>	<b>5</b>
2.1 Possibili applicazioni di smart data in mobilità .....	5
2.2 Analisi comportamentale degli utenti stradali e analisi della dinamica giornaliera.....	9
2.3 Teoria della razionalità limitata e modelli comportamentali.....	15
<b>3 Approccio metodologico .....</b>	<b>33</b>
3.1 Ricostruzione del percorso .....	37
3.2 Definizione delle alternative di scelta del percorso .....	42
<b>4 Applicazione al caso di studio .....</b>	<b>49</b>
4.1 Analisi preliminare dei dati .....	50
4.1 Evidenze dall'analisi preliminare dei dati .....	51
4.2 Analisi relativa al grafo stradale .....	57
4.2.1 Analisi della variazione dei tempi di viaggio .....	65
4.2.2 Analisi della dinamica giornaliera .....	87
4.2.3 Calibrazione di un modello di scelta del percorso .....	100
<b>5 Conclusioni.....</b>	<b>112</b>
<b>Bibliografia.....</b>	<b>117</b>

# 1 Introduzione

Gli ambiti di applicazione dei c.d. “big data” nel campo dei sistemi di trasporto sono numerosi e molto diversi tra loro.

Possono essere utilizzati per conoscere lo stato del veicolo, le prestazioni del sistema di trasporto, le posizioni degli utenti e le preferenze o le abitudini di questi ultimi, fino a fornire un supporto ai decisori in tema di regolamentazione e definizione di policies.

Con riferimento all’analisi del comportamento degli utenti stradali, i Floating Car Data (di seguito FCD), in particolare, rappresentano una preziosa fonte di informazioni tra cui, ad esempio: origine e destinazione degli spostamenti e localizzazione delle fermate intermedie nel caso di viaggi “a catena”.

Offrono, inoltre, una rappresentazione accurata degli effettivi meccanismi di scelta del percorso da parte degli utenti che può essere utilizzata per calibrare e validare modelli comportamentali, finora spesso basati su informazioni raccolte da piccole flotte di veicoli (ad esempio i taxi) o da campioni di guidatori coinvolti in esperimenti di laboratorio o mirate campagne di indagine e, più raramente, su larga scala.

I FCD possono, infatti, essere sfruttati nelle diverse fasi di calibrazione di modelli di utilità casuale multilivello basati su percorsi:

dall’identificazione del set di alternative di percorso, specificando una struttura matematica che catturi la correlazione tra percorsi diversi, alla stima dei valori più probabili dei coefficienti del modello di scelta.

## 1.1 Obiettivo della tesi

L'elaborazione di FCD rispetto alle informazioni raccolte da piccole flotte di veicoli o attraverso esperimenti di laboratorio pone alcune criticità dovute, da un lato alla dimensione dei database oggetto di analisi e, dall'altra, alla natura stessa dei dati, caratterizzati da una sequenza di posizioni sparse e svincolate dal grafo stradale sottostante.

Per tale ragione, al fine di condurre una ampia analisi del sistema degli spostamenti in area urbana - attraverso l'elaborazione di un set esteso di FCD - è stata definita una metodologia generale che prevede:

- la ricostruzione sul grafo stradale del percorso seguito dagli utenti in ciascun viaggio effettuato, a partire dalle posizioni sparse registrate;
- l'analisi dell'intero insieme di percorsi ricostruiti relativi ad una stessa coppia Origine-Destinazione (di seguito O-D) zona-zona. Due percorsi appartenenti alla stessa coppia O-D, infatti, possono differire di quantità trascurabili o per brevi deviazioni. In questo modo l'insieme delle alternative diventa difficilmente gestibile. Per tale ragione è stato sviluppato un processo per la suddivisione del set di percorsi in più cluster e la selezione del percorso più rappresentativo per ogni cluster. I percorsi rappresentativi costituiscono il set finale di alternative a disposizione dell'utente stradale di un modello comportamentale (percorsi espliciti).

La suddivisione dei percorsi in cluster e quindi l'identificazione di un set ristretto di alternative ha, inoltre, come finalità l'"analisi della dinamica giornaliera" delle scelte degli utenti stradali, che viene condotta verificando la tendenza degli individui a:

- cambiare il percorso abituale in favore di uno appartenente ad un altro cluster nel caso in cui durante l'ultimo viaggio effettuato sia

stato sperimentato un ritardo significativo rispetto al tempo di percorrenza atteso.

- oppure anticipare l'orario di partenza nel caso in cui durante l'ultimo viaggio effettuato sia stato sperimentato un ritardo significativo rispetto al tempo di percorrenza atteso.

La metodologia definita nell'ambito della ricerca è stata applicata a:

- un primo set di FCD relativi al mese di maggio 2010 registrati nell'area metropolitana di Roma;
- un set più esteso e recente di dati (gennaio-giugno 2015) registrati nella stessa area.

Sulla base dei risultati ottenuti, è stata approfondita la possibilità di modellare i comportamenti di scelta del percorso mediante modelli di scelta discreta.

Sono state, a tal fine, calibrate diverse forme funzionali di utilità sistematica di un modello logit di scelta del percorso.

Nel capitolo 2 della presente tesi di dottorato è riportata una overview delle possibili applicazioni di smart data in ambito mobilità presenti in letteratura: a partire da analisi di tipo descrittivo (deviazioni degli utenti dallo "shortest path", tendenza a selezionare un percorso abituale su una specifica coppia O-D, cambio di orario di partenza o di percorso, etc.) fino alla definizione di modelli comportamentali di scelta del percorso.

Nel capitolo 3 si riporta la descrizione dettagliata della metodologia utilizzata per l'analisi del set di FCD a disposizione e, nel capitolo 4, i risultati ottenuti dall'applicazione di tale approccio e dalla calibrazione del modello.

## 2 Stato dell'arte

### 2.1 Possibili applicazioni di smart data in mobilità

Negli ultimi anni, grazie ai progressi in ambito ICT, al sempre più diffuso utilizzo delle tecnologie e ai processi di digitalizzazione in atto, grandi moli di dati sono stati collezionati, elaborati, analizzati.

I c.d. 'big data' rappresentano una importante fonte di informazioni in una ampia gamma di settori.

In ambito trasportistico i FCD, in particolare - oggetto delle analisi condotte nell'ambito della presente ricerca - stanno acquisendo una rilevanza sempre maggiore, da un lato per la molteplicità di applicazioni per cui possono essere utilizzati e dall'altro perché attraverso tali dati è possibile superare alcuni problemi derivanti dall'uso dei sensori fisici (ad esempio, limitata distribuzione geografica, disomogeneità nelle misurazioni, copertura imperfetta delle strade secondarie, etc.).

Come già menzionato, sulla base dei FCD, possono essere eseguite diverse analisi del traffico a supporto dei servizi di mobilità (Ajmar, A., Arco, E., Boccardo, P., Perez, F., 2019).

La conoscenza del traffico stradale sulla rete e la sua evoluzione nel breve termine, infatti, rappresentano componenti fondamentali di applicazioni ATIS (Advanced Traveler Information Systems) e ATMS (Advanced Traffic Management System).

A tal fine, i FCD si caratterizzano come uno strumento efficace per acquisire importanti informazioni rispetto ai tempi di viaggio e alle velocità sulla rete stradale e rendere sempre più accurate le previsioni a breve termine delle condizioni di viaggio (De Fabritiis, C., Ragona, R., Valenti, G., 2008).

Riguardo a quest'ultimo aspetto, per citare alcuni esempi, C. De Fabritiis et al. (De Fabritiis, C., Ragona, R., Valenti, G., 2008) definiscono una metodologia utile alla determinazione delle velocità su ciascun arco stradale dell'area analizzata. L'algoritmo è suddiviso in tre step: in primo luogo si effettua una operazione di map-matching, di ricostruzione del percorso tra due posizioni registrate e, infine, si stima la velocità di viaggio sull'arco.

G. Fusco et al. (Fusco, G., Colombaroni, C., Comelli, L., Isaenko, N., 2015) utilizzano due dataset per testare le performance di previsione delle velocità sia di modelli espliciti (i due modelli di assegnazione dinamica del traffico Dynasmart e QDTA), che impliciti (Reti Neurali Artificiali e Reti Bayesiane).

Approfondiscono (Fusco, G., Colombaroni, C., Isaenko, N., 2016) inoltre l'approccio implicito, validando modelli di reti neurali artificiali e reti bayesiane per previsioni a breve termine, sfruttando l'ampio database di velocità medie registrate nell'area metropolitana di Roma durante sette mesi, attraverso FCD.

A differenza dell'analisi sopra riportata, in cui le velocità medie storiche sono utilizzate come stima a priori, lo stesso set di dati grezzi è poi utilizzato per ottenere la migliore stima a priori possibile, sulla base di osservazioni pregresse (Fusco, G., Colombaroni, C., Isaenko, N., 2016).

L. Mannini et al. (Mannini, L., Carrese, S., Cipriani, E., Crisalli, U., 2015) definiscono una metodologia per la previsione dei tempi di viaggio, stimati sulla base di dati di traffico sia storici (FCD) che in real-time (forniti da rilevatori a spira) e la applicano a un tratto di autostrada urbana di Roma.

Anche M. Tulic et al. (Tulic, M., Bauer, D., Scherrer, W., 2014) utilizzano FCD raccolti su larga scala per la previsione a breve termine dei tempi di viaggio su archi e percorsi.

Per quanto riguarda, invece, le analisi di performance della rete stradale, numerosi approcci sono proposti per il monitoraggio delle velocità sulla rete (Diependaele, K., Riguelle, F., Temmerman, P., 2016), (Rahmani, 2010), (Li, Q., Zhang, T., Yu, Y., 2011), (Rambaldi, S., Marchioni, M., Bazzani, A., Giorgini, B., 2012) e i tempi di viaggio.

Rispetto a quest'ultimo tema, M. Rhamani et al. (Rahmani, M., Jenelius, E., Koutsopoulos, H. N., 2015), (Rahmani, M., Jenelius, E., Koutsopoulos, H. N., 2013) sviluppano un metodo non parametrico per la stima della distribuzione dei tempi di viaggio su specifici percorsi, utilizzando FCD a basso campionamento. Il modello corregge possibili errori associati all'utilizzo di tali dati di input (copertura parziale del percorso sulla rete con i dati disponibili, sovrapposizione con altri archi della rete non inclusi nel percorso, rappresentatività del campione, etc.) ed è applicato al caso studio di una flotta di taxi nella città di Stoccolma. I risultati ottenuti sono confrontati con i dati registrati sui percorsi dotati di telecamere ANPR (Automatic Number Plate Recognition).

Z. Yong-Chuan et al. (Yong-chuan, Z., Xiao-qing, Z., Zhen-ting, C., 2011) e L. Xu et al. (Xu, L., Yue, Y., Li, Q., 2013) utilizzano i FCD per l'individuazione di situazioni di congestione stradale.

Nel primo caso, circa 6 milioni di dati generati da 500 taxi in un giorno, vengono processati attraverso una routine di map-matching e di stima della velocità di viaggio per ottenere una rappresentazione della distribuzione della congestione stradale.

Anche nel secondo caso, i FCD sono oggetto di una fase di pre-processing e di map-matching al fine di determinare le velocità e lo stato di traffico su ciascuno degli archi stradali. A tale risultato, viene poi applicato un modulo "data-cubing" che permette di identificare le situazioni di



congestione e aggregarle, in modo da avere una sintesi schematica del modello di congestione ricorrente.

M. A. Waller et al. (Waller, 2013) riportano una serie di esempi che mostrano come il processo di gestione della supply chain si colleghi con le analisi predittive e i Big Data. Tra questi la previsione delle situazioni di congestione a beneficio del “transport management” da parte dell’operatore del trasporto e del commerciante.

Sempre attraverso i dati FCD è possibile, inoltre:

- identificare e localizzare incidenti sulla rete stradale (ad esempio (Kerner, B.S., Demir, C., Herrtwich, R.G., Klenov, S.L., Rehborn, H., Aleksic, M., Haug, A., 2005));
- effettuare una analisi dinamica dell’accessibilità a diversi Points Of Interest (di seguito POIs) vincolati alla rete, a partire dalle velocità registrate in tempo reale (Li, Q., Zhang, T., Wang, H., Zeng, Z., 2011);
- valutare i consumi di carburante sulla base delle caratteristiche del traffico stradale. G. Song et al. (Song, G., Yu, L., 2009), ad esempio, propongono un metodo per il calcolo dell’efficientamento dei consumi sulla base di indicatori che rappresentano da un lato il consumo di carburante e, dall’altro, il consumo di carburante alla velocità di maggiore efficienza. L’approccio è stato applicato per derivare la distribuzione temporale dell’efficienza dei consumi nella città di Pechino.
- Supportare il processo di regolazione semaforica (ad esempio (Astarita, V., Giofrè, V. P., Guido, G., Vitale, A., 2017) e (Protschky, V., Ruhhammer, C., Feit, S., 2015)).
- Effettuare la stima dinamica di matrici O-D (ad esempio (Nigro, M., Cipriani, E., Del Giudice, A., 2018)).

- Approfondire aspetti specifici dei Transport System Models (TSMs), quali ad esempio la definizione della zonizzazione e la costruzione del grafo (Croce, A.I., Musolino, G., Rindone, C., Vitetta, A., 2009).

Come si può notare, gli ambiti di applicazione dei FCD sono numerosi e molto diversi tra loro.

Nei paragrafi seguenti sono approfonditi alcuni ulteriori possibili utilizzi dei dati, che ricadono nell'ambito della presente ricerca.

## 2.2 Analisi comportamentale degli utenti stradali e *analisi della dinamica giornaliera*

In aggiunta alle possibili applicazioni di FCD in mobilità menzionate nel paragrafo precedente, tali dati permettono di compiere una analisi delle scelte effettuate dagli utenti stradali in termini di percorso selezionato - tra i diversi possibili che collegano l'origine alla destinazione del guidatore - e orario di partenza e osservare le variazioni in queste scelte nel corso del tempo (analisi della dinamica giornaliera).

Attraverso la comprensione degli spostamenti degli individui nel tempo e nello spazio è possibile, infatti, non solo:

- individuare fattori che spieghino il comportamento di viaggio e ne scoprano le cause;
- identificare dei pattern nelle scelte dei guidatori;

ma anche rappresentarne il comportamento di viaggio attraverso modelli che permettano di prevedere i comportamenti umani (Chen, C., Ma, J., Susilo, Y., Liu, Y., Wang, M., 2016).

Tuttavia, come già menzionato, i percorsi effettivamente seguiti dai guidatori sulla rete stradale e le loro scelte in termini di orario di partenza sono stati molto raramente analizzati su larga scala e più spesso sulla base

di informazioni raccolte attraverso piccole flotte di veicoli (ad esempio taxi o campioni ristretti di guidatori coinvolti in esperimenti di laboratorio (Selten, R., Chamura, T., Pitz, T., Kube, S., Schreckenberg, M., 2007), (Tang, W.Y., Cheng, L., 2016)) oppure attraverso specifiche campagne di indagine.

Per citare alcuni esempi, S. Zhu et al. (Zhu, S., Levinson, D., 2015), attraverso l'analisi di dati GPS e GIS riguardanti gli spostamenti degli utenti su un orizzonte temporale di 13 settimane, evidenziano come gli individui spesso non selezionino il percorso più breve per raggiungere la destinazione, da ciò la necessità di perfezionare i modelli di scelta del percorso. I percorsi effettivamente seguiti dagli utenti sono confrontati con il percorso minimo fornito dal modello: in questo modo si evidenzia il divario tra l'ipotesi di comportamento di scelta e il comportamento reale. Anche L. Hainan et al. (Li, H., Guensler, R., Ogle, J., Wang, J., 2004) utilizzano dati GPS disaggregati relativi a 56 conducenti per la durata di una settimana, per analizzare il comportamento degli utenti stradali. In particolare, lo studio è incentrato sulla variabilità giornaliera dei viaggi casa-lavoro (orario di partenza, scelta del percorso e viaggi a catena). Lo studio mostra come nei viaggi casa-lavoro, immaginati altamente ripetitivi e, dunque, prevedibili, in realtà gli utenti tendano a cambiare orario di partenza, piuttosto che il percorso e come i c.d. "viaggi a catena" influenzino la scelta dell'orario di partenza e del percorso da seguire. Dall'analisi è emerso come circa il 30% degli utenti effettui fermate nel viaggio casa-lavoro. Per studiare l'orario di partenza e di arrivo, sono individuate tutte le situazioni in cui il valore assoluto della differenza tra orario di arrivo-partenza rispetto al valore mediano supera una certa soglia limite. I risultati evidenziano come, fissando la soglia a 30 minuti, circa il 30% dei viaggi presenti cambiamenti di orario e come cambiamenti nella scelta del percorso siano meno frequenti rispetto a cambiamenti

nell'orario di partenza, ma più frequenti nel caso di utenti che effettuano viaggi casa-lavoro "a catena".

P. Stopher et al. (Stopher, P., Clifford, E., Montes, M., 2008) utilizzano un set di dati GPS (raccolti in tre "wave" – la prima utilizzando un campione di 50 utenti per 28 giorni, la seconda di 46 per 28 giorni, la terza di 36 per 15 giorni) per analizzare le variazioni rispetto al valore medio al variare del set di dati e nelle stime di varianza rispetto alle variabili seguenti: distanza percorsa per ciascun viaggio, numero di viaggi per persona al giorno, tempo di viaggio per persona al giorno, distanza percorsa per persona al giorno o per passeggero e chilometri di viaggio al giorno.

A. Tawfik et al. (Tawfik, A. M., Rakha, H. A., Miller, S. D., 2010) analizzano il comportamento di un campione di 50 utenti stradali a cui è richiesto di compilare un questionario iniziale contenente una serie di informazioni personali (età, genere, problemi alla vista, etc.) e uno finale, a seguito della guida (numero di volte in cui sia necessario ripetere il viaggio casa-lavoro per individuare un percorso preferito, etc.).

I risultati confermano l'ipotesi che il comportamento degli utenti dipenda dall'esperienza e, in particolare, evidenziano quattro possibili comportamenti:

- utenti che non cambiano mai percorso;
- utenti che selezionano un percorso arbitrariamente, ripetono la loro scelta una serie di volte e non ne risultano soddisfatti. Tali utenti cambiano percorso e, ancora una volta non sono soddisfatti, tornano alla scelta iniziale e non cambiano più percorso;
- utenti che alternano le scelte di percorso continuamente, pur avendo un "percorso preferito";
- utenti che alternano scelte di percorso diverse senza mai individuare un "percorso preferito".

La stessa metodologia è utilizzata dagli autori (Tawfik, A. M., Rakha, H. A., Miller, S. D., 2010) per analizzare percezione ed esperienza degli utenti stradali nelle scelte di mobilità. I risultati, in questo caso, mostrano come:

- le percezioni dei conducenti differiscano significativamente dalle loro esperienze reali e le loro scelte siano meglio spiegate dalle percezioni piuttosto che dalle esperienze;
- i conducenti siano più sensibili alle velocità di spostamento, piuttosto che ai tempi di viaggio;
- le velocità di spostamento percepite sembrano influenzare la scelta del percorso più che i tempi di percorrenza percepiti;
- il comportamento di scelta del percorso dei conducenti differisca tra diversi gruppi di conducenti.

A. Vacca (Vacca, A., 2015) presenta un caso studio a cui ha partecipato un totale di 109 utenti, monitorati 24/24 h per un periodo complessivo di 14 giorni, durante i quali, grazie al sistema GPS negli smartphone, è stato possibile registrare tutti i percorsi effettuati e tutte le caratteristiche ad essi associate (tempi, distanze, modalità, motivo dello spostamento, etc.), per un totale di 8831 spostamenti.

I dati così raccolti sono utilizzati per analizzare la differenza tra i percorsi osservati e quelli di minimo costo, il fenomeno della variabilità dei percorsi effettuati da uno stesso utente che, nel caso di spostamenti tra la stessa coppia O-D, spesso non utilizza lo stesso itinerario e, infine, per la costruzione di un modello di scelta del percorso che include la definizione di tecniche di generazione dell'insieme di scelta.

Il fenomeno del cambio di percorso da parte degli utenti stradali in viaggio sulla stessa relazione O-D - con l'obiettivo di comprendere i fattori chiave alla base di tale meccanismo - è poi approfondito applicando i modelli di scelta del percorso implementati ai dati sulle preferenze rivelate costituiti da 677 percorsi selezionati da 77 utenti, raccolti

attraverso il GPS nella città di Cagliari in Italia (Vacca, A., Prato, C.G., Meloni, I., 2019).

I risultati ottenuti dall'applicazione di tali modelli mostrano che variabili quali l'abitudine e l'apprendimento rivestono un ruolo rilevante nel fenomeno del cambio di percorso.

C. Meneguzzer et al. (Meneguzzer, C., Olivieri, A., 2013) conducono una analisi delle dinamiche di scelta del percorso giorno per giorno su una rete di soli tre percorsi significativamente differenti l'uno dall'altro per caratteristiche geometriche e funzionali, attraverso i dati raccolti da 30 utenti su 50 intervalli rappresentativi ciascuno di una giornata. I partecipanti sono informati dei propri tempi di percorrenza sui percorsi scelti, al termine di ciascuno degli intervalli temporali. I risultati evidenziano come la frequenza di cambi nella scelta del percorso sia strettamente correlata al tempo medio impiegato per il viaggio e come entrambe le variabili non mostrino correlazione con le caratteristiche personali dei partecipanti.

W. Ciscal-Terry et al. (Ciscal-Terry, W., Dell'Amico, M., Hadjidimitriou, N.S., Iori, M., 2016) analizzano le scelte di 89 guidatori che hanno effettuato 42000 percorsi nella provincia di Reggio-Emilia, durante 17 mesi oggetto di analisi. I risultati evidenziano come gli utenti stradali tendano a preferire la strada che permette una velocità maggiore piuttosto che la più breve, ma che il comportamento degli utenti vari in funzione della tipologia di area geografica in cui l'utente effettua lo spostamento (pianura, montagna).

E. Yao et al. (Yao, 2013) – utilizzando FCD relativi a una flotta di taxi in Beijing nel 2010 – estraggono le informazioni relative a ciascun viaggio registrato, applicano un algoritmo di map-matching e identificano una serie di percorsi minimi rappresentativi del set di alternative a disposizione del guidatore. Queste informazioni sono poi utilizzate per la

calibrazione di un modello logit multinomiale che conferma la tendenza degli utenti a selezionare percorsi dalla maggiore velocità di guida, minor numero di svolte a sinistra e prevalenza di strade a scorrimento veloce.

D.J. Sun et al. (Sun, 2014) attraverso un set di FCD relativi ad una flotta di taxi, analizzano l'assunzione generale secondo cui gli utenti selezionano il percorso più breve / veloce, identificano i principali fattori che influenzano la scelta del percorso e propongono un modello predittivo di scelta del percorso applicato alla città di Shenzhen in Cina.

Come evidenziato anche in letteratura, tuttavia, sebbene i sondaggi in cui si richiede di segnalare il comportamento di viaggio su un solo giorno, siano funzionali ad una indicazione del comportamento medio di una famiglia, l'analisi di variabilità nelle scelte di mobilità giornaliera richiede campioni di dimensioni maggiori. Tuttavia, la raccolta di dati su più giorni risulta poco efficiente, se condotta attraverso sondaggi, piuttosto che utilizzando le nuove tecnologie di registrazione passiva dei dati (Stopher, PR., Kockelman, K., Greaves, SP., Clifford, E., 2008).

Alcuni autori hanno, pertanto, approfondito l'analisi comportamentale degli utenti stradali sulla base di dati raccolti passivamente.

A. Lima et al. (Lima, A., Stanojevic, R., Papagiannaki, D., Rodriguez, P., González, M. C., 2016), ad esempio, riportano un'analisi su 92419 traiettorie GPS su un orizzonte temporale di 18 mesi. I viaggi sono stati raggruppati per zone di origine e destinazione e si è osservato come gli individui tendano ad utilizzare solo alcuni percorsi "preferiti" per i viaggi frequenti e un terzo degli utenti abbia un solo percorso preferito.

In contrasto con l'assunzione del minimo costo, si è notato come una frazione significativa dei percorsi seguiti dagli utenti risulti essere non ottima.

Sono state, dunque, analizzate le deviazioni dallo "shortest path" ed è stata introdotta una densità di probabilità spaziale che i percorsi siano

confinati all'interno di una ellissi avente origine e destinazione nei punti focali di origine e destinazione del viaggio.

### 2.3 Teoria della razionalità limitata e modelli comportamentali

Le informazioni raccolte attraverso l'analisi comportamentale degli utenti stradali permettono di calibrare e validare numerosi modelli in grado di effettuare previsioni sulle scelte del percorso degli utenti.

Come già menzionato nel paragrafo precedente, la letteratura fornisce evidenza di come gli individui spesso non selezionino il percorso più breve tra le alternative a disposizione.

Da ciò deriva la necessità di perfezionare i modelli di scelta del percorso.

La classica teoria dell'utilità, infatti, prevede l'esistenza di un soggetto razionale che persegue la massimizzazione dell'utilità (guadagno di tempo, di denaro, etc.). L'individuo conosce, dunque, i costi legati a ciascuna delle alternative e reagisce selezionando l'alternativa migliore. In tale scenario non vi è alcun processo cognitivo, elaborazione delle informazioni, apprendimento o sviluppo di percezioni personali.

HA. Simon, per primo, si propone di fornire una rappresentazione più realistica del processo decisionale umano, introducendo la teoria della razionalità limitata. Individua (Simon, HA., 1972) un esempio di teoria della razionalità nella classica teoria aziendale, secondo cui l'obiettivo è massimizzare il profitto, definito come differenza tra l'incasso lordo delle vendite e i costi di produzione. In formule:

$$1) \quad \textit{Profitto} = R - C = pq - C(q)$$

In cui,  $R$  è l'incasso lordo dato dal prodotto tra prezzo ( $p$ ) e quantità di prodotto venduta ( $q$ ) e  $C$  i costi di produzione, funzione della quantità prodotta.



Ipotizzando che quantità venduta e quantità prodotta coincidano e indicandole entrambe con  $q$ , in caso di profitto massimo, si ottiene:

$$2) \quad \frac{d(R-C)}{dq} = 0$$

Il risultato è, tuttavia, calcolabile solo nel caso in cui si abbia una conoscenza perfetta dei vincoli (dati dal costo e dalla domanda) e l'equazione sia risolvibile.

La teoria della razionalità limitata nasce dall'introduzione di vincoli vari alla formulazione riportata sopra, quali l'introduzione di aspetti come rischio e incertezza nelle funzioni di costo e domanda (ad esempio, supponendo che uno o più parametri nelle loro funzioni siano variabili casuali di cui si conosce solo la distribuzione), oppure supponendo che l'individuo abbia solo informazioni incomplete riguardo le alternative, oppure ancora accrescere a tal punto il grado di complessità della funzione di costo, da rendere impossibile l'individuazione dell'alternativa migliore.

Oltre che introducendo nuovi vincoli, la teoria dell'utilità classica può essere modificata, secondo HA. Simon, trasformando la natura dell'obiettivo da raggiungere: viene riportata come esempio la teoria di Baumol, secondo cui l'azienda debba massimizzare le vendite, rispettando il vincolo per cui il profitto deve essere maggiore di un certo valore soddisfacente prefissato. Le equazioni diventano dunque:

$$3) \quad \frac{dR}{dq} = 0 \quad \text{Profitto} = R - C \geq P^*$$

Nella seconda parte dell'articolo, viene analizzato il gioco degli scacchi come microcosmo i cui processi decisionali possono essere assimilati a quelli degli individui nel mondo reale.

Particolare attenzione viene riservata al numero di strategie esaminate da ciascun giocatore: delle  $10^{140}$  mosse possibili, nella mente del giocatore

ne viene generato e analizzato solo un campione ridotto e la scelta viene effettuata non appena si individua una opzione soddisfacente.

L'individuo non può analizzare tutte le possibili alternative disponibili e questo è dovuto sia a limiti cognitivi, sia all'inadeguatezza delle informazioni a disposizione, sia alla ristrettezza dell'intervallo di tempo in cui prendere la decisione. La sua razionalità non è perfetta, assoluta, ma limitata.

Il concetto di soddisfazione fa riferimento ai processi decisionali in cui viene fissato un certo livello di aspirazione: la ricerca della soluzione viene portata avanti finché non si identifica una alternativa che soddisfi il livello di aspirazione predefinito (che può essere dinamicamente aggiornato in base agli input forniti dall'ambiente esterno).

In situazioni in cui complessità e incertezza rendono la razionalità impossibile, conclude HA. Simon, entrano in gioco soddisfazione e ottimizzazione.

L'ottimizzazione viene, però, trasformata in una versione approssimata: il caso reale viene semplificato e reso tale che il decisore possa facilmente individuare una soluzione (approssimativamente) ottima.

Nel caso della soddisfazione, d'altra parte, il caso reale viene mantenuto tale (o quasi), ma la soluzione è soddisfacente, piuttosto che ottima.

Nel caso di procedimento "search and satisficing", l'individuo genera l'insieme di scelta delle alternative secondo un procedimento euristico e le passa in rassegna terminando il processo una volta identificatane una che risulti accettabile (Simon, HA., 1979). Ovvero quando viene presa in considerazione l'alternativa a cui corrisponde un valore soddisfacente della funzione obiettivo, questa viene scelta come soluzione del processo decisionale e le altre opzioni vengono scartate.

B. D. Jones (Jones, B. D., 1999), riassume la teoria della razionalità limitata di HA. Simon nel modo seguente:

- incapacità dell'individuo nel pianificare lunghe sequenze comportamentali, causata sia dai suoi limiti cognitivi, sia dalla complessità dell'ambiente all'interno del quale opera;
- tendenza a fissare dei livelli di aspirazione per ciascuno degli obiettivi prefissati (i livelli di aspirazione fissati rispetto a ciascuno degli obiettivi del decisore, in particolare, sono approfonditi da R. Selten (Selten, R., 1999) attraverso la c.d. "aspiration adaptation theory", che rappresenta un tentativo di inquadrare in un approccio modellistico coerente il tema della razionalità limitata);
- tendenza a concentrarsi sugli obiettivi in maniera sequenziale anziché simultanea;
- comportamento che tende alla soddisfazione, anziché all'ottimizzazione.

Sottolinea, inoltre, come i modelli di utilità attesa siano poco consoni a descrivere il comportamento degli individui nei processi decisionali.

Alcuni dei motivi sono legati al decisore, ad esempio:

- l'individuo prende in considerazione solo gli aspetti più rilevanti connessi ad una decisione, per semplificare quanto più possibile il processo di scelta;
- meccanismi psicologici e fisiologici sono alla base della ricerca di attributi e alternative;
- alternative identiche, ma descritte in modo diverso sono percepite diversamente dagli utenti ("cognitive illusion and framing");
- fenomeno del self-control: l'individuo si pone dei limiti per proibirsi alcuni comportamenti in futuro;
- gli individui sono "bayesiani incompleti": in situazioni di incertezza, non aggiornano le proprie scelte in base alle informazioni in arrivo riguardanti le probabilità dei risultati, secondo le predizioni della teoria della probabilità;

- in situazioni di decisioni ripetute nel tempo, gli individui presentano la tendenza a identificare la soluzione ottima con una sub-ottima o con un valore medio e a diventare troppo conservativi per sperimentare altri, più efficienti, metodi di analisi.

Altri alla natura dell'ambiente in cui il decisore agisce:

- l'incertezza che caratterizza l'ambiente si riflette sulle preferenze degli individui: nei modelli di scelta razionali, il decisore massimizza la probabilità di ottenere un determinato risultato. Se lo stato finale è ambiguo, ecco che lo diventano anche le preferenze del decisore.
- a causa dell'interazione tra individuo e ambiente, le preferenze del decisore devono essere considerate mutevoli, non fisse, in quanto si modificano al cambiare delle circostanze.

Numerosi sono stati, in anni recenti, i casi di applicazione della teoria della razionalità limitata alle diverse fasi dei modelli di domanda di trasporto. La teoria della razionalità perfetta, infatti, è stata ampiamente utilizzata nella modellazione del comportamento di viaggio. Viceversa, la teoria della razionalità limitata ha solo recentemente riacquisito l'attenzione dei ricercatori, per via del suo potere nella modellizzazione dei comportamenti di viaggio e nella stima di previsioni più realistiche (Di, X., Liu, H.X., 2016).

Ad esempio, E. Hato et al. (Hato, E., Asakura, Y., 2000) considerano la selezione del percorso da parte dell'utente stradale un processo euristico articolato in tre fasi.

Nella prima, "network recognition", il guidatore riconosce un certo arco (o nodo) sulla rete stradale reale con una probabilità  $Q_{ij}$ , pari a:

$$4) \quad Q_{ij} = \text{prob} (C_{og_{ij}} > S)$$

Ovvero, l'arco è riconosciuto dall'individuo quando l'indice di percezione dell'arco stesso ( $Cog_{ij}$ ) supera un certo valore di soglia ( $S$ ).

Gli archi (e i nodi) a cui corrispondono valori più elevati di probabilità concorrono allo sviluppo di una mappa mentale, semplificazione di quella reale, a partire dalla quale l'individuo è effettivamente in grado di procedere con la generazione delle alternative e la selezione del percorso finale.

La seconda fase, appunto, è composta dal procedimento di enumerazione delle alternative seguito dall'associazione, a ciascuna di esse, di determinati attributi.

Infine, tutte le alternative i cui attributi non rispecchiano le condizioni imposte dal decisore sono eliminate dallo spettro delle possibili soluzioni. Il terzo step consiste nel processo di confronto delle alternative tramite i valori di utilità ad esse associate e la selezione di quella a cui corrisponde l'utilità massima, secondo il modello comportamentale standard dell'utilità attesa.

La teoria della razionalità limitata è stata, dunque, integrata nel modello di scelta del percorso per ovviare all'assunzione, tipica dei modelli di razionalità assoluta, secondo cui il decisore possa prendere in considerazione e analizzare tutte le alternative possibili dell'insieme di scelta. I limiti cognitivi impongono che possa scegliere da un cluster limitato che è creato in base alla personale percezione della rete stradale.

Altri autori hanno, invece, incorporato la teoria della razionalità limitata ai modelli comportamentali, ipotizzando che per ogni utente stradale esista una banda di indifferenza rispetto a miglioramenti o peggioramenti del tempo di percorrenza, superata la quale, gli utenti sono spinti a cambiare percorso o orario di partenza: (Sun, C., Li, M., Cheng, L., Zhu, S., Chu., Z., 2016), (Di, X., Liu, H.X., Zhu, S., Levinson, D.M., 2017),

(Eikenbroek, O.A., Still, G.J., van Berkum, E.C., Kern, W., 2018), (Watling, D.P., Rasmussen, T.K., Prato, C.G., Nielsen, O.A., 2018), (Batista, S. F. A., Zhao, C.-L., Leclercq, L., 2018), (Wang, D., Liao, F., Gao, Z., Timmermans, H., 2019), (Zhang, X., Guan, H., Zhu, H., Zhu, J., 2019), (Yu, Y., Han, K., Ochieng, W., 2020).

H. S. Mahmassani et al. (Mahmassani, H.S., Liu, Y.-H., 1999) integrano il concetto di “banda di indifferenza” in un modello probit multinomiale, al fine di tenere in considerazione la capacità di apprendimento degli utenti dalle esperienze di viaggio passate e individuare correlazioni tra decisioni ripetute prese dallo stesso utente.

Il cambiamento di orario di partenza viene schematizzato ipotizzando che il viaggiatore  $i$  non cambi l’orario di partenza nei giorni successivi fintanto che il ritardo sperimentato nel giorno  $t$ ,  $SD_{it}$ , (differenza tra il tempo di arrivo preventivato  $PAT_i$  e quello realmente impiegato  $AT_{it}$ ) rimanga nell’intervallo di indifferenza  $IBD_{it}$  (con due diverse componenti  $EDB_{it}$  e  $LBD_{it}$ , che definiscono l’arrivo in anticipo e in ritardo, rispettivamente).

Una variabile  $\delta$  viene utilizzata per indicare variazioni di orario di partenza. Vale  $1$  quando l’utente  $i$  cambia percorso dopo il ritardo del giorno  $t-1$ , uguale a  $-1$  altrimenti.

$EDB_{it}$  e  $LBD_{it}$  sono due variabili casuali con componente sistematica e casuale. La componente sistematica dipende dagli attributi dell’utente e dalla sua esperienza pregressa. I termini casuali si assume abbiano distribuzione normale.

Il meccanismo di cambiamento di percorso viene, invece, schematizzato supponendo che il viaggiatore  $i$  non cambi percorso nei giorni successivi fintanto che il guadagno di tempo sperimentato nel giorno  $t$ ,  $TTS_{ijt}$ , (differenza di tempo di percorrenza tra il percorso attuale  $TTC_{ijt}$  e quello minimo  $TTB_{ijt}$ ) rimanga contenuto nell’intervallo di indifferenza  $IBD_{it}$ .

Una variabile  $\Phi_{ijt}$  viene utilizzata come indicatore di cambio di percorso. È uguale a 1 quando l'utente  $i$  cambia nel giorno  $t$  il percorso, in seguito al viaggio nel giorno  $t-1$ , uguale a -1 in caso contrario.

L'intervallo di indifferenza è dato in questo caso da:

$$5) \quad IBR_{ijt} = \max [\eta_{ijt} TTC_{ijt} \pi_{ijt}]$$

In cui,  $\eta_{ijt}$  e  $\pi_{ijt}$  sono due variabili casuali con componente sistematica e casuale. La componente sistematica dipende dagli attributi dell'utente e dalla sua esperienza pregressa. I termini casuali si assume abbiano distribuzione normale.

Il modello è stato poi calibrato attraverso il metodo della massima verosimiglianza.

I risultati mostrano come gli utenti tollerino i ritardi in maniera diversificata in base all'età o al genere, come l'affidabilità delle informazioni in tempo reale influenzi il comportamento degli utenti, la loro tendenza ad effettuare cambiamenti nell'orario o nel percorso in seguito ad aver sperimentato un ritardo nell'orario di arrivo, piuttosto che un anticipo e, infine, la propensione a tollerare i ritardi nel caso in cui si siano recentemente sperimentati significativi aumenti nel tempo di viaggio dopo una piccola modifica dell'orario di partenza.

R-C. Jou et al. (Jou, R., Lam, S., Liu, Y., Chen, K., 2005) sviluppano un modello ispirato alla razionalità limitata di H. S. Mahmassani in grado di rappresentare i cambiamenti dinamici di comportamento dei guidatori sulle autostrade in seguito ad informazioni di traffico in real-time.

L'assunzione è che gli utenti tendano a non cambiare percorso se le condizioni di traffico nel segmento di autostrada che stanno percorrendo ricadono nell'intervallo di indifferenza per il cambiamento di percorso (IBR). L'IBR è definito come il range massimo tollerabile di guadagno per cui gli utenti non sono portati a cambiare percorso e quest'ultimo come il percorso di costo generalizzato minimo in prossimità dell'area di studio.

La differenza con il modello di H. S. Mahmassani risiede nella combinazione di tempi di viaggio e costi.

Il modello viene formulato nel modo seguente:

$$6) \quad GTS_{ij} = \alpha_{ij}TCS_{ij} + \beta_{ij}TTS_{ij}$$

In cui,  $GTS_{ij}$  è il guadagno di costo generalizzato,  $TCS_{ij}$  è il guadagno in termini di costo e  $TTS_{ij}$  il guadagno in termini di tempo di percorrenza.

$\alpha_{ij}$  e  $\beta_{ij}$  sono i due pesi.

Una variabile  $\Phi_{ijt}$  viene utilizzata come indicatore di cambio di percorso.

È uguale a 1 quando l'utente i cambia percorso, uguale a -1 in caso contrario.

L'intervallo di indifferenza è dato nuovamente da:

$$7) \quad IBR_{ijt} = \max [\eta_{ijt}TTC_{ijt}\pi_{ijt}]$$

In cui,  $\eta_{ijt}$  e  $\pi_{ijt}$  sono due variabili casuali con componente sistematica e casuale. La componente sistematica dipende dagli attributi dell'utente, dalla sua esperienza pregressa e da quattro tipi di informazioni che gli vengono fornite. I termini casuali si assume abbiano distribuzione normale.

L'aspetto principale di questo modello è che combinando tempi di viaggio e costi in un unico costo generalizzato, presuppone che gli utenti decidano di cambiare percorso tenendo in considerazione entrambe le variabili.

K. Han et al. (Han, K., Szeto, W.Y., Friesz, T.L., 2015), recentemente, approfondiscono i modelli di "dynamic user equilibrium" (DUE) a cui viene applicato il concetto di razionalità limitata (BR) proposto da H.A. Simon, in modo da incorporare il comportamento non completamente razionale degli utenti stradali nei modelli comportamentali di viaggio. Nei modelli BR-DUE la percezione dei costi di viaggio è basata su un intervallo di indifferenza che rispecchia la tolleranza dei guidatori nello



sperimentare una differenza tra il percorso effettuato e quello ottimo e la tolleranza è legata alla coppia O-D.

Una estensione di questi modelli è rappresentata dai “variable tolerance bounded rational dynamic user equilibrium” (VT-BR-DUE), in cui la tolleranza degli utenti è legata non solo alla coppia O-D, ma anche a fattori “esterni” come qualità della strada, distanza di viaggio, panorama, familiarità con il percorso, condizioni di traffico, orario di partenza.

Il modello sviluppato rappresenta un approccio unificato per modellizzare sia BR-DUE sia VT-BR-DUE.

W. Shang et al. (Shang, W., Han, K., Ochieng, W., Angeloudis, P., 2017) sviluppano un modello in grado di esplorare gli impatti delle informazioni di viaggio fornite a utenti stradali. Il modello è costituito da due layer, il “cyber layer” che rappresenta la rete di comunicazione attraverso cui le informazioni sono condivise tra guidatori, e il “physical layer” rappresenta l’evoluzione, giorno per giorno, della rete stradale. In studi passati si era supposto che gli utenti venissero informati tramite sistemi del tipo network operator, ATIS, Google Map, etc. In questo caso, si prende in considerazione una nuova fonte di informazione, che è quella acquisita dall’individuo nell’effettuare gli spostamenti giornalieri e diffusa tramite nuovi sistemi di informazione, quali ad esempio, gli strumenti che offrono la possibilità di navigare in internet (come quelli offerti da TomTom o Garmin) oppure le app per smartphone con caratteristiche di social network (Waze).

In questo modello si assume che, invece di conoscere esattamente costi/tempi di viaggio per ogni percorso, gli utenti formino gruppi all’interno dei quali condividere le informazioni di viaggio. In questo modo le loro scelte non sono solo influenzate dalla propria esperienza di viaggio, ma anche da quelle degli altri componenti del gruppo di appartenenza.

Anche in questo caso, la teoria della razionalità limitata viene introdotta nel modello di scelta del percorso con la stessa formulazione di H. S. Mahmassani: se la differenza tra il costo di viaggio percepito dall'utente nel percorso attuale e il costo di un tragitto alternativo non supera una certa soglia, egli è spinto a riconfermare la scelta corrente, altrimenti considererà tutte le alternative sulla base di un modello logit.

L'ipotesi secondo cui le scelte degli utenti stradali sono influenzate da quale percorso e quale orario di partenza selezionano gli altri guidatori era già stata approfondita, ad esempio da Q. Han et al. (Han, Q., Timmermans, H., 2006): assumendo vera questa ipotesi, allora la valutazione delle alternative di viaggio non viene fatta tenendo in conto esclusivamente i tempi di percorrenza sperimentati in prima persona, ma anche quelli sperimentati da altri utenti stradali su altri cammini (acquisiti tramite sistemi di informazione in real-time).

Quindi, l'utilità, per un utente  $i$  qualsiasi, di aver scelto l'alternativa  $j$  è:

$$8) \quad U_{ij} = U_j^{exogenous} + U_j^{endogenous} + \varepsilon_{ij}$$

In cui  $U_j^{exogenous}$  rappresenta la componente specifica dell'alternativa e  $U_j^{endogenous}$  la componente di interazione che cattura l'impatto atteso dei comportamenti di scelta degli altri viaggiatori.

Togliendo  $U_j^{endogenous}$ , l'utilità torna a essere quella di un convenzionale modello di scelta discreta, ma prendendolo in considerazione, si introduce nel modello l'ipotesi che l'attrattività di un'alternativa sia funzione anche delle scelte degli altri utenti.

Gli autori considerano un modello di apprendimento piuttosto semplice (reinforcement learning), secondo cui un utente aggiorna la propria opinione su un'alternativa, basandosi sulle proprie esperienze pregresse:

$$9) \quad E_{ij}(t) = \frac{\lambda \cdot N(t-1) \cdot E_{ij}(t-1) + U[s_{ij}, s_{-i}(t)]}{\lambda \cdot N(t-1) + 1}$$

$$10) \quad N(t) = \lambda \cdot N(t-1) + 1$$

$U[s_{ij}, s_{-i}(t)]$  rappresenta l'utilità per l'utente  $i$  di scegliere l'alternativa  $j$ , mentre contemporaneamente gli altri utenti selezionano le proprie strategie.

$N(t)$  è l'indicatore delle esperienze passate messe in relazione con quella attuale.  $N(0)$  è il peso della prima esperienza e riflette la convinzione dell'utente che la propria convinzione sia riconfermata: se il suo valore è basso, l'esperienza sostituirà subito la convinzione iniziale, altrimenti la percezione iniziale verrà mantenuta durante tutto il processo di apprendimento.

$E_{ij}(t)$  è l'aspettativa attuale, è calcolata come somma pesata delle aspettative passate e delle esperienze passate, con pesi  $\lambda \cdot \frac{N(t-1)}{[\lambda \cdot N(t-1)+1]}$  e  $\frac{1}{[\lambda \cdot N(t-1)+1]}$  rispettivamente.

Il parametro  $\lambda$  riflette la tendenza a diminuire il peso associato a esperienze passate e aumentare quello associato ad altre più recenti. Questo è dovuto al fatto che gli utenti possono dimenticare esperienze lontane nel tempo o comunque, essendo coscienti della mutevolezza dell'ambiente circostante, tendono a considerare più probabile che esperienze recenti possano riflettere meglio lo scenario corrente. La probabilità che l'utente  $i$  scelga l'alternativa  $j$  è dunque funzione dell'utilità attuale attesa, costantemente aggiornata dalle esperienze dell'utente, secondo un modello logit:

$$11) \quad P_{ij} = \frac{\exp\left[\frac{E_{ij}(\lambda)}{\mu}\right]}{\sum_{j'} \exp\left[\frac{E_{ij'}(\lambda)}{\mu}\right]}$$

In cui  $\mu$  è un parametro che rispecchia l'errore dovuto all'incertezza delle convinzioni relative al comportamento degli altri utenti. Data l'assunzione che gli individui non siano né perfettamente razionali ( $\mu=0$ ), né completamente irrazionali ( $\mu=\infty$ ), i valori intermedi di  $\mu$  sono i più adatti a rappresentare il comportamento umano. Dunque, il parametro

rappresenta uno strumento per modellizzare la “razionalità limitata” degli utenti stradali.

Anche E. A. I. Bogers et al. (Bogers, 2005), presentano un modello di scelta del percorso mixed multinomial logit model che tiene conto di fattori quali abitudini degli individui, apprendimento dalle esperienze pregresse, attitudine al rischio e importanza di informazioni di viaggio in real-time.

I risultati degli esperimenti di laboratorio mostrano come gli individui scelgano l’alternativa migliore nel caso in cui le informazioni fornitegli siano più elaborate, come abitudine e inerzia insieme a informazioni in real-time lungo il percorso giocino un ruolo fondamentale nella scelta del cammino e come l’apprendimento degli attributi relativi alle varie alternative rappresenti un aspetto essenziale nei primi giorni, ma con il passare del tempo venga surclassato da abitudine e sistemi di informazione.

T. Arentze et al. (Arentze, T., Timmermans, H., 2003) sviluppano un modello generale in grado di rappresentare le scelte dinamiche degli utenti, tenendo conto la teoria dell’apprendimento per rinforzo e dell’adattamento.

L’ipotesi di base è che il sistema di trasporto e l’ambiente urbano siano caratterizzati da dinamicità, non-stazionarietà e incertezza, mentre gli utenti stradali debbano effettuare scelte principalmente basate sull’esperienza acquisita giorno per giorno. Le esperienze pregresse costituiscono il mezzo attraverso cui gli individui si adattano alla mutevolezza dell’ambiente esterno e vengono immagazzinate nella mente dell’individuo assieme a un valore associatogli, detto “ricompensa”.

Il goal è massimizzare la ricompensa ricevuta entro un certo orizzonte di tempo.

Tranne per la dinamicità del modello, il concetto è fortemente equiparabile alla funzione di utilità nei modelli di scelta di utilità casuale.

La ricompensa può variare nel tempo o essere soggetta a variazioni casuali.

Ogni esperienza che produce una ricompensa ed è immagazzinata nella memoria è un evento. L'evento è definito come:

$$12) \quad e = e(x, r, t, w, \rho)$$

In cui:

- $x$  sono gli attributi, ovvero tutte le caratteristiche che descrivono l'evento stesso. Generalmente l'uomo, tra tutti gli attributi possibili, ne considera un sottoinsieme che potrebbe risultare potenzialmente rilevante nell'effettuare la scelta.
- $r$  è la ricompensa: è un valore numerico che indica l'intrinseca desiderabilità di un evento. In questo caso è stata formulata come la somma di una funzione, che varia nel tempo, degli attributi e una componente di errore stocastica.
- $w$  è il vettore dei pesi: rappresenta il legame con cui l'individuo congiunge attributi e ricompensa: per un peso uguale a zero, l'individuo crederà che il valore della ricompensa non sia sistematicamente legato agli attributi, mentre per un peso uguale a 1, che la stessa ricompensa si ottenga nel caso in cui si ripeta l'evento.
- $\rho$  è la "recuperabilità", ovvero la facilità con cui un evento passato viene ripescato dalla memoria.

La scelta dell'utente ricadrà dunque sulla soluzione che massimizza la ricompensa attesa  $Q(c, a)$ , ovvero la ricompensa di scegliere l'azione  $a$  nello stato  $c$ .

E. Avineri et al. (Avineri, E., Prashker, J.N., 2005) comparano quattro modelli predittivi, di cui due statici (massima utilità casuale e teoria del prospetto cumulativa) e due dinamici (modello di apprendimento Bayesiano e modello di apprendimento per rinforzo):

1. il primo modello è caratterizzato dall'attribuzione di una componente casuale all'utilità associata a una alternativa: questa variabile casuale rappresenta l'incertezza nella mente del decisore (errori di valutazione o informazioni insufficienti riguardo le alternative), o incertezze nella mente del modellatore (attributi che non sono stati presi in considerazione o errori di misura).
2. La teoria del prospetto è stata sviluppata da D. Kahneman e A. Tversky (1979), per spiegare le contraddizioni frequenti alle predizioni dei modelli di massima utilità. Gli esperimenti svolti da D. Kahneman e A. Tversky si concentravano principalmente sull'analisi dell'attitudine al rischio degli utenti: è opportuno scegliere una nuova alternativa dall'utilità incognita oppure ripetere la scelta a cui si è avvezzi per non rischiare di sperimentare ritardi? I risultati mostrano come gli utenti siano scoraggiati molto più da una potenziale perdita, di quanto non siano attratti da una potenziale vincita.

Le alternative vengono valutate in un processo a due stadi: "editing" ed "evaluation".

Nella prima fase, le alternative vengono classificate come potenziali vincite o perdite in base a un valore di soglia stabilito dal decisore. Nella seconda, il valore dell'alternativa viene calcolato utilizzando due funzioni, una funzione di valore  $v(\bullet)$  e una funzione di ponderazione  $\pi(\bullet)$ . Considerando una alternativa con tre possibili risultati:  $x$  con una probabilità  $p$ ,  $y$  con una probabilità

$q$  e lo status quo con una probabilità  $1-p-q$  il valore dell'alternativa secondo la teoria del prospetto è dato da:

$$13) \quad \pi(p)v(x) + \pi(q)v(y)$$

D. Kahneman e A. Tversky hanno poi sviluppato una nuova versione di teoria del prospetto cumulativa che considera una distribuzione cumulata delle probabilità (rappresentate dalla funzione di ponderazione).

3. Per quanto riguarda i modelli dinamici, il modello di apprendimento Bayesiano suppone che il processo di scelta del percorso sia iterativo e, a ogni step, il guidatore utilizzi le frequenze storiche dei tempi di percorrenza sperimentati per calcolare un valore atteso del tempo di percorrenza per ciascuno dei percorsi alternativi, in base al quale selezionare quello che massimizzi la sua utilità.
4. Un modello che tiene conto di come gli utenti si adattino alla variabilità dei tempi di percorrenza, infine, è il modello di apprendimento per rinforzo (I. Erev et al., 1999), che può essere sintetizzato in tre fasi (Gigerenzer, G., Selten, R. Editors, 2002):
  - a. L'individuo ha una propensione iniziale a scegliere una certa strategia  $j$ , pari a  $q_j(1)$ . Al tempo  $t = 1$  (prima di acquisire qualunque esperienza), è indifferente a qualunque alternativa. Si suppone che, se tutti gli individui selezionano una alternativa in modo casuale e tutte le alternative hanno la stessa probabilità di essere selezionate, la ricompensa attesa per l'individuo  $n$  sia  $A_n(1)$ . Ovvero, la propensione a selezionare la strategia  $j$  è pari alla ricompensa attesa in seguito a scelte casuali:
 
$$14) \quad q_{nj}(1) = A_n(1) \text{ per ogni strategia } j$$
  - b. La propensione dell'individuo  $n$  a scegliere la strategia  $j$  al tempo  $t + 1$ , è la somma pesata della propensione iniziale

$(q_{nj}(1))$  e la ricompensa media ottenuta nello scegliere  $j$  fino al tempo  $t$  ( $AVE_{nj}(t)$ ).

Il peso assegnato alla propensione iniziale è una funzione del parametro che rappresenta la forza della propensione iniziale ( $N(1)$ ) e il numero di volte in cui l'individuo  $n$  ha scelto la strategia  $j$  fino al tempo  $t$ , ( $C_{nj}(t)$ ):

$$15) \quad q_{nj}(t+1) = q_{nj}(1) \frac{N(1)/m_n}{C_{nj}(t)+N(1)/m_n} + AVE_{nj}(t) \frac{C_{nj}(t)}{C_{nj}(t)+N(1)/m_n}$$

In cui  $m_n$  è il numero di strategie pure dell'individuo.

L'assunzione dell'aggiornamento della media, può essere descritta come segue: se l'individuo  $n$  sceglie la strategia  $j$  al tempo  $t$  e riceve la ricompensa  $x$ , allora la propensione dell'individuo  $n$  a scegliere la strategia  $j$  viene aggiornata:

$$16) \quad q_{nj}(t+1) = q_{nj}(1) W(t) + x(1 - W(t))$$

In cui:

$$17) \quad W(t) = \frac{C_{nj}(t)+N(1)/m_n}{C_{nj}(t)+N(1)/m_n+1}$$

Questo procedimento per tentativi rivela che la propensione a scegliere l'alternativa  $j$  aumenta con ricompense positive ( $x < q_{nj}(t)$ ) e decresce con basse ricompense ( $x > q_{nj}(t)$ ). Infine, si sottolinea come la velocità di aggiornamento diminuisca con l'esperienza ( $W(t)$  aumenti con  $C_{nj}(t)$ ).

- c. La probabilità che l'individuo  $n$  scelga la strategia  $j$  al tempo  $t$  è data dalla regola di risposta esponenziale:

$$18) \quad P_{nj}(t) = \frac{EXP\left[\frac{q_{nj}(t)\lambda}{S_n(t)}\right]}{\sum EXP\left[\frac{q_{nk}(t)\lambda}{S_n(t)}\right]}$$

In cui la somma viene effettuata su tutte le  $k$  strategie pure possibili per l'individuo,  $\lambda$  è un parametro che determina la sensibilità al rinforzo e  $S_n(t)$  è la misura della deviazione



standard delle ricompense che l'individuo ha sperimentato fino all'istante  $t$ .

Dunque, la probabilità di scegliere una strategia aumenta all'aumentare della propensione a selezionarla (che aumenta con la ricompensa media derivata dalle scelte passate).

La deviazione standard è stimata come la differenza assoluta media tra la ricompensa recente ( $x$  al tempo  $t$ ) e la ricompensa media accumulata fino a  $t$  ( $A_n(t)$ ).

$$19) \quad S_n(t+1) = S_n(t)W'(t) + |A_n(t) - x|(1 - W'(t))$$

$$20) \quad W'(t) = \frac{t+N(1)}{t+N(1)+1}$$

Il valore iniziale  $S_n(1)$  è la differenza attesa della ricompensa dal valore ottenuto in seguito a scelte casuali e il valore atteso date le scelte casuali (da notare che il modello è definito solo per  $S_n(1)$  positivi).

$A_n(t)$  si calcola similmente:

$$21) \quad A_n(t+1) = A_n(t)W'(t) + A_n(t)(1 - W'(t))$$

### 3 Approccio metodologico

Come già menzionato nelle sezioni introduttive della presente tesi di dottorato, al fine di analizzare le variazioni nelle scelte degli utenti stradali (analisi della dinamica giornaliera) attraverso l'elaborazione di un set esteso di FCD, è stata definita una metodologia generale che prevede:

1. la ricostruzione del percorso seguito dall'utente a partire dalle posizioni sparse registrate;
2. la generazione di un set ristretto di alternative di viaggio per ciascuna coppia O-D, basata su una analisi di "dissimilarità" dei percorsi selezionati dagli utenti stradali.

Per quanto riguarda il punto 1) del precedente elenco, come già evidenziato, la gestione di grandi set di FCD comporta necessarie pre-elaborazioni (Fusco, G., Colombaroni, C., Comelli, L., Isaenko, N., 2015). Tali dati sono, infatti, rappresentati da una serie di coordinate geografiche sparse e devono, in primo luogo, essere messi in relazione con il grafo stradale (map-matching) per possibili applicazioni nei modelli di trasporto (Isaenko, N., Colombaroni, C., Fusco, G., 2017).

Sebbene molti algoritmi di map-matching siano stati sviluppati negli ultimi anni in particolare per i sistemi di navigazione, questi ultimi non sono adatti per analisi di modellazione. I FCD, infatti, vengono solitamente raccolti con frequenze inferiori a quelle applicate dai sistemi di navigazione di bordo. Si pone, così, il tema della ricostruzione del percorso seguito dal veicolo tra due punti di campionamento successivi, raccolti anche ogni 1 o 2 minuti (Chen, F., Shen, M., Tang, Y., 2011). M. Rahmani et al. (Rahmani, M., Koutsopoulos, H., 2013), ad esempio, sviluppano un metodo a due-step per individuare un insieme di collegamenti candidati e poi selezionare il percorso più probabile.

E. Frejinger et al. (Frejinger, E., Bierlaire, M., 2007) introducono il concetto di “sotto-rete”, in grado di catturare la correlazione tra percorsi simili sulla rete. La “sotto-rete” rappresenta una semplificazione della rete stradale che include solo i percorsi facilmente identificabili e ragionevolmente rilevanti per gli utenti.

M. Bierlaire et al. (Bierlaire, M., Chen, J., Newman, J., 2013) risolvono il problema implementando un algoritmo di map-matching probabilistico che genera un set di percorsi potenziali e associa a ciascuno di essi un valore di verosimiglianza.

Chen et al. (Chen, F., Shen, M., Tang, Y., 2011) propongono un algoritmo di map-matching basato su una ricerca di percorso “locale”: l’informazione dell’ultima posizione registrata è utilizzata per ridurre lo spazio di ricerca ad un’area di confidenza quadrata, consentendo così di determinare la traiettoria del veicolo.

J. Li et al. (Li, J., Xie, L., Lai, X., 2013) introducono un metodo di probabilità a due livelli per la ricostruzione dei percorsi a partire dai FCD, che risolve da un lato gli errori del map-matching dovuti alla complessità del grafo e all’accuratezza del GPS e, dall’altro, il problema dei collegamenti mancanti a causa della bassa frequenza di campionamento, ricostruendo il percorso tra due posizioni consecutive con l'algoritmo del percorso più breve.

X. Liu et al. (Liu, X., Liu, K., Li, M., Lu, F., 2016) implementano un metodo di map-matching per i dati a bassa frequenza di campionamento (ad esempio, un punto GPS ogni 1-2 minuti). L'algoritmo tiene conto di diversi fattori, come l'accuratezza del posizionamento spaziale dei punti GPS e le informazioni topologiche della rete stradale sottostante, la coerenza della direzione di guida di una traiettoria GPS, etc.

La ricostruzione dei percorsi seguiti dagli utenti stradali attraverso dati FCD, però, può fornire come risultato una quantità notevole di itinerari di connessione tra una origine e una destinazione, che spesso differiscono per piccole porzioni di percorso gli uni dagli altri.

In letteratura, il tema della selezione di un numero limitato di percorsi significativamente diversi che rappresentano le scelte di percorso dei conducenti con il livello di precisione richiesto dai modelli di traffico è stato affrontato, anche recentemente, da numerosi autori.

Ad esempio, J. Kim et al. (Kim, J., Mahmassani, H.S., 2015) definiscono un metodo di clustering dei percorsi, finalizzato a identificare degli schemi nelle abitudini di viaggio degli utenti stradali, che non richiede lo step del map-matching, ma analizza i dati di traiettoria dei veicoli senza le informazioni della rete stradale sottostante.

Numerosi sono inoltre, in letteratura, gli indicatori che misurano la dissimilarità delle alternative di percorso, con l'obiettivo di ridimensionare il set di scelta.

Tra questi, V. Akgün et al. (Akgün, V., Erkut, E., Batta, R., 2000) introducono la dissimilarità in termini di lunghezza della porzione condivisa di percorso tra due itinerari.

P. Dell'Olmo et al. (Dell'Olmo, P., Gentili, M., Scozzari, A., 2005) utilizzano il concetto di "zona cuscinetto" per caratterizzare percorsi eterogenei.

R. Martí et al. (Martí, R., Velarde, J.L., Duarte, A., 2009) propongono un indicatore che supera le criticità legate alla "zona cuscinetto" e tiene conto del comportamento di scelta dei conducenti.

Il problema della dissimilarità dei percorsi è, poi, strettamente correlato all'analisi di covarianza.

E. Cascetta et al. (Cascetta, E., Nuzzolo, A., Russo, F., Vitetta, A., 1996) catturano, tra i primi, la correlazione tra le alternative di percorso in modo

esplicito introducendo un fattore nella componente deterministica della formulazione del modello logit, che è proporzionale alla lunghezza degli archi condivisi da ciascun percorso con gli altri percorsi nel set di scelta. Ulteriori contributi sono dovuti, tra gli altri, a S. Bekhor et al. che adattano un modello di logit kernel al problema della scelta del percorso (Bekhor, 2002) e a E. Cascetta et al. che assimilano l'insieme di scelte ad un "fuzzy set" in un modello implicito di disponibilità / percezione delle alternative di scelta (Cascetta, E., Papola, A., 2001).

Nei capitoli seguenti è descritto l'approccio sperimentale utilizzato per la ricostruzione dei percorsi e la generazione dell'insieme di scelta per l'utente stradale.

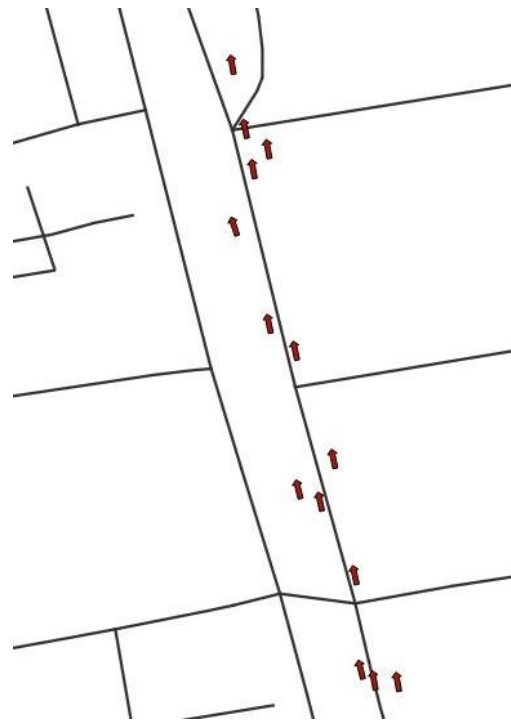
L'output della metodologia è rappresentato da un set ristretto di percorsi effettivamente percepiti come "alternativi" dagli utenti stradali.

L'approccio è stato sviluppato, appunto, nell'ottica di integrare la teoria della razionalità limitata nella modellizzazione della scelta del percorso e ovviare, così, all'assunzione secondo cui il decisore possa prendere in considerazione e analizzare tutte le alternative possibili dell'insieme di scelta.

I risultati ottenuti dall'applicazione della metodologia sono stati, poi, utilizzati per effettuare l'analisi della dinamica giornaliera (il viaggiatore rivede la sua scelta precedente se si aspetta di ottenere un vantaggio dal cambiamento? Oppure, viceversa, conferma la scelta abituale individuata come "migliore" sulla base della propria esperienza di viaggio?).

### 3.1 Ricostruzione del percorso

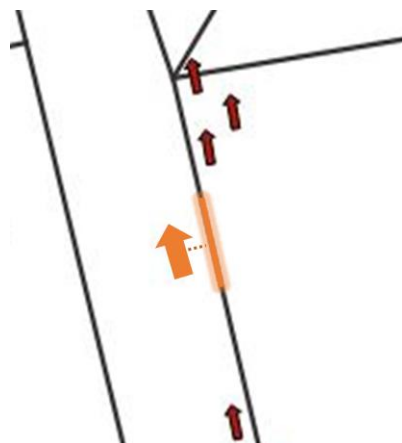
Come già menzionato, i FCD sono rappresentati da una serie di record geolocalizzati che includono alcune informazioni, tra cui: la velocità puntuale, la direzione di marcia e data e ora di registrazione del dato.



*Figura 1 – Esempio di FCD grezzo*

Come si può notare dalle immagini - per possibili applicazioni nei modelli di trasporto - i dati grezzi generalmente richiedono delle elaborazioni preliminari per essere messi in relazione con il grafo stradale sottostante da cui risultano svincolati (map-matching).

Questo procedimento prevede una prima fase di assegnazione di ciascun dato all'arco del grafo stradale cui più probabilmente appartiene.



*Figura 2 – Rappresentazione esemplificativa della procedura di assegnazione del FCD all'arco del grafo di appartenenza*

I FCD utilizzati nell'ambito della presente ricerca sono già assegnati agli archi del grafo stradale e, in particolare, risultano caratterizzati da alcuni campi, tra cui:

- "gid": codice identificativo del record;
- "x": latitudine nel sistema di riferimento WGS84;
- "y": longitudine nel sistema di riferimento WGS84;
- "dir": direzione di marcia misurata in gradi in senso orario rispetto al Nord;
- "vel": velocità puntuale in km/h;
- "dt": timestamp, ovvero data e ora della registrazione;
- "stato": stato del motore del veicolo (0: accensione, 1: moto, 2: spegnimento);
- "quality": indicatore qualitativo del segnale GPS (1: no segnale, 2: basso, 3: buono);
- "id\_veicolo": codice identificativo del veicolo (periodicamente aggiornato per ragioni di privacy);
- "classe\_veicolo": tipologia di veicolo (A: auto privata, B: bus, C: veicolo commerciale);
- "odo": distanza percorsa dal veicolo;
- "mm\_fid": codice identificativo dell'arco stradale a cui il dato appartiene<sup>1</sup>.

Pertanto, a partire dal database a disposizione, sono stati identificati dei "sottoinsiemi di dati" che costituiscono viaggi effettuati dagli utenti stradali, ovvero sequenze di record registrate durante la stessa giornata e caratterizzate dallo stesso "ID veicolo".

---

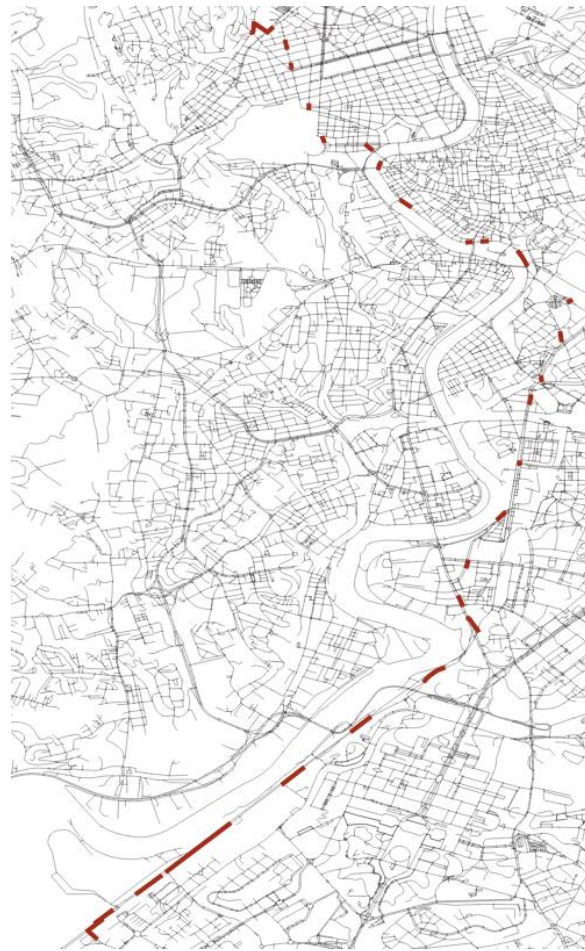
<sup>1</sup> I dati sono assegnati agli archi stradali del grafo Teleatlas 2013.

Sono stati, infine, selezionati gli archi a cui tali FCD risultano assegnati e, per ciascuna coppia di archi consecutivi appartenenti allo stesso sottoinsieme di dati (allo stesso viaggio), è stato ricostruito il percorso seguito dall'utente (traiettorie tratteggiate in nero nell'immagine seguente), ipotizzando che quest'ultimo abbia selezionato l'alternativa più breve in termini di distanza percorsa.

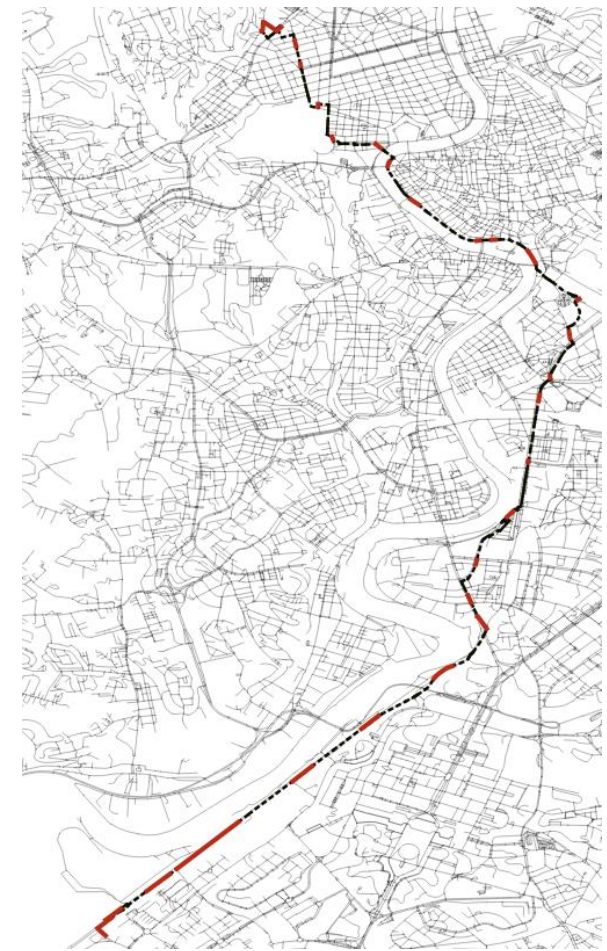




*Figura 3 – Sottoinsieme di FCD rappresentativo di un viaggio*



*Figura 4 – Sequenza di archi del grafo a cui i FCD sono associati*



*Figura 5 – Ricostruzione del percorso minimo tra coppie di archi consecutivi*

Il processo prevede i seguenti step:

- *Step 1:* selezione del nodo iniziale e finale di ciascuno degli archi a cui risultano assegnati i FCD dell'itinerario da ricostruire, ovvero identificazione della sequenza di nodi noti attraversati dall'utente;
- *Step 2:* ricostruzione del percorso minimo tra il nodo finale di un arco a cui è assegnato il FCD e il nodo iniziale dell'arco a cui è assegnato il successivo;
- *Step 3:* iterazione della procedura di ricostruzione del percorso minimo finché non è stata completata la ricostruzione dell'intero percorso (sequenza di archi del grafo).

La ricostruzione del percorso minimo è stata effettuata attraverso la funzione "graphshortestpath" di MATLAB, che identifica univocamente il percorso minimo tra due nodi del grafo secondo l'algoritmo di Dijkstra. Il grafo stradale - dato di input della routine - deve essere caratterizzato attraverso una matrice sparsa, come set di archi caratterizzati da un nodo iniziale, un nodo finale e un peso (la lunghezza).

I risultati hanno mostrato una rappresentazione accurata degli itinerari percorsi dagli utenti stradali: nell'80% dei casi, sul totale dei percorsi ricostruiti, la lunghezza del percorso ricostruito differisce in media del 6% (circa 900 m.) rispetto alla lunghezza del percorso fornita dal FCD (campo "odo").

Si sottolinea, infine, come spesso i tempi necessari a processare grandi database di FCD rappresentino una criticità.

Il tempo per la ricostruzione di un percorso attraverso la routine implementata varia con il livello di dettaglio del grafo e con la lunghezza del percorso. La routine impiega 10.5 secondi per la ricostruzione di un percorso lungo circa 6 km sul grafo Teleatlas 2013 esteso al comune della città di Roma, utilizzando un i7-6500U 2.50 GHz con 12 GB di RAM.

### 3.2 Definizione delle alternative di scelta del percorso

Come già menzionato, i percorsi seguiti dagli utenti stradali da una zona di origine a una di destinazione possono essere numerosi e differire gli uni dagli altri per porzioni molto limitate.

Per tale ragione, è stata identificata una procedura per la ripartizione dei percorsi in cluster effettivamente percepiti come alternativi dagli utenti stradali, ovvero una procedura in grado di generare un insieme ristretto di scelta del percorso per il guidatore.

Il processo prende in input i risultati della routine di ricostruzione del percorso (l'insieme dei percorsi che gli utenti hanno selezionato durante viaggi differenti per spostarsi da una zona di origine a una di destinazione).

Il problema di identificare dei sottoinsiemi di itinerari che possono essere percepiti dall'utente stradale come alternativi è risolto tramite un algoritmo euristico che assegna i percorsi a diversi cluster e seleziona un percorso rappresentativo per ciascuno di questi ultimi.

Il criterio utilizzato per la clusterizzazione consiste nel massimizzare la dissimilarità tra percorsi appartenenti a cluster diversi e minimizzare la dissimilarità tra percorsi appartenenti allo stesso set.

A tal proposito, è stato utilizzato il seguente indice di dissimilarità  $D(i, j)$  tra il percorso  $i$  e il percorso  $j$ :

$$22) \quad D(i, j) = 1 - \frac{1}{2} \left( \frac{L(P_i \cap P_j)}{L(P_i)} + \frac{L(P_i \cap P_j)}{L(P_j)} \right)$$

In cui:

- $L(P_i)$  è la lunghezza del percorso  $i$ ;
- $L(P_i \cap P_j)$  è la lunghezza del percorso che  $i$  e  $j$  hanno in comune.

Bassi valori dell'indice indicano percorsi con larghe porzioni in comune, mentre valori dell'indice prossimi all'unità indicano percorsi completamente differenti tra loro.



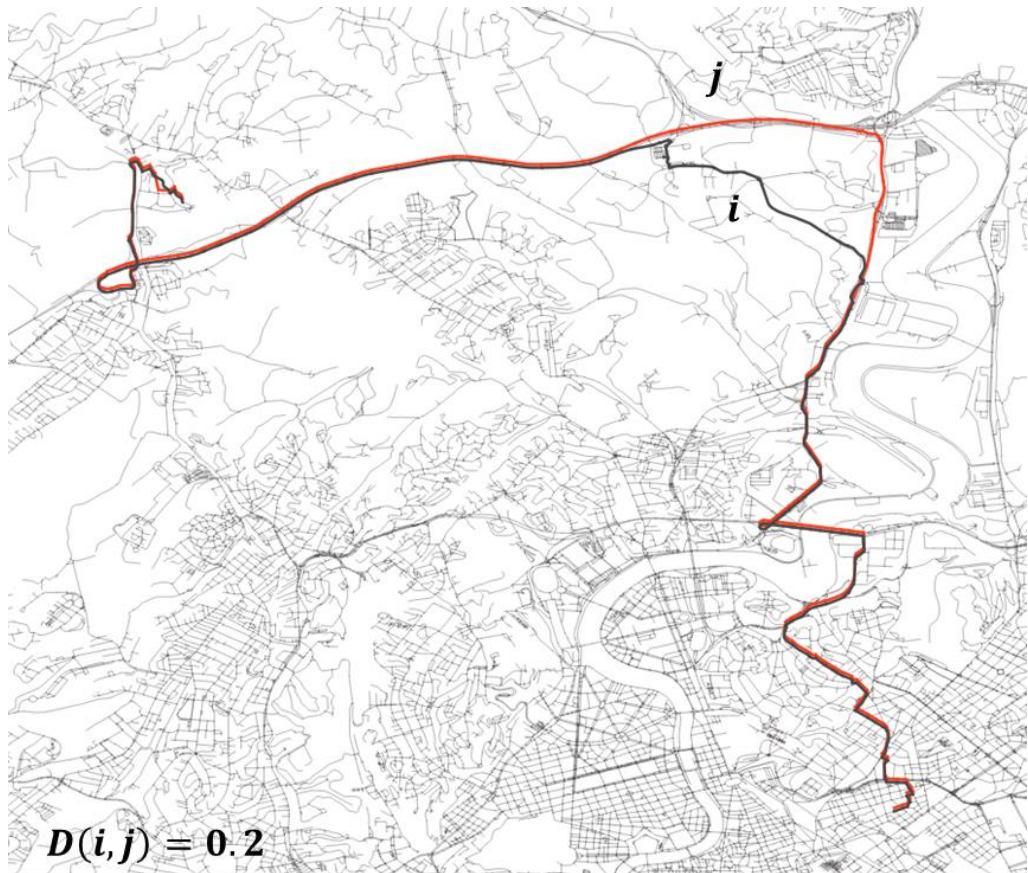


Figura 6 – Esempio di percorsi con basso indice di dissimilarità

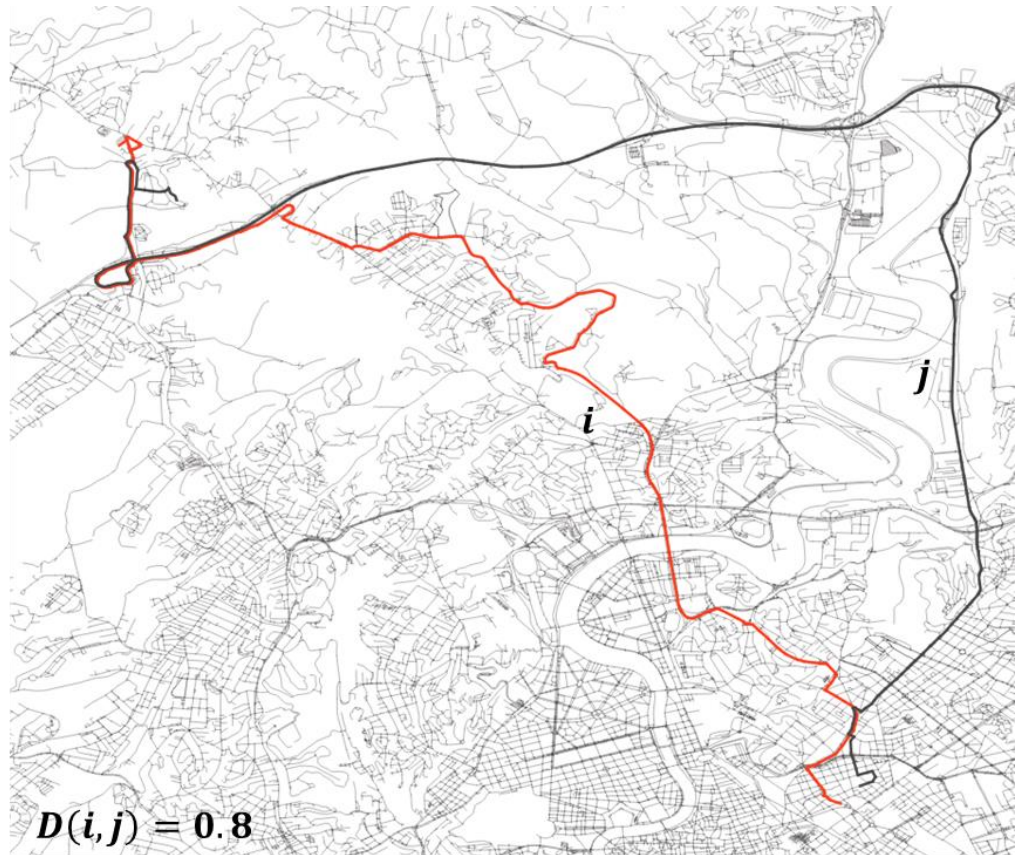


Figura 7 – Esempio di percorsi con alto indice di dissimilarità

In letteratura sono introdotti numerosi indicatori dal maggior grado di complessità, che tengono in considerazione tempi di viaggio, numero e categoria degli archi in comune.

Tuttavia, questo presuppone una conoscenza estesa delle velocità sugli archi della rete in diverse ore del giorno e una attenzione alle variazioni settimanali e stagionali molto complessa da ottenere.

Per questa ragione è stato selezionato un indicatore basato unicamente sulla distanza.

Una volta identificati i cluster finali di scelta, per ciascuno di essi viene selezionato il percorso più rilevante rispetto al grafo stradale e che meglio rappresenta le scelte effettuate degli utenti.

Più nel dettaglio, il processo prevede i seguenti step:

- *Step 0 (Inizializzazione)*: si identifica il set  $P = \{P_i; i=1, 2, \dots, n\}$  degli  $n$  percorsi ricostruiti e si assume che esso coincida con il primo cluster, ovvero  $S=P$  e il numero dei cluster è  $m = 1$ .  $M$  rappresenta il massimo numero di cluster in cui ripartire i percorsi oggetto di analisi.
- *Step 1 (Analisi di dissimilarità)*: per ogni coppia di percorsi  $P_i$  e  $P_j$  di  $S$ , si identificano gli archi stradali condivisi, si calcola la lunghezza cumulata della porzione di percorso condiviso  $L(P_i \cap P_j)$  e la dissimilarità tra i due percorsi tramite l'indice di dissimilarità  $D(i, j)$  dell'equazione 22).
- *Step 2 (Primo split)*: si seleziona la coppia di percorsi più dissimili del cluster  $S$ ,  $h$  e  $h'$ :

$$23) \quad (h, h') = \underset{i,j}{\operatorname{argmax}}\{D(i, j)\} \quad i, j = 1, 2, \dots, n$$

$h$  e  $h'$  divengono ciascuno il primo elemento di un nuovo cluster.

$S_1$  e  $S_2$  sono i due nuovi cluster di percorsi e il numero dei cluster si aggiorna in  $m = m + 1$ .

- *Step 3 (Classificazione dei percorsi)*: per ciascuno dei percorsi di partenza  $P_i, i=1, \dots, n$  si seleziona il cluster  $S_l \in \{S_1, S_2, \dots, S_m\}$  di minor dissimilarità con  $P_i$ :

$$24) \quad l_i = \underset{j_k}{\operatorname{argmin}}\{D(i, j_k)\}$$

$$i = 1, 2, \dots, n \quad j_k = 1, 2, \dots, n_k; \quad k = 1, 2, \dots, m; \quad k \neq i$$

Si assegna  $P_i$  a  $S_l$ .

Se  $m < M$ , si passa allo Step 4, altrimenti allo Step 5.

Se  $P_i$  è completamente dissimile da tutti i cluster identificati, è temporaneamente assegnato al cluster  $S_0$ .

- *Step 4 (Identificazione di un nuovo set)*: per ogni set  $S_k \in \{S_1, S_2, \dots, S_m\}$  si calcola l'indice di dissimilarità  $D(i_k, j_k)$  tra ciascuna coppia di percorsi appartenenti allo stesso set ("dissimilarità interna"), si seleziona il percorso  $P_{q_k}$  per cui l'indice di dissimilarità è massimo per ciascun set  $k$  e, successivamente, il percorso per cui l'indice è massimo tra tutti i set, indicato con  $p$  nella equazione seguente:

$$25) \quad D_{q_k} = \underset{i_k, j_k}{\operatorname{argmax}}\{D(i_k, j_k)\}$$

$$(i_k, j_k) = 1, 2, \dots, n_k \quad k = 1, 2, \dots, m; \quad k \neq i$$

$$p = \underset{k}{\operatorname{argmax}}\{D_{q_k}\}$$

Se l'indice di dissimilarità  $D_p$  è maggiore di un valore soglia  $\eta$ , si definisce un nuovo cluster di percorsi  $S_p$  e  $P_p$  si assegna a  $S_p$ .

Il numero dei cluster si aggiorna a  $m = m + 1$  e si passa allo Step 3, altrimenti  $m = m$  e si passa allo Step 5.

Il valore soglia di  $\eta=0.5$  è stato selezionato in quanto media dei possibili valori dell'indice di dissimilarità interna. Se  $D_k < 0.5$  non c'è necessità di creare un nuovo cluster perché i percorsi appartenenti al set sono già classificati come simili.

Il numero di cluster è variabile e aumenta finché non viene raggiunto un valore di dissimilarità interna sufficiente basso ( $\eta < 0.5$ ), oppure finché il numero di cluster non diventa pari a  $M$ .

- *Step 5 (Selezione dei percorsi rappresentativi)*: per ciascun set di percorsi  $S_k \in \{S_1, S_2, \dots, S_m\}$  viene selezionato il percorso  $P_{r_k}$  che massimizza la funzione seguente:

$$26) \quad r_k = \operatorname{argmax} \frac{\sum_{a \in P_k} w_a f_a l_a}{\sum_{a \in P_k} l_a} \quad k = 1, 2, \dots, M$$

In cui:

- $f_a$  e  $l_a$  sono rispettivamente la frequenza di scelta e la lunghezza dell'arco  $a$ ;
- $w_a$  è il peso dell'arco  $a$ , funzione della classificazione gerarchica dell'arco sulla rete.

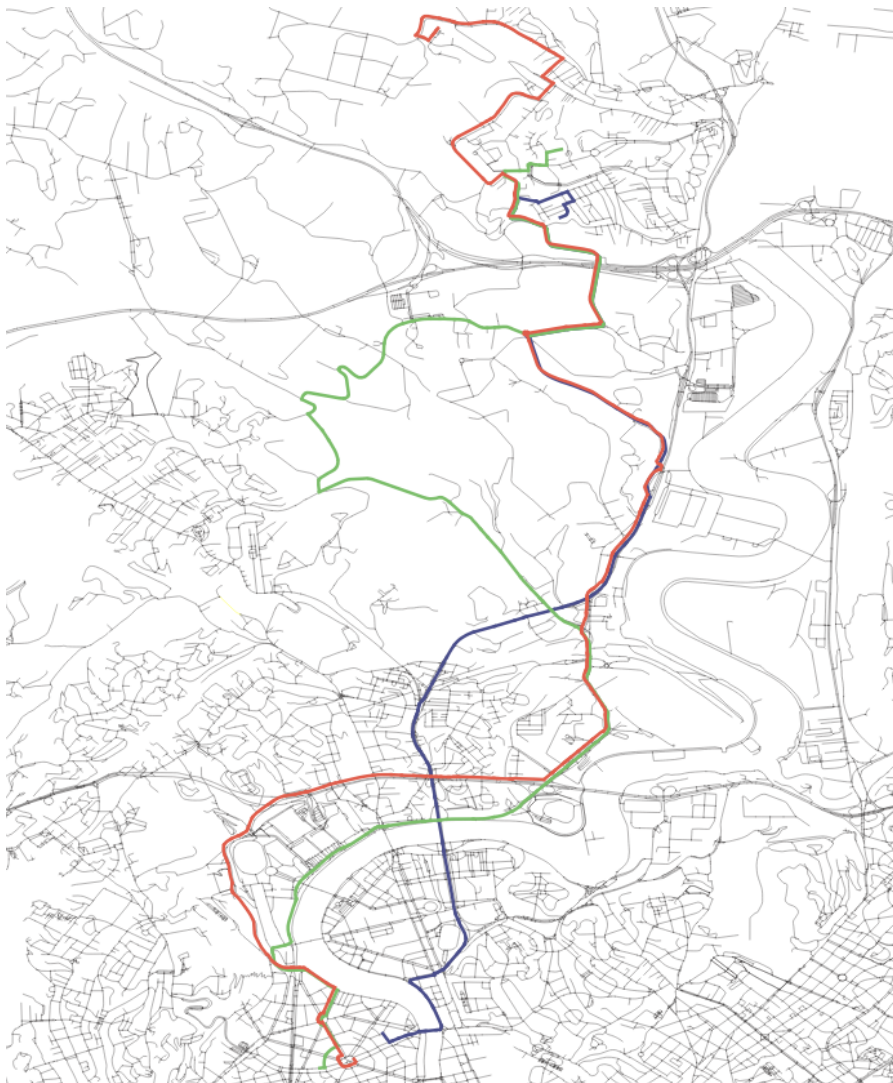
Si seleziona il percorso come rappresentativo del cluster  $S_k$ .

Nella pagina seguente si riporta una rappresentazione esemplificativa dei risultati ottenuti dall'applicazione della routine a una serie di percorsi effettuati dagli utenti.









Come si può notare dalle figure della pagina precedente, i percorsi effettuati dagli utenti sono stati ripartiti in tre cluster, in modo tale da ridurre al minimo il valore di dissimilarità interna, ovvero da costituire dei set di percorsi simili tra loro.

Per ciascuno dei cluster così definiti è stato selezionato un percorso rappresentativo (figura a sinistra).

*Figura 11 – Rappresentazione esemplificativa, percorsi rappresentativi dei cluster individuati*

## 4 Applicazione al caso di studio

La metodologia descritta nel capitolo precedente è stata applicata in primo luogo ad un set di FCD raccolti nel mese di maggio 2010, relativi a circa 600 conducenti, che hanno effettuato 1450 viaggi su 28 coppie O-D nella città di Roma (Fusco, G., Bracci, A., Caligiuri, T., Colombaroni, C., Isaenko, N., 2018).

I risultati ottenuti hanno permesso di condurre una prima analisi del comportamento di viaggio degli utenti stradali: verificarne la tendenza a selezionare il percorso più breve, indagare le variazioni nelle scelte del percorso e nella scelta dell'orario di partenza da un giorno all'altro, etc. Le elaborazioni hanno confermato la propensione degli utenti a seguire il "percorso abituale" per arrivare a destinazione, sebbene il 12% degli utenti abbia cambiato in favore di un itinerario alternativo.

Tuttavia, le osservazioni non hanno rivelato l'esistenza di alcuna correlazione sistematica tra il cambio di percorso / dell'orario di partenza e il tempo di viaggio sperimentato il giorno precedente.

È stata, infatti, evidenziata la predominanza di una componente casuale che nasconde ogni correlazione attesa tra il comportamento degli utenti e i tempi di viaggio sperimentati.

Ipotizzando che i casi in cui gli utenti stradali scelgono un percorso alternativo rispetto all'itinerario abituale siano eventi rari e correlati al verificarsi di circostanze eccezionali, la metodologia è stata applicata a un secondo set di FCD relativi ad un periodo di analisi più lungo.

Di seguito sono descritte nel dettaglio le elaborazioni eseguite e i principali risultati ottenuti.

## 4.1 Analisi preliminare dei dati

Questo secondo database di FCD – raccolti nei primi sei mesi dell’anno 2015 e già associati agli archi del grafo stradale Teleatlas 2013 – è stato elaborato in modo da definire un set di “viaggi di interesse”, selezionando:

- i soli dati relativi a utenti stradali (categoria “auto privata”);
- una volta ricostruiti i viaggi effettuati da ciascun utente (ovvero la sequenza di record con stesso “ID veicolo”, per ciascun giorno oggetto di analisi) - i soli viaggi di lunghezza superiore ai 2 km (attraverso il campo “odo”) in modo da escludere spostamenti tra origini e destinazioni molto vicine tra loro;
- i viaggi effettuati durante l’ora di punta mattutina (dalle 7:00 alle 9:00), ipotizzando così di ottenere spostamenti prevalentemente sistematici, per cui sia possibile osservare il comportamento di viaggio dell’utente su più giorni consecutivi.

A valle di tali elaborazioni, è stato ottenuto un totale di 145883 viaggi, di una durata media di 30 minuti circa.

*Tabella 1 - Dettaglio dei “viaggi di interesse” ottenuti per ciascun mese*

	Gen.	Feb.	Mar.*	Apr.*	Mag.	Giu.
<b>Viaggi totali</b>	19014	19060	4942	9313	47567	45987
<b>Utenti</b>	4328	5360	3042	5193	10406	10897
<b>Avg. viaggi per utente</b>	4	4	2	2	5	4
<b>Avg. durata viaggio per utente</b>	30 min.	31 min.	30 min.	30 min.	32 min.	31 min.

*\*Per marzo e aprile 2015, i dati non sono disponibili per ogni giorno del mese.*

A questo punto, sono stati stimati l’orario medio di partenza e la durata media – in termini di tempi di viaggio - per ciascun utente.

Questi valori sono stati utilizzati per una analisi preliminare, svincolata dall’informazione del grafo stradale, e confrontati con i valori associati a ciascun viaggio effettuato dagli utenti per stimare:

- l'aumento o la riduzione del tempo di percorrenza sperimentata da ciascun utente nei diversi viaggi effettuati rispetto al tempo di viaggio medio;
- l'anticipo o il ritardo rispetto all'orario medio di partenza di ciascun utente.

In questo modo, sono stati analizzati:

- i casi in cui l'utente anticipasse o ritardasse l'orario medio di partenza a seguito di un aumento o una riduzione del tempo di percorrenza sperimentato durante l'ultimo viaggio effettuato (caso 1);
- i casi in cui l'utente anticipasse o ritardasse l'orario medio di partenza a seguito di un aumento o una riduzione del tempo di percorrenza sperimentato il giorno precedente (caso 2).

Le due casistiche sono state analizzate separatamente, poiché è stato ipotizzato che un ritardo sperimentato durante il viaggio effettuato il giorno precedente (caso 2) abbia un impatto diverso sulle scelte di mobilità di un utente, rispetto ad un ritardo sperimentato durante l'ultimo viaggio effettuato (ad esempio una settimana prima).

#### 4.1 Evidenze dall'analisi preliminare dei dati

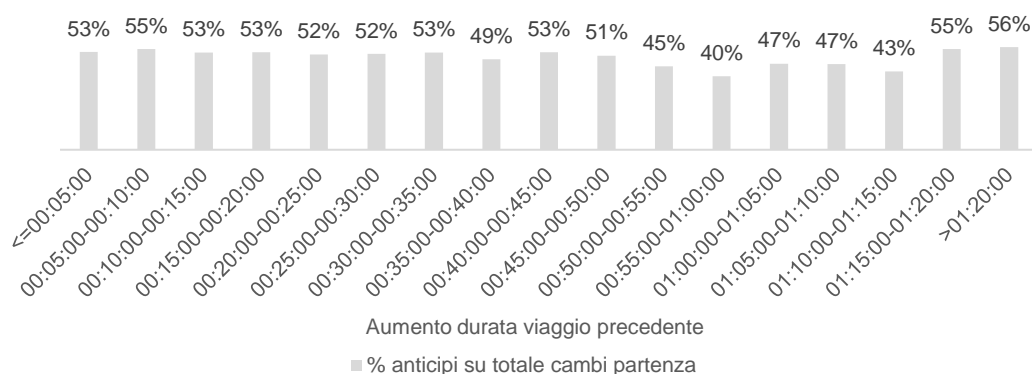
Per i mesi di gennaio-giugno 2015, nel caso 1, sono stati registrati 29015 casi in cui - a seguito di un aumento della durata dell'ultimo viaggio effettuato - l'utente ha anticipato la partenza e 25527 casi in cui l'ha ritardata. Sono poi stati registrati 39329 casi in cui, a fronte di una riduzione della durata dell'ultimo viaggio effettuato - l'utente ha anticipato la partenza e 36455 casi in cui l'ha ritardata.

Rispetto agli aumenti della durata del tempo di viaggio è emerso quanto riportato nella tabella e nel grafico seguenti: non si evidenzia un aumento significativo della percentuale di utenti che tende ad anticipare la

partenza del viaggio successivo all'aumentare del tempo di viaggio sperimentato nel viaggio precedente, ovvero non è immediatamente identificabile la c.d. "banda di indifferenza" rispetto a peggioramenti del tempo di percorrenza, superata la quale, gli utenti sono spinti ad anticipare l'orario di partenza.

**Tabella 2 – Anticipi e ritardi dell'orario di partenza a seguito di un aumento della durata sperimentato durante l'ultimo viaggio effettuato**

Aumento durata ultimo viaggio	Anticipi partenza viaggio successivo	Ritardi partenza viaggio successivo
<b>Totale</b>	<b>29015</b>	<b>25527</b>
<b>%</b>	<b>53%</b>	<b>47%</b>

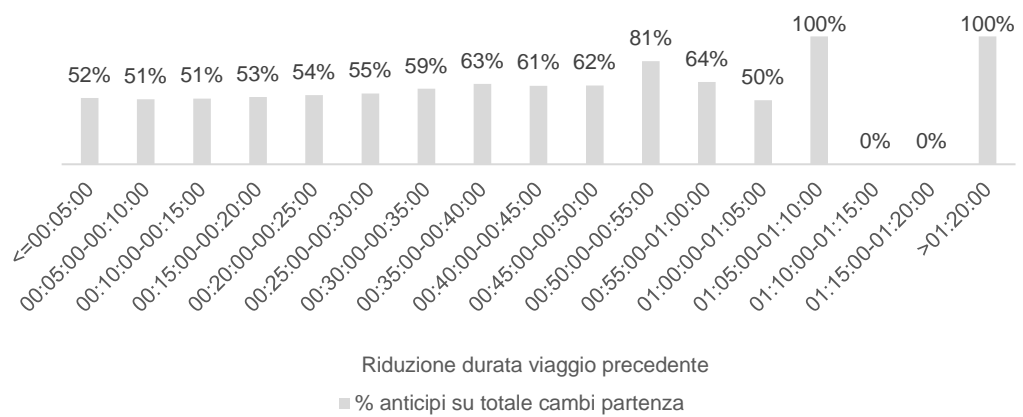


**Figura 12 – Percentuale di anticipi sul totale dei cambi di orario di partenza in funzione dell'aumento della durata sperimentato durante l'ultimo viaggio effettuato**

La stessa analisi è stata effettuata per le riduzioni della durata del tempo di viaggio.

**Tabella 3 – Anticipi e ritardi dell'orario di partenza a seguito di una riduzione della durata sperimentata durante l'ultimo viaggio effettuato**

Riduzione durata ultimo viaggio	Anticipi partenza viaggio successivo	Ritardi partenza viaggio successivo
<b>Totale</b>	<b>39329</b>	<b>36455</b>
<b>%</b>	<b>52%</b>	<b>48%</b>



*Figura 13 – Percentuale di anticipi sul totale dei cambi di orario di partenza in funzione della riduzione della durata sperimentata durante l'ultimo viaggio effettuato*

Per quanto riguarda l'entità degli incrementi / delle riduzioni di durata dei tempi di viaggio, in oltre il 50% dei casi, si tratta di variazioni entro i 5 minuti rispetto alla durata media del viaggio di ciascun utente.

Lo stesso vale per l'entità delle variazioni apportate dagli utenti all'orario di partenza: in più del 35% dei casi nella si tratta di variazioni contenute (entro i 5 minuti).

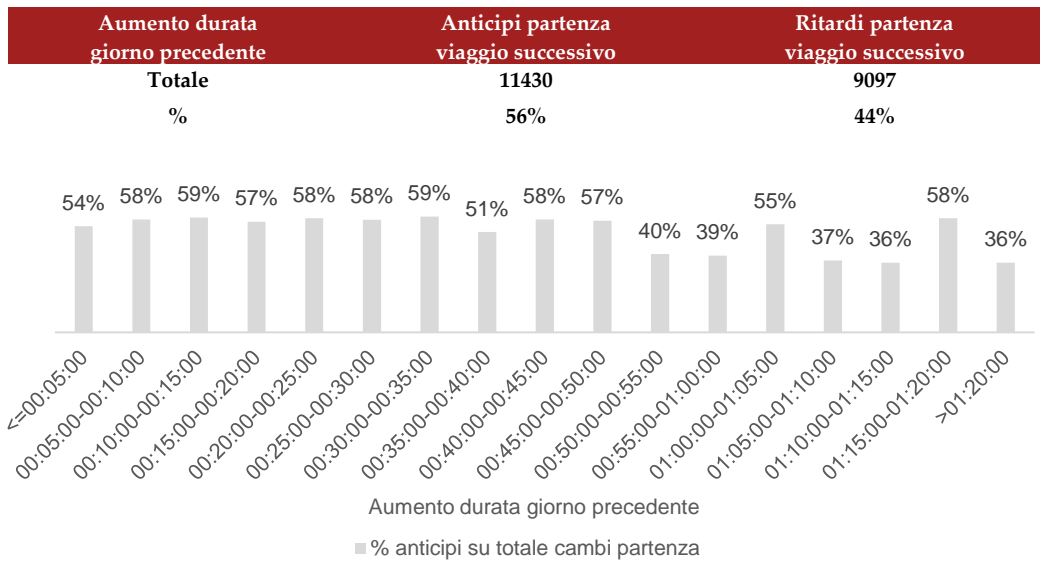
La stessa analisi è stata condotta considerando soltanto i casi in cui l'utente anticipasse o ritardasse l'orario medio di partenza a seguito di un aumento o una riduzione del tempo di percorrenza sperimentato durante il viaggio effettuato il giorno precedente (caso 2), ovvero sono stati analizzati solo i viaggi compiuti in giorni consecutivi dagli utenti.

Per i mesi di gennaio-giugno 2015, sono stati registrati 11430 casi in cui - a seguito di un aumento della durata dell'ultimo viaggio effettuato - l'utente ha anticipato la partenza e 9097 casi in cui l'ha ritardata.

Sono poi stati registrati 14107 casi in cui, a fronte di una riduzione della durata dell'ultimo viaggio effettuato - l'utente ha anticipato la partenza e 13714 casi in cui l'ha ritardata.

Rispetto agli aumenti della durata del tempo di viaggio è emerso quanto riportato di seguito.

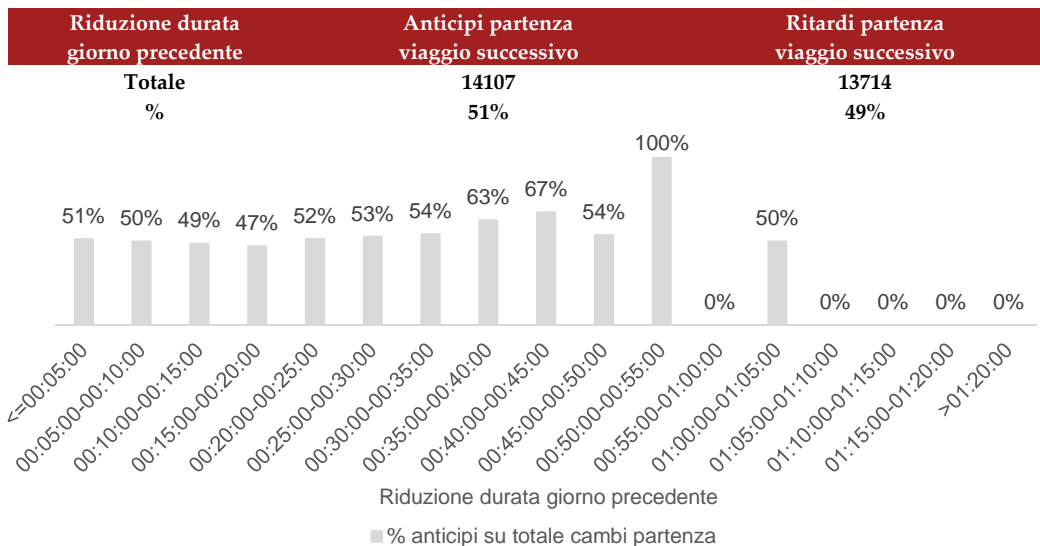
**Tabella 4 – Anticipi e ritardi dell'orario di partenza a seguito di un aumento della durata sperimentato il giorno precedente**



**Figura 14 - Percentuale di anticipi sul totale dei cambi di orario di partenza in funzione dell'aumento della durata sperimentato il giorno precedente**

La stessa analisi è stata effettuata per le riduzioni della durata del tempo di viaggio.

**Tabella 5 – Anticipi e ritardi dell'orario di partenza a seguito di una riduzione della durata sperimentata il giorno precedente**



**Figura 15 - Percentuale di anticipi sul totale dei cambi di orario di partenza in funzione della riduzione della durata sperimentata il giorno precedente**

Per quanto riguarda l'entità degli incrementi / delle riduzioni di durata dei tempi di viaggio, in oltre il 55% dei casi, si tratta di variazioni entro i 5 minuti rispetto alla durata media del viaggio di ciascun utente.

Lo stesso vale per le variazioni apportate all'orario di partenza: in più del 40% dei casi si tratta di variazioni contenute apportate all'orario medio di partenza di ciascun utente (entro i 5 minuti).



Si riporta di seguito uno specchio riepilogativo delle principali evidenze ottenute rispetto ai sei mesi analizzati e ai due casi di analisi.

*Tabella 6 – Caso 1 (viaggi consecutivi)*

Aumento/riduzione durata ultimo viaggio	Anticipo/ritardo partenza viaggio successivo	Gennaio		Febbraio		Marzo		Aprile		Maggio		Giugno		Totale	% su Totale
		Num. casi	% su totale	Num. casi	% su totale	Num. casi	% su totale	Num. casi	% su totale	Num. casi	% su totale	Num. casi	% su totale		
Aumento	Anticipo	3148	54%	3474	52%	981	55%	1494	51%	11035	55%	8883	52%	29015	53%
Aumento	Ritardo	2648	46%	3268	48%	796	45%	1423	49%	9090	45%	8302	48%	25527	47%
Riduzione	Anticipo	4384	49%	4814	49%	1411	54%	2140	50%	12974	55%	13606	51%	39329	52%
Riduzione	Ritardo	4506	51%	5067	51%	1224	46%	2118	50%	10635	45%	12905	49%	36455	48%
<b>Totale</b>		<b>14686</b>		<b>16623</b>		<b>4412</b>		<b>7175</b>		<b>43734</b>		<b>43696</b>		<b>130326</b>	

*Tabella 7 – Caso 2 (viaggi effettuati in giorni consecutivi)*

Aumento/riduzione durata giorno t-1	Anticipo/ritardo partenza giorno t	Gennaio		Febbraio		Marzo		Aprile		Maggio		Giugno		Totale	% su Totale
		Num. casi	% su totale	Num. casi	% su totale	Num. casi	% su totale	Num. casi	% su totale	Num. casi	% su totale	Num. casi	% su totale		
Aumento	Anticipo	1628	56%	1236	52%	426	61%	410	55%	4603	57%	3127	55%	11430	56%
Aumento	Ritardo	1276	44%	1125	48%	269	39%	336	45%	3533	43%	2558	45%	9097	44%
Riduzione	Anticipo	2040	48%	1567	47%	517	57%	539	49%	4823	54%	4621	50%	14107	51%
Riduzione	Ritardo	2180	52%	1802	53%	395	43%	562	51%	4107	46%	4668	50%	13714	49%
<b>Totale</b>		<b>7124</b>		<b>5730</b>		<b>1607</b>		<b>1847</b>		<b>17066</b>		<b>14974</b>		<b>48348</b>	

Dalle tabelle si può notare una lieve tendenza degli utenti ad anticipare l'orario di partenza sia nel caso in cui abbiano sperimentato un aumento del tempo di viaggio il giorno precedente, sia nel caso in cui abbiano sperimentato una riduzione.

Si sottolinea, tuttavia, come tale analisi sia svincolata dal percorso seguito dall'utente: al fine di analizzare le variazioni apportate dagli utenti all'orario di partenza in relazione ad uno stesso itinerario, si rimanda ai paragrafi successivi.

## 4.2 Analisi relativa al grafo stradale

Come già menzionato, l'analisi preliminare descritta nel paragrafo precedente è completamente svincolata dal percorso seguito dall'utente.

Al fine di analizzare le variazioni apportate dagli utenti all'orario di partenza in relazione ad uno stesso itinerario percorso più volte nell'arco dei mesi oggetto di analisi è stata applicata la metodologia descritta nel capitolo precedente.

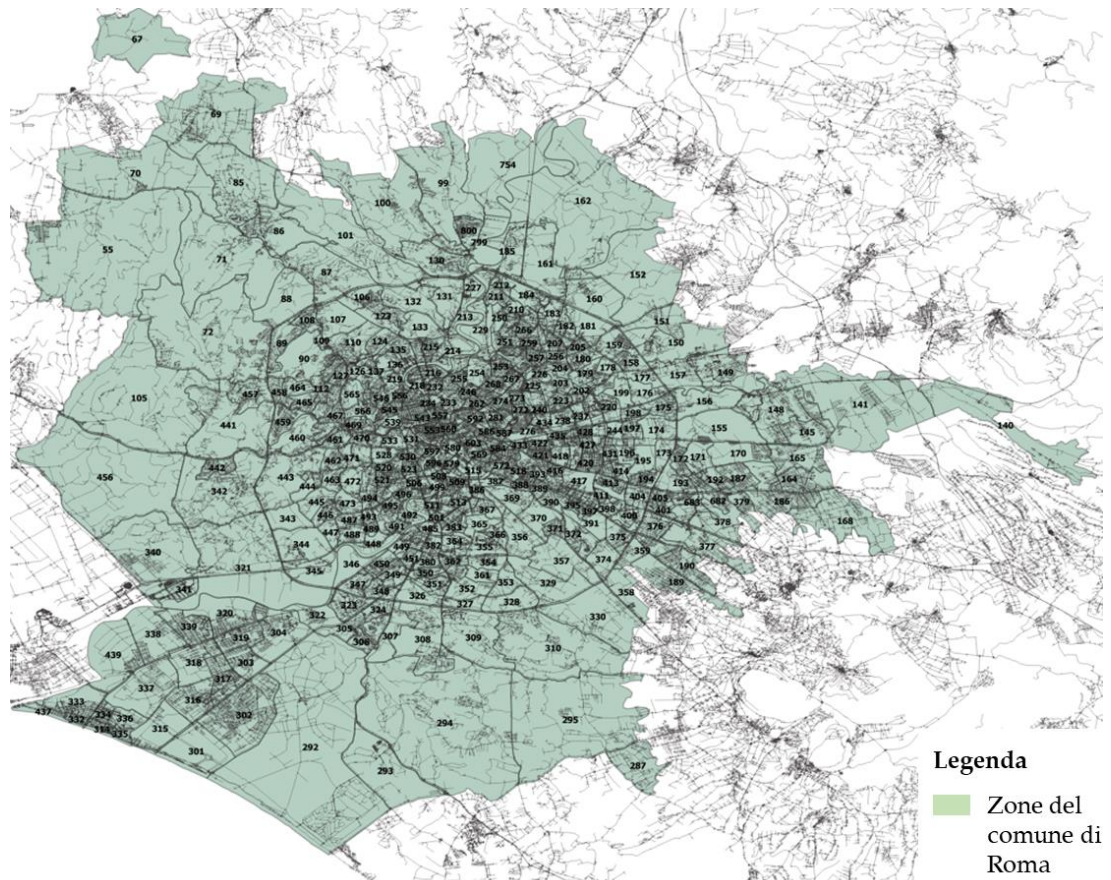
Anche per questa analisi, è stato utilizzato il database dei FCD raccolti nei mesi di gennaio-giugno 2015, elaborato in modo da definire un set di "viaggi di interesse", secondo quanto specificato nel paragrafo 4.1.

A ciascun FCD è stata, in primo luogo, associata la zona di appartenenza e sono stati eliminati:

- i viaggi con origine e destinazione nella stessa zona, poiché di scarso interesse ai fini dell'analisi, in modo da escludere spostamenti tra origini e destinazioni molto vicine tra loro;
- i viaggi con origine o destinazione (o entrambe) nelle zone esterne rispetto al comune di Roma, in modo da limitare l'analisi agli spostamenti effettuati all'interno della città.

L'operazione di associazione di ciascun FCD alla zona di appartenenza è stata effettuata in QGIS attraverso un tool di unione degli elementi sulla base di attributi di posizione di questi ultimi.

La zonizzazione utilizzata per le elaborazioni è rappresentata nella figura seguente: il comune di Roma - all'interno dell'omonima città metropolitana - è articolato in 463 zone.



*Figura 16 – Zonizzazione adottata per l’area di studio (comune di Roma)*

L’identificazione delle coppie O-D oggetto di analisi è stata effettuata calcolando, per ciascuna coppia zona-zona, il numero di viaggi dalla durata minima di 20 minuti e dalla lunghezza minima di 6 km.

Le coppie O-D con almeno 30 viaggi che rispettavano tali vincoli sono state selezionate per l’analisi.

Il risultato di tale selezione sono le 201 coppie O-D riportate nella tabella successiva.

Alcune delle zone selezionate, sono analizzate in origine, altre in destinazione, altre ancora in origine e in destinazione.

ID O-D	Coppia O-D	Numero viaggi analizzati	ID O-D	Coppia O-D	Numero viaggi analizzati
1	205-272	62	102	322-514	44
2	458-505	71	103	302-333	48
3	196-358	76	104	302-327	42
4	306-495	67	105	339-325	40
5	130-215	95	106	325-382	40
6	518-364	67	107	332-304	57
7	461-550	82	108	430-359	48
8	427-446	75	109	123-282	40
9	408-509	77	110	506-269	40
10	339-498	62	111	537-136	41
11	339-334	94	112	194-175	43
12	319-509	86	113	519-217	40
13	267-547	63	114	175-395	48
14	186-190	77	115	325-484	99
15	159-434	81	116	376-309	49
16	130-235	67	117	127-380	36
17	218-208	62	118	171-238	40
18	87-263	55	119	176-280	40
19	99-216	57	120	184-577	37
20	148-184	90	121	198-410	41
21	149-225	56	122	217-351	38
22	180-270	64	123	220-400	38
23	193-490	56	124	268-109	45
24	237-362	58	125	294-451	42
25	302-382	58	126	302-332	74
26	320-345	56	127	302-347	37
27	347-505	68	128	302-380	37
28	349-552	59	129	320-483	38
29	367-538	60	130	324-427	41
30	377-330	71	131	325-366	56
31	460-449	57	132	325-519	38
32	467-364	56	133	347-441	39
33	521-229	55	134	359-304	37
34	561-382	55	135	442-495	43
35	563-358	59	136	495-347	38
36	573-174	73	137	246-483	34
37	420-366	56	138	243-185	43
38	270-131	50	139	100-210	35
39	509-246	52	140	150-202	36
40	253-327	54	141	334-345	36
41	196-344	52	142	130-308	35
42	494-602	59	143	325-513	42
43	159-542	50	144	323-504	44
44	186-194	57	145	303-367	36
45	193-352	56	146	350-569	38
46	208-547	50	147	413-387	42
47	239-406	53	148	367-345	48
48	294-351	62	149	361-334	34
49	304-354	61	150	351-124	35
50	347-547	57	151	375-347	36
51	392-345	50	152	487-449	35

ID O-D	Coppia O-D	Numero viaggi analizzati	ID O-D	Coppia O-D	Numero viaggi analizzati
52	392-383	54	153	463-362	35
53	421-449	56	154	459-217	36
54	430-280	55	155	445-484	62
55	431-377	75	156	432-341	37
56	493-534	64	157	110-243	34
57	171-213	45	158	155-220	33
58	434-502	47	159	189-380	35
59	383-256	49	160	260-287	43
60	260-263	50	161	272-164	32
61	196-278	49	162	302-366	34
62	196-231	47	163	304-196	33
63	159-518	46	164	308-389	34
64	149-192	47	165	319-425	34
65	164-420	58	166	336-383	34
66	376-157	49	167	349-378	34
67	521-483	89	168	375-155	87
68	306-383	52	169	377-404	68
69	494-483	50	170	431-231	34
70	325-449	72	171	443-501	34
71	314-449	49	172	457-524	34
72	322-483	63	173	502-216	34
73	351-448	57	174	504-380	44
74	555-157	45	175	521-486	33
75	543-158	49	176	522-483	36
76	432-220	64	177	89-349	32
77	206-593	45	178	158-189	32
78	184-149	52	179	183-216	32
79	377-403	48	180	192-603	32
80	379-373	45	181	204-214	39
81	220-233	43	182	205-407	31
82	432-213	42	183	217-278	46
83	683-426	59	184	223-565	31
84	682-271	44	185	294-398	31
85	518-177	42	186	302-381	32
86	427-229	52	187	319-157	34
87	465-273	43	188	324-484	61
88	130-90	43	189	332-336	80
89	446-155	48	190	338-322	46
90	503-376	42	191	347-508	33
91	428-486	44	192	492-355	107
92	317-483	44	193	503-282	32
93	347-345	50	194	533-257	32
94	462-352	43	195	574-268	42
95	186-400	60	196	602-362	33
96	379-214	46	197	148-591	30
97	310-592	40	198	221-376	31
98	148-125	42	199	336-127	30
99	322-185	40	200	346-547	30
100	420-151	48	201	522-397	30
101	457-235	44			



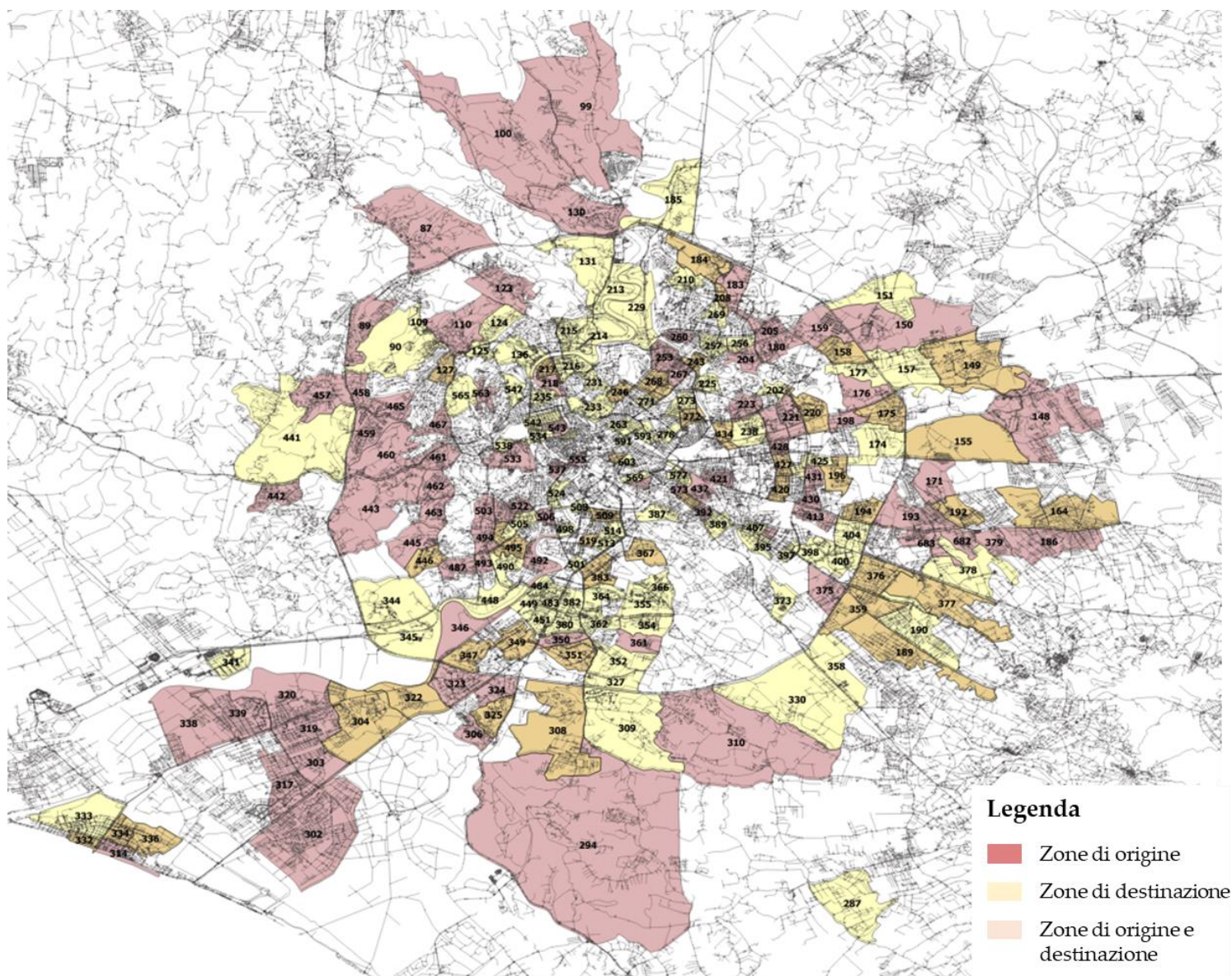


Figura 17 – Zone oggetto di analisi

Una volta selezionate le 201 coppie O-D zona-zona, sono stati ottenuti un totale di 9882 viaggi effettuati da 339 utenti.

*Tabella 8 – Sintesi dei risultati ottenuti dal processo di selezione delle O-D oggetto di analisi*

<b>Totale coppie OD</b>		
201		
<b>Totale utenti analizzati</b>	<b>Viaggi medi per utente</b>	
339	29	
<b>Viaggi totali analizzati</b>	<b>Tempo di viaggio medio (min)</b>	<b>Lunghezza media (km)</b>
9882	36	16

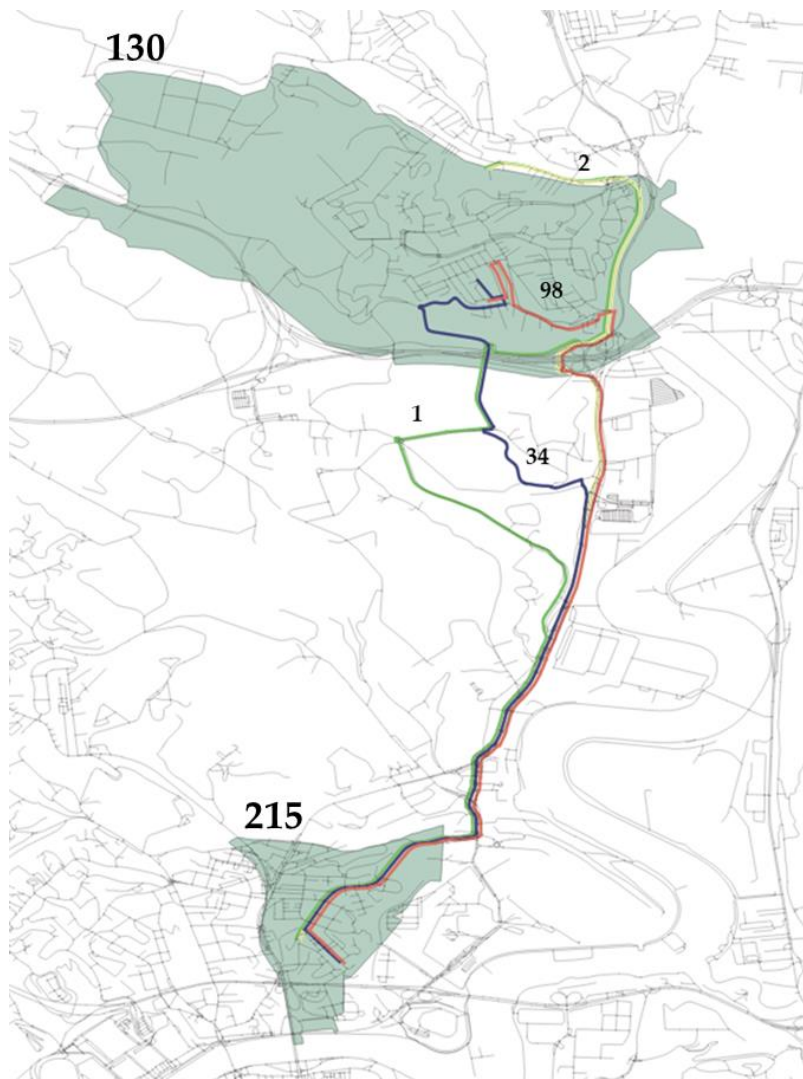
A questo punto, ai viaggi compiuti dagli utenti su ciascuna relazione O-D, è stata applicata la routine di ricostruzione del percorso.

Come già approfondito nel capitolo precedente, ogni viaggio è caratterizzato da una serie di posizioni sparse assegnate agli archi del grafo stradale e caratterizzate dallo stesso "ID veicolo".

La routine implementata permette di ricostruire l'intera sequenza di archi attraversati dall'utente, utilizzando un algoritmo di percorso minimo tra ciascuna coppia di posizioni registrate.

Una volta ricostruiti i viaggi effettuati dagli utenti su ciascuna coppia O-D zona-zona, attraverso l'algoritmo di dissimilarità, è stato individuato il set di percorsi che possono essere percepiti come alternativi dagli utenti stradali in viaggio su ciascuna relazione.

Si riporta di seguito, a titolo esemplificativo, il risultato ottenuto per la coppia O-D 130-215.



*Figura 18 – Risultati dell’algoritmo di dissimilarità per la coppia 130-215*

I 95 viaggi con origine nella zona 130 e destinazione nella zona 215, sono stati classificati in quattro cluster, per ciascuno dei quali è stato identificato un percorso rappresentativo. I quattro percorsi rappresentativi sono rappresentati nella figura di sinistra.

Il 55% dei viaggi è stato assegnato al cluster del percorso rappresentativo 2 (in giallo), un 17% al cluster del percorso rappresentativo 1 (in verde), un 17% al cluster del percorso rappresentativo 34 (in blu) e il restante 11% al cluster del percorso rappresentativo 98 (in rosso).

I valori degli indici di dissimilarità per la coppia O-D oggetto di analisi, evidenziano i risultati della procedura di clustering.

Cluster	Rappr.	Lungh. [km]	Num. Scelte	Diss. Int. Avg.	Diss. Est. avg
1	1 (verde)	9.77	16	0.31	0.62
2	2 (giallo)	8.20	52	0.08	0.58
3	98 (rosso)	7.95	11	0.20	0.47
4	34 (blu)	7.82	16	0.29	0.59



Lo stesso procedimento è stato seguito per tutte le coppie O-D zona-zona selezionate e sono stati ottenuti i valori riportati nella tabella seguente:

- nella prima colonna è riportato il valore medio di “dissimilarità interna” ottenuto: il valore è piuttosto basso e attesta che percorsi simili sono stati efficientemente assegnati allo stesso cluster;
- viceversa, il valore medio di “dissimilarità esterna” del 50%, indica che un percorso appartenente ad un cluster risulta essere significativamente diverso da percorsi assegnati ad altri cluster;
- nella terza colonna della tabella è riportata la dissimilarità media tra i percorsi generici di un cluster e il percorso rappresentativo dello stesso cluster. Un valore basso indica che il percorso rappresentativo ben caratterizza gli itinerari appartenenti al proprio cluster;
- nella quarta colonna, infine, è riportata la dissimilarità media tra percorsi rappresentativi. Il valore medio del 50% indica che i percorsi rappresentativi di ciascun cluster sono significativamente differenti gli uni dagli altri.

*Tabella 9 – Sintesi dei risultati ottenuti dal processo di ricostruzione e clusterizzazione dei “viaggi di interesse”*

Diss. interna avg	Diss. esterna avg	Dissimilarità media tra perc. generico e rappresentativo	Dissimilarità media tra percorsi rappresentativi
16%	50%	14%	50%

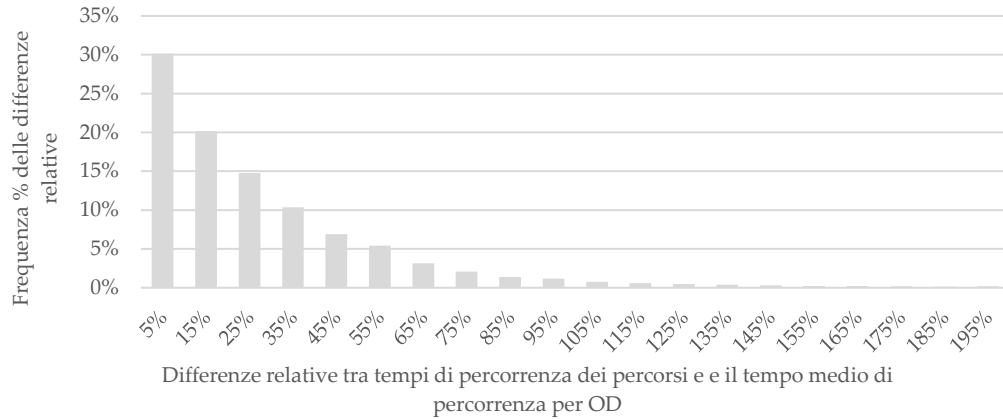
#### 4.2.1 Analisi della variazione dei tempi di viaggio

Il comportamento di viaggio degli utenti è strettamente correlato con i tempi di viaggio: il I principio di Wardrop stabilisce, infatti, che in condizioni di equilibrio, il costo di un cammino effettivamente utilizzato su una coppia O-D è minore o uguale al costo di qualsiasi altro cammino che congiunge la stessa coppia (Wardrop, J. G., 1952).

Tuttavia, i risultati ottenuti attraverso la ricostruzione e la suddivisione dei percorsi in cluster evidenziano come gli utenti possano selezionare anche alternative di viaggio dal tempo di percorrenza diverso dal minimo.

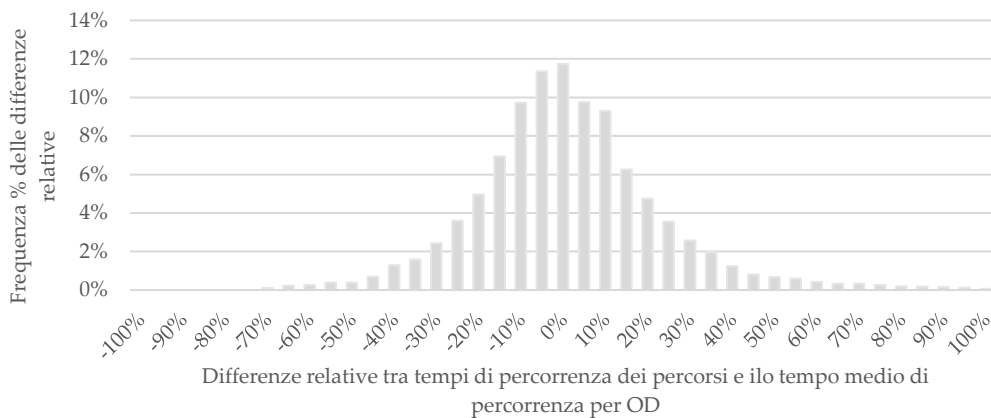
Per tale ragione, al fine valutare la frequenza con cui vengono selezionati i percorsi più lunghi da parte degli utenti, è stata calcolata la differenza relativa tra ciascun tempo di viaggio e il tempo di viaggio minimo per O-D. Poiché le analisi sono condotte relativamente ad un periodo di sei mesi - in cui i tempi di viaggio possono variare sensibilmente da un giorno all'altro - per escludere tempi di percorrenza particolarmente ridotti, il tempo minimo di viaggio per coppia O-D è stato definito come il 20° percentile dei tempi di viaggio osservati per ciascuna O-D.

La distribuzione di frequenza delle differenze relative tra tempi di percorrenza osservati e il tempo minimo di percorrenza per coppia O-D ha evidenziato come il 75% dei percorsi differisca dal minimo di meno del 35%.



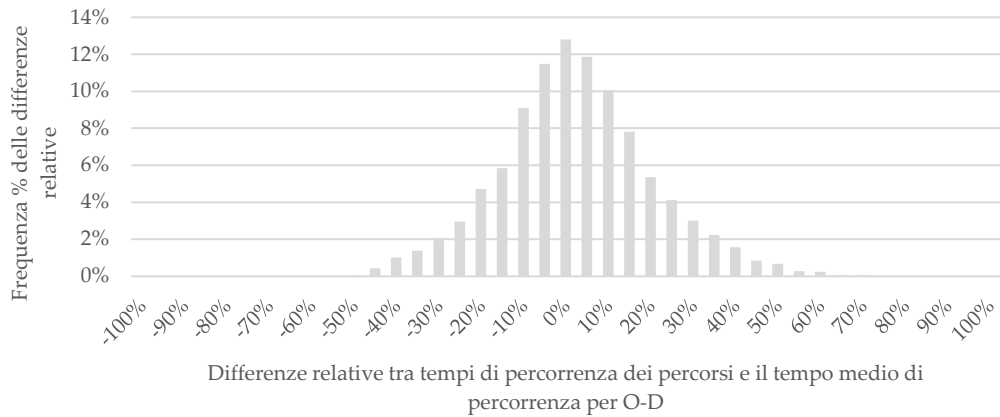
**Figura 19 – Frequenza delle differenze relative tra tempi di percorrenza e tempo minimo di percorrenza per coppia O-D**

D’altro canto, calcolando la distribuzione di frequenza della differenza relativa tra i tempi di viaggio e il tempo medio di viaggio per O-D, si può osservare come più dell’80% dei percorsi differisca dal valore medio di meno del 25% e solo il 5% di oltre il 50% (figura seguente).



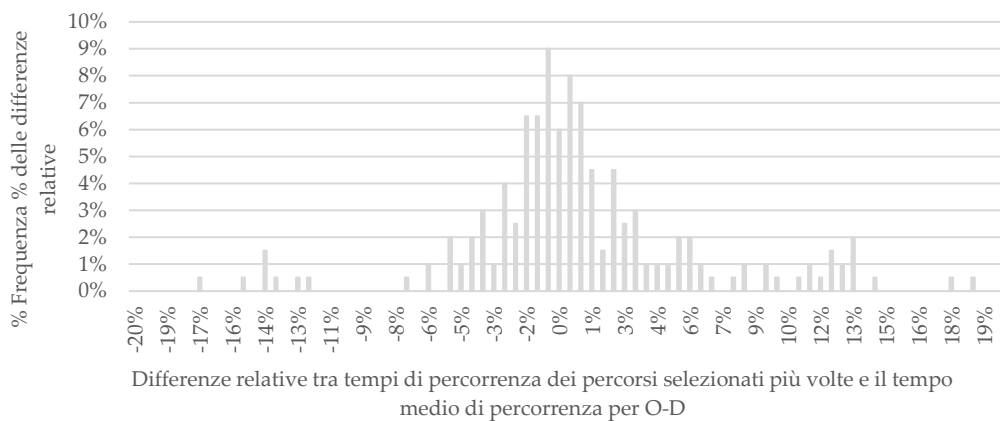
**Figura 20 - Frequenza delle differenze relative tra tempi di percorrenza e tempo medio di percorrenza per coppia O-D**

Trascurando il 5% dei percorsi il cui tempo di viaggio differisce dal valore medio di oltre il 50%, si ottiene il risultato rappresentato nella figura seguente.



**Figura 21 - Frequenza delle differenze relative tra tempi di percorrenza e tempo medio di percorrenza per coppia O-D, trascurando il 5% dei percorsi**

Calcolando, infine, la frequenza delle differenze relative tra il tempo di viaggio medio del cluster selezionato più volte dagli utenti in viaggio su una coppia O-D e il tempo di viaggio medio sulla stessa coppia, è stato ottenuto quanto rappresentato nella figura seguente: quasi il 90% dei tempi di percorrenza differisce dal percorso medio di meno del 10%.



**Figura 22 - Frequenza delle differenze relative tra tempi di percorrenza dei percorsi maggiormente selezionati e i tempi medi di percorrenza per coppia O-D**

È stata, inoltre, indagata la variabilità dei tempi di viaggio all'interno di ciascuna coppia O-D e, per ogni coppia, all'interno di ciascun cluster. Per ricavare statistiche sulla variabilità complessiva del tempo di viaggio all'interno di ciascuna coppia O-D è stata introdotta la deviazione standard relativa media (cioè il coefficiente di variazione medio) come rapporto tra la

deviazione standard del tempo di viaggio di tutti i viaggi della coppia O-D e il tempo medio di percorrenza per coppia.

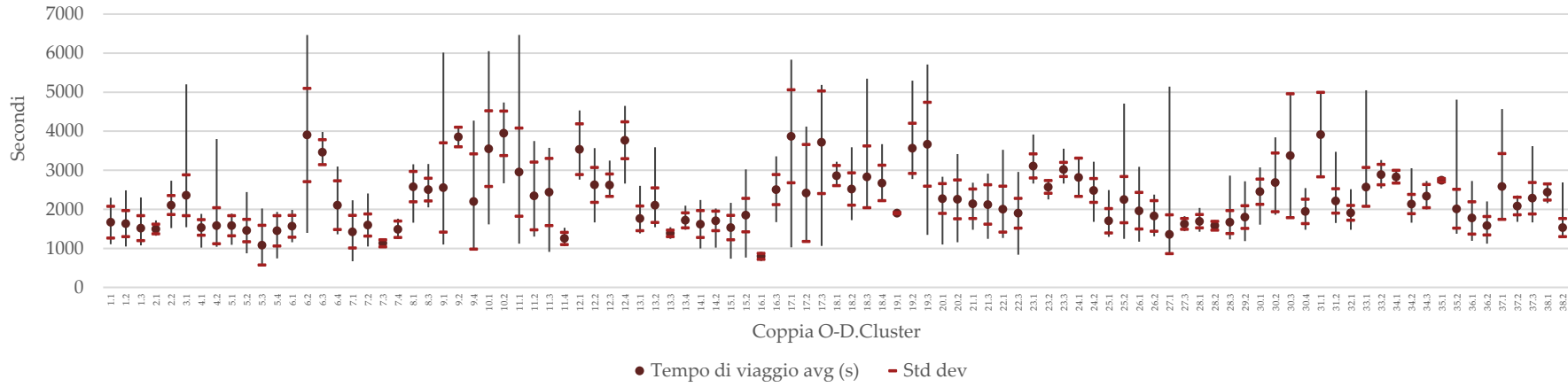
È stato ottenuto un valore medio del 24% su tutte le coppie O-D.

Lo stesso calcolo è stato effettuato per i cluster: è stato ottenuto un valore medio del 19% su tutte le coppie O-D.

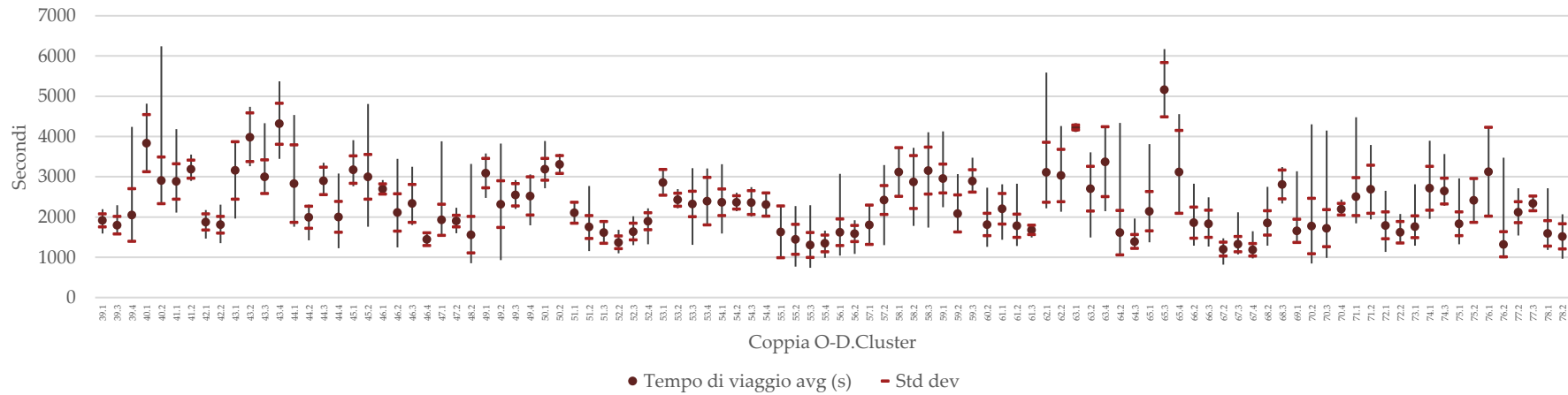
Per quanto riguarda, invece, la valutazione dell'approssimazione introdotta nella rappresentazione statistica dei tempi di viaggio con l'introduzione dei cluster, sono stati analizzati, per ciascun cluster di ciascuna coppia O-D, il valore massimo, minimo e medio dei tempi di percorrenza, nonché la loro deviazione standard.

Da tale analisi sono stati esclusi i cluster contenenti un solo percorso ed è stato ottenuto quanto segue.

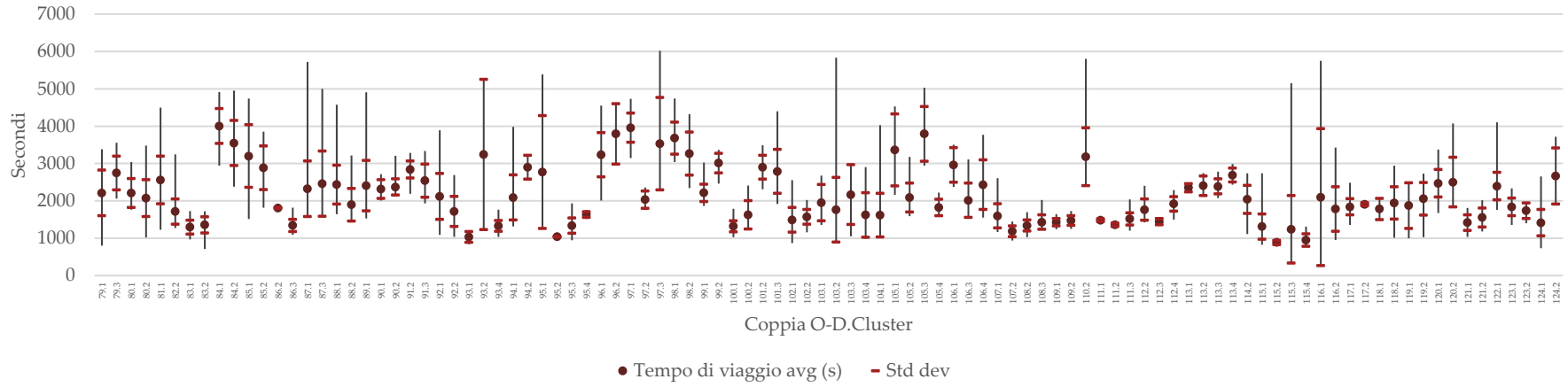
Variazione del tempo di viaggio per O-D e cluster



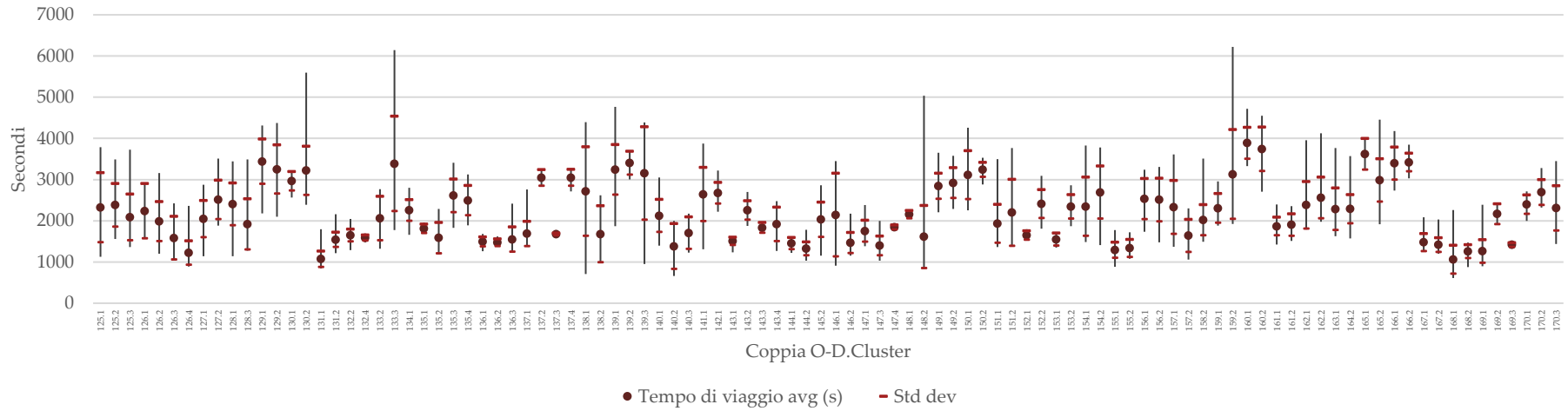
Variazione del tempo di viaggio per O-D e cluster



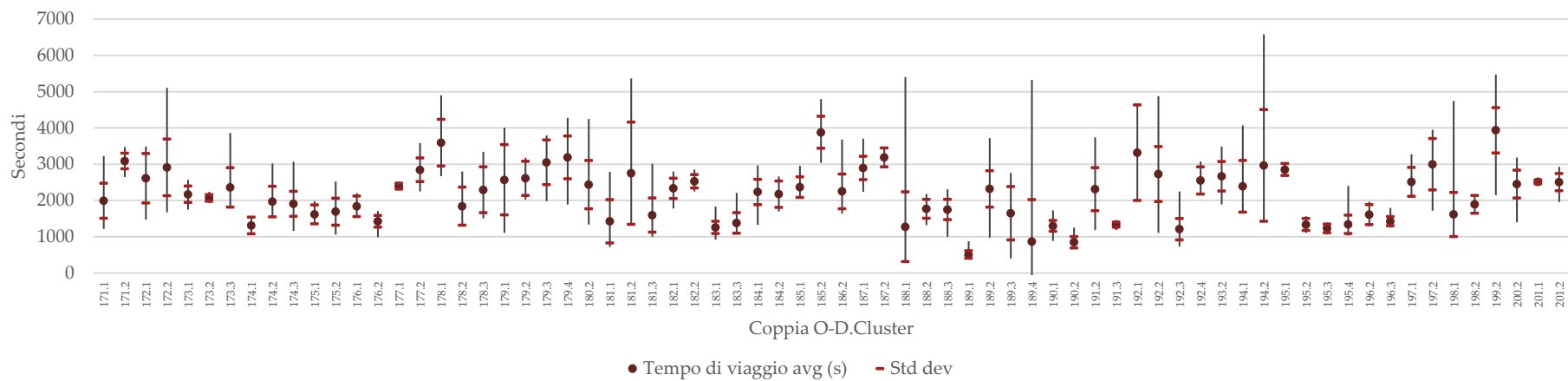
Variazione del tempo di viaggio per O-D e cluster



Variazione del tempo di viaggio per O-D e cluster



Variazione del tempo di viaggio per O-D e cluster



Come si può notare, in accordo con il principio di Wardrop, i tempi di viaggio medi di cluster appartenenti alla stessa coppia O-D risultano piuttosto simili tra loro.



L'operazione di clusterizzazione dei percorsi riduce la dispersione dei tempi di viaggio e introduce dei percorsi rappresentativi, anziché tenere conto della totalità delle alternative di viaggio, spesso simili tra loro. Tuttavia, anche analizzando un set di viaggi piuttosto esteso, può accadere che a un cluster sia assegnato un numero molto ridotto di percorsi.

Da qui nasce il tema della significatività statistica dei risultati relativi all'analisi dei tempi di viaggio.

Nella tabella seguente sono riportati i risultati dell'analisi statistica condotta sulle coppie O-D e i cluster oggetto di analisi, da cui nuovamente sono stati esclusi i cluster contenenti un solo percorso.

Tabella 10 – Risultati dell'analisi statistica per coppie O-D e cluster

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
1	1	5	1666	408	No	1599	67	4%
1	2	40	1627	336	Si	1599	27	2%
1	3	17	1516	319	Si	1599	-84	5%
2	1	11	1488	124	Si	2007	-519	26%
2	2	60	2103	245	Si	2007	95	5%
3	1	75	2355	523	Si	2355	0	0%
4	1	39	1529	199	Si	1548	-20	1%
4	2	28	1576	461	Si	1548	27	2%
5	1	16	1576	257	Si	1428	148	10%
5	2	52	1451	290	Si	1428	23	2%
5	3	11	1078	508	No	1428	-350	24%
5	4	16	1444	387	Si	1428	16	1%
6	1	8	1561	279	Si	3459	-1898	55%
6	2	50	3898	1197	Si	3459	439	13%
6	3	4	3460	320	Si	3459	1	0%
6	4	5	2100	622	No	3459	-1359	39%
7	1	22	1421	415	Si	1526	-106	7%
7	2	54	1594	284	Si	1526	68	4%
7	3	3	1125	91	Si	1526	-402	26%
7	4	3	1484	212	Si	1526	-42	3%
8	1	61	2576	386	Si	2563	13	1%
8	3	13	2500	291	Si	2563	-62	2%
9	1	70	2553	1143	Si	2568	-15	1%
9	2	2	3848	251	Si	2568	1280	50%
9	4	4	2193	1222	No	2568	-375	15%
10	1	50	3547	970	Si	3624	-77	2%
10	2	12	3943	571	Si	3624	319	9%
11	1	62	2947	1128	Si	2667	280	11%
11	2	17	2338	868	Si	2667	-329	12%
11	3	8	2436	861	No	2667	-231	9%
11	4	7	1248	155	Si	2667	-1419	53%
12	1	10	3533	651	Si	3045	488	16%
12	2	45	2622	446	Si	3045	-423	14%
12	3	7	2613	288	Si	3045	-432	14%
12	4	24	3760	472	Si	3045	715	23%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
13	1	21	1762	313	Si	1847	-85	5%
13	2	27	2099	442	Si	1847	251	14%
13	3	9	1380	83	Si	1847	-467	25%
13	4	6	1714	191	Si	1847	-134	7%
14	1	57	1616	347	Si	1637	-21	1%
14	2	20	1698	249	Si	1637	60	4%
15	1	38	1527	313	Si	1696	-169	10%
15	2	43	1845	430	Si	1696	149	9%
16	1	4	791	76	Si	2397	-1606	67%
16	3	62	2501	385	Si	2397	104	4%
17	1	52	3868	1191	Si	3759	109	3%
17	2	4	2414	1241	No	3759	-1345	36%
17	3	6	3715	1313	No	3759	-44	1%
18	1	6	2857	257	Si	2663	194	7%
18	2	24	2515	413	Si	2663	-148	6%
18	3	14	2829	792	Si	2663	166	6%
18	4	11	2669	455	Si	2663	6	0%
19	1	2	1898	13	Si	3577	-1679	47%
19	2	13	3557	642	Si	3577	-21	1%
19	3	42	3664	1076	Si	3577	86	2%
20	1	43	2271	379	Si	2261	10	0%
20	2	47	2252	497	Si	2261	-9	0%
21	1	24	2136	377	Si	2126	10	0%
21	3	31	2118	507	Si	2126	-8	0%
22	1	13	2001	588	Si	1916	85	4%
22	3	50	1894	382	Si	1916	-22	1%
23	1	26	3106	306	Si	2994	112	4%
23	2	8	2569	164	Si	2994	-425	14%
23	3	22	3016	182	Si	2994	22	1%
24	1	2	2816	489	No	2489	327	13%
24	2	56	2477	304	Si	2489	-12	0%
25	1	24	1698	312	Si	2018	-320	16%
25	2	34	2244	593	Si	2018	226	11%
26	1	52	1957	468	Si	1948	9	0%
26	2	4	1826	393	No	1948	-122	6%
27	1	63	1354	498	Si	1370	-16	1%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
27	3	4	1620	136	Si	1370	250	18%
28	1	17	1687	170	Si	1668	19	1%
28	2	3	1574	112	Si	1668	-94	6%
28	3	39	1667	292	Si	1668	-1	0%
29	2	59	1797	291	Si	1797	0	0%
30	1	45	2446	321	Si	2364	82	3%
30	2	6	2684	752	No	2364	320	14%
30	3	2	3368	1587	No	2364	1004	42%
30	4	18	1942	310	Si	2364	-423	18%
31	1	2	3909	1081	No	2267	1642	72%
31	2	55	2207	314	Si	2267	-60	3%
32	1	55	1902	189	Si	1902	0	0%
33	1	48	2567	499	Si	2607	-41	2%
33	2	7	2885	264	Si	2607	278	11%
34	1	2	2829	165	Si	2165	664	31%
34	2	50	2129	245	Si	2165	-37	2%
34	3	3	2332	298	Si	2165	166	8%
35	1	2	2739	44	Si	2033	705	35%
35	2	57	2009	496	Si	2033	-25	1%
36	1	39	1772	414	Si	1680	92	5%
36	2	34	1574	237	Si	1680	-105	6%
37	1	8	2581	841	No	2288	293	13%
37	2	10	2078	224	Si	2288	-210	9%
37	3	38	2281	402	Si	2288	-7	0%
38	1	3	2438	206	Si	1585	853	54%
38	2	46	1530	233	Si	1585	-56	4%
39	1	31	1914	160	Si	1932	-18	1%
39	3	7	1795	219	Si	1932	-137	7%
39	4	13	2049	653	Si	1932	117	6%
40	1	3	3832	712	No	2959	873	30%
40	2	50	2907	579	Si	2959	-52	2%
41	1	47	2881	439	Si	2910	-29	1%
41	2	5	3185	224	Si	2910	275	9%
42	1	22	1873	199	Si	1831	42	2%
42	2	37	1806	207	Si	1831	-25	1%
43	1	5	3154	711	No	3679	-524	14%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
43	2	3	3980	603	No	3679	301	8%
43	3	19	2999	416	Si	3679	-680	18%
43	4	23	4315	510	Si	3679	636	17%
44	1	6	2827	960	No	2149	678	32%
44	2	6	1990	276	Si	2149	-159	7%
44	3	4	2895	340	Si	2149	746	35%
44	4	41	2000	384	Si	2149	-149	7%
45	1	8	3174	342	Si	3020	154	5%
45	2	48	2995	556	Si	3020	-26	1%
46	1	5	2696	127	Si	2174	523	24%
46	2	36	2110	463	Si	2174	-64	3%
46	3	7	2339	471	Si	2174	165	8%
46	4	2	1445	159	No	2174	-729	34%
47	1	32	1930	387	Si	1917	13	1%
47	2	21	1897	142	Si	1917	-20	1%
48	2	61	1558	451	Si	1558	0	0%
49	1	5	3086	366	Si	2427	659	27%
49	2	43	2317	578	Si	2427	-110	5%
49	3	8	2549	278	Si	2427	121	5%
49	4	5	2519	473	Si	2427	92	4%
50	1	53	3185	271	Si	3193	-8	0%
50	2	4	3302	222	Si	3193	109	3%
51	1	2	2103	258	No	1759	344	20%
51	2	46	1750	284	Si	1759	-9	0%
51	3	2	1614	272	No	1759	-145	8%
52	2	14	1368	158	Si	1657	-289	17%
52	3	20	1635	208	Si	1657	-22	1%
52	4	19	1893	208	Si	1657	236	14%
53	1	2	2857	321	No	2364	493	21%
53	2	11	2427	164	Si	2364	63	3%
53	3	40	2319	316	Si	2364	-44	2%
53	4	3	2391	590	No	2364	28	1%
54	1	33	2362	329	Si	2357	5	0%
54	2	10	2362	166	Si	2357	5	0%
54	3	8	2355	294	Si	2357	-2	0%
54	4	4	2308	286	Si	2357	-49	2%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
55	1	2	1628	643	No	1366	262	19%
55	2	26	1443	372	Si	1366	77	6%
55	3	36	1303	311	Si	1366	-63	5%
55	4	11	1343	207	Si	1366	-23	2%
56	1	46	1619	329	Si	1610	9	1%
56	2	18	1586	200	Si	1610	-23	1%
57	1	2	1803	490	No	2390	-587	25%
57	2	43	2418	360	Si	2390	27	1%
58	1	2	3116	605	No	3120	-4	0%
58	2	5	2868	656	No	3120	-252	8%
58	3	40	3151	583	Si	3120	32	1%
59	1	32	2951	358	Si	2814	138	5%
59	2	7	2084	459	Si	2814	-730	26%
59	3	9	2892	279	Si	2814	78	3%
60	2	49	1810	276	Si	1810	0	0%
61	1	15	2202	381	Si	1900	302	16%
61	2	28	1778	289	Si	1900	-122	6%
61	3	5	1678	113	Si	1900	-222	12%
62	1	39	3109	743	Si	3095	14	0%
62	2	8	3029	649	Si	3095	-67	2%
63	1	2	4222	60	Si	2813	1409	50%
63	2	40	2700	554	Si	2813	-112	4%
63	4	3	3372	869	No	2813	559	20%
64	2	29	1609	554	Si	1528	81	5%
64	3	17	1389	173	Si	1528	-139	9%
65	1	48	2142	490	Si	2511	-369	15%
65	3	6	5162	674	Si	2511	2651	106%
65	4	3	3119	1032	No	2511	608	24%
66	2	13	1858	385	Si	1838	20	1%
66	3	35	1831	337	Si	1838	-7	0%
67	2	14	1199	173	Si	1247	-48	4%
67	3	39	1322	188	Si	1247	76	6%
67	4	35	1181	156	Si	1247	-65	5%
68	2	47	1851	301	Si	1926	-75	4%
68	3	4	2808	358	Si	1926	882	46%
69	1	48	1654	285	Si	1654	0	0%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
70	2	24	1774	691	Si	1771	2	0%
70	3	42	1720	462	Si	1771	-51	3%
70	4	5	2191	144	Si	1771	420	24%
71	1	41	2501	471	Si	2531	-30	1%
71	2	8	2686	599	Si	2531	155	6%
72	1	53	1789	333	Si	1762	27	2%
72	2	10	1618	270	Si	1762	-144	8%
73	1	56	1757	270	Si	1757	0	0%
74	1	18	2714	544	Si	2671	43	2%
74	3	26	2641	320	Si	2671	-30	1%
75	1	47	1829	294	Si	1853	-24	1%
75	2	2	2411	544	No	1853	558	30%
76	1	2	3123	1101	No	1375	1748	127%
76	2	62	1319	310	Si	1375	-56	4%
77	2	41	2118	260	Si	2128	-10	0%
77	3	2	2334	181	Si	2128	206	10%
78	1	33	1592	314	Si	1564	28	2%
78	2	19	1515	311	Si	1564	-48	3%
79	1	36	2208	613	Si	2333	-125	5%
79	3	11	2742	454	Si	2333	409	18%
80	1	7	2207	386	Si	2093	115	5%
80	2	38	2072	493	Si	2093	-21	1%
81	1	42	2555	639	Si	2555	0	0%
82	2	41	1709	337	Si	1709	0	0%
83	1	23	1291	189	Si	1328	-37	3%
83	2	36	1352	214	Si	1328	24	2%
84	1	13	3999	465	Si	3679	321	9%
84	2	31	3544	601	Si	3679	-134	4%
85	1	33	3196	835	Si	3128	68	2%
85	2	9	2881	582	Si	3128	-248	8%
86	2	2	1806	14	Si	1355	451	33%
86	3	49	1337	165	Si	1355	-18	1%
87	1	26	2321	744	Si	2372	-51	2%
87	3	16	2455	874	Si	2372	83	4%
88	1	25	2430	519	Si	2205	225	10%
88	2	18	1893	437	Si	2205	-313	14%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
89	1	47	2404	678	Si	2404	0	0%
90	1	4	2312	253	Si	2361	-50	2%
90	2	38	2366	215	Si	2361	5	0%
91	2	33	2837	228	Si	2767	70	3%
91	3	10	2538	444	Si	2767	-230	8%
92	1	25	2115	614	Si	1940	175	9%
92	2	19	1710	400	Si	1940	-230	12%
93	1	3	1028	147	Si	1385	-357	26%
93	2	2	3238	2010	No	1385	1852	134%
93	4	44	1325	147	Si	1385	-60	4%
94	1	41	2087	604	Si	2125	-38	2%
94	2	2	2900	319	No	2125	775	36%
95	1	4	2765	1512	No	1429	1336	93%
95	2	2	1039	8	Si	1429	-391	27%
95	3	52	1334	206	Si	1429	-95	7%
95	4	2	1627	77	Si	1429	198	14%
96	1	44	3231	589	Si	3255	-24	1%
96	2	2	3789	808	No	3255	534	16%
97	1	21	3955	392	Si	3564	391	11%
97	2	5	2028	229	Si	3564	-1536	43%
97	3	14	3527	1237	Si	3564	-37	1%
98	1	18	3676	429	Si	3437	239	7%
98	2	24	3257	577	Si	3437	-179	5%
99	1	24	2212	232	Si	2531	-318	13%
99	2	16	3008	263	Si	2531	477	19%
100	1	42	1314	147	Si	1352	-38	3%
100	2	6	1619	381	Si	1352	267	20%
101	2	19	2897	318	Si	2835	62	2%
101	3	24	2786	591	Si	2835	-49	2%
102	1	22	1486	330	Si	1526	-40	3%
102	2	22	1566	197	Si	1526	40	3%
103	1	5	1950	487	No	1762	188	11%
103	2	26	1757	867	Si	1762	-5	0%
103	3	3	2161	801	No	1762	399	23%
103	4	14	1618	594	Si	1762	-144	8%
104	1	41	1611	585	Si	1611	0	0%



ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
105	1	3	3360	967	No	2375	985	41%
105	2	22	2082	388	Si	2375	-293	12%
105	3	6	3790	734	Si	2375	1415	60%
105	4	9	1818	223	Si	2375	-556	23%
106	1	3	2956	462	No	2362	593	25%
106	3	10	2011	457	Si	2362	-351	15%
106	4	26	2429	666	Si	2362	67	3%
107	1	33	1592	324	Si	1420	173	12%
107	2	24	1182	142	Si	1420	-238	17%
108	2	27	1336	150	Si	1374	-38	3%
108	3	20	1425	193	Si	1374	51	4%
109	1	22	1424	97	Si	1444	-20	1%
109	2	18	1468	131	Si	1444	24	2%
110	2	39	3177	775	Si	3177	0	0%
111	1	4	1478	39	Si	1500	-22	1%
111	2	2	1356	49	Si	1500	-145	10%
111	3	34	1511	163	Si	1500	11	1%
112	2	23	1760	284	Si	1808	-48	3%
112	3	2	1438	82	Si	1808	-370	20%
112	4	17	1916	194	Si	1808	108	6%
113	1	2	2347	111	Si	2435	-88	4%
113	2	3	2400	260	Si	2435	-35	1%
113	3	28	2382	203	Si	2435	-53	2%
113	4	7	2685	187	Si	2435	251	10%
114	2	46	2038	376	Si	2038	0	0%
115	1	63	1306	335	Si	1249	58	5%
115	2	4	882	55	Si	1249	-367	29%
115	3	26	1235	903	No	1249	-13	1%
115	4	6	944	167	Si	1249	-304	24%
116	1	5	2095	1833	No	1809	286	16%
116	2	44	1777	597	Si	1809	-32	2%
117	1	33	1836	217	Si	1841	-6	0%
117	2	3	1902	42	Si	1841	61	3%
118	1	2	1779	287	No	1933	-154	8%
118	2	38	1941	433	Si	1933	8	0%
119	1	8	1871	610	No	2015	-144	7%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
119	2	32	2050	437	Si	2015	36	2%
120	1	30	2467	371	Si	2473	-6	0%
120	2	7	2497	665	Si	2473	24	1%
121	1	29	1412	207	Si	1453	-40	3%
121	2	12	1550	254	Si	1453	98	7%
122	1	37	2391	371	Si	2391	0	0%
123	1	34	1832	237	Si	1822	11	1%
123	2	4	1732	207	Si	1822	-90	5%
124	1	42	1411	354	Si	1494	-84	6%
124	2	3	2664	753	No	1494	1169	78%
125	1	6	2323	842	No	2246	77	3%
125	2	18	2380	526	Si	2246	134	6%
125	3	18	2087	558	Si	2246	-160	7%
126	1	2	2237	666	No	1474	762	52%
126	2	20	1982	480	Si	1474	508	34%
126	3	4	1583	524	No	1474	109	7%
126	4	48	1222	290	Si	1474	-252	17%
127	1	14	2045	448	Si	2333	-289	12%
127	2	23	2509	473	Si	2333	176	8%
128	1	19	2404	514	Si	2174	230	11%
128	3	17	1916	617	Si	2174	-257	12%
129	1	21	3439	546	Si	3354	85	3%
129	2	17	3248	591	Si	3354	-105	3%
130	1	4	2963	231	Si	3193	-230	7%
130	2	37	3218	591	Si	3193	25	1%
131	1	20	1072	192	Si	1374	-302	22%
131	2	36	1542	180	Si	1374	168	12%
132	2	24	1649	150	Si	1628	21	1%
132	4	12	1585	67	Si	1628	-43	3%
133	2	4	2058	533	No	3246	-1188	37%
133	3	34	3385	1154	Si	3246	140	4%
134	1	36	2256	257	Si	2256	0	0%
135	1	3	1811	107	Si	2288	-477	21%
135	2	9	1583	371	Si	2288	-705	31%
135	3	12	2610	403	Si	2288	323	14%
135	4	19	2493	363	Si	2288	205	9%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
136	1	16	1492	115	Si	1516	-24	2%
136	2	4	1471	90	Si	1516	-45	3%
136	3	18	1547	299	Si	1516	31	2%
137	1	23	1686	301	Si	1964	-278	14%
137	2	2	3046	198	Si	1964	1081	55%
137	3	4	1673	26	Si	1964	-291	15%
137	4	5	3045	199	Si	1964	1081	55%
138	1	40	2716	1079	Si	2643	73	3%
138	2	3	1677	686	No	2643	-967	37%
139	1	21	3242	608	Si	3228	14	0%
139	2	3	3403	282	Si	3228	175	5%
139	3	11	3153	1129	Si	3228	-75	2%
140	1	30	2120	395	Si	2024	97	5%
140	2	3	1378	550	No	2024	-646	32%
140	3	3	1704	385	No	2024	-320	16%
141	1	35	2642	652	Si	2642	0	0%
142	1	34	2674	257	Si	2674	0	0%
143	1	11	1507	97	Si	1998	-491	25%
143	2	24	2254	230	Si	1998	256	13%
143	3	2	1833	123	Si	1998	-165	8%
143	4	5	1918	414	Si	1998	-80	4%
144	1	4	1450	142	Si	1333	118	9%
144	2	40	1321	162	Si	1333	-12	1%
145	2	35	2030	423	Si	2030	0	0%
146	1	4	2142	1011	No	1537	605	39%
146	2	34	1466	250	Si	1537	-71	5%
147	1	16	1751	260	Si	1557	194	12%
147	3	23	1397	233	Si	1557	-160	10%
147	4	2	1844	55	Si	1557	287	18%
148	1	2	2153	92	Si	1633	520	32%
148	2	46	1611	760	Si	1633	-23	1%
149	1	24	2841	310	Si	2864	-23	1%
149	2	10	2919	366	Si	2864	56	2%
150	1	23	3113	587	Si	3157	-44	1%
150	2	12	3241	176	Si	3157	84	3%
151	1	22	1928	467	Si	2033	-105	5%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
151	2	14	2198	805	Si	2033	165	8%
152	1	3	1646	108	Si	2346	-700	30%
152	2	32	2411	342	Si	2346	66	3%
153	1	3	1548	152	Si	2275	-727	32%
153	2	32	2343	290	Si	2275	68	3%
154	1	16	2345	711	Si	2531	-186	7%
154	2	19	2687	638	Si	2531	156	6%
155	1	54	1288	189	Si	1294	-6	0%
155	2	8	1335	213	Si	1294	41	3%
156	1	13	2532	490	Si	2517	15	1%
156	2	24	2509	525	Si	2517	-8	0%
157	1	22	2329	645	Si	2085	244	12%
157	2	12	1638	395	Si	2085	-447	21%
158	2	31	2017	373	Si	2017	0	0%
159	1	15	2305	352	Si	2775	-470	17%
159	2	20	3127	1084	Si	2775	352	13%
160	1	29	3886	381	Si	3840	46	1%
160	2	13	3737	533	Si	3840	-103	3%
161	1	21	1865	220	Si	1877	-13	1%
161	2	11	1901	264	Si	1877	24	1%
162	1	10	2380	573	Si	2506	-125	5%
162	2	24	2558	500	Si	2506	52	2%
163	1	32	2283	510	Si	2283	0	0%
164	2	33	2286	348	Si	2286	0	0%
165	1	2	3619	380	No	3021	598	20%
165	2	32	2984	520	Si	3021	-37	1%
166	1	17	3394	395	Si	3406	-12	0%
166	2	17	3418	219	Si	3406	12	0%
167	1	12	1477	212	Si	1437	40	3%
167	2	22	1415	171	Si	1437	-22	2%
168	1	72	1059	345	Si	1093	-34	3%
168	2	15	1256	162	Si	1093	163	15%
169	1	64	1259	280	Si	1291	-31	2%
169	2	2	2166	247	No	1291	875	68%
169	3	2	1417	50	Si	1291	126	10%
170	1	6	2397	228	Si	2402	-5	0%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
170	2	7	2690	310	Si	2402	289	12%
170	3	21	2307	542	Si	2402	-95	4%
171	1	10	1989	482	Si	2761	-772	28%
171	2	24	3083	211	Si	2761	322	12%
172	1	6	2610	681	No	2854	-244	9%
172	2	28	2906	782	Si	2854	52	2%
173	1	9	2167	224	Si	2237	-70	3%
173	2	8	2065	96	Si	2237	-172	8%
173	3	17	2355	546	Si	2237	118	5%
174	1	2	1307	229	No	1904	-598	31%
174	2	19	1965	419	Si	1904	60	3%
174	3	23	1907	345	Si	1904	2	0%
175	1	3	1611	262	No	1680	-69	4%
175	2	30	1686	371	Si	1680	7	0%
176	1	3	1835	282	No	1456	379	26%
176	2	33	1422	162	Si	1456	-34	2%
177	1	3	2391	79	Si	2799	-408	15%
177	2	29	2841	327	Si	2799	42	2%
178	1	12	3588	640	Si	2705	883	33%
178	2	5	1838	522	No	2705	-867	32%
178	3	15	2288	632	Si	2705	-417	15%
179	1	8	2566	967	No	2948	-383	13%
179	2	3	2606	471	No	2948	-342	12%
179	3	6	3048	613	Si	2948	100	3%
179	4	15	3181	588	Si	2948	233	8%
180	2	31	2432	668	Si	2432	0	0%
181	1	10	1419	597	No	1725	-306	18%
181	2	6	2746	1411	No	1725	1020	59%
181	3	23	1592	473	Si	1725	-133	8%
182	1	19	2331	279	Si	2405	-74	3%
182	2	12	2522	183	Si	2405	117	5%
183	1	35	1252	170	Si	1279	-27	2%
183	3	10	1375	282	Si	1279	96	7%
184	1	25	2232	347	Si	2220	12	1%
184	2	6	2168	362	Si	2220	-52	2%
185	1	9	2363	285	Si	3435	-1072	31%

ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
185	2	22	3874	438	Si	3435	439	13%
186	2	31	2246	479	Si	2246	0	0%
187	1	32	2894	321	Si	2911	-17	1%
187	2	2	3183	263	Si	2911	272	9%
188	1	25	1271	962	No	1551	-280	18%
188	2	9	1768	260	Si	1551	217	14%
188	3	26	1745	281	Si	1551	194	12%
189	1	19	504	102	Si	1451	-947	65%
189	2	28	2314	503	Si	1451	863	59%
189	3	17	1640	734	Si	1451	189	13%
189	4	16	864	1152	No	1451	-587	40%
190	1	37	1295	147	Si	1207	88	7%
190	2	9	845	158	Si	1207	-362	30%
191	2	21	2309	594	Si	1973	336	17%
191	3	11	1331	70	Si	1973	-641	33%
192	1	2	3313	1314	No	2034	1279	63%
192	2	53	2723	762	Si	2034	689	34%
192	3	49	1205	296	Si	2034	-829	41%
192	4	3	2550	375	Si	2034	516	25%
193	2	31	2662	404	Si	2662	0	0%
194	1	8	2386	712	Si	2816	-430	15%
194	2	24	2959	1537	Si	2816	143	5%
195	1	2	2848	165	Si	1396	1452	104%
195	2	4	1333	163	Si	1396	-63	5%
195	3	5	1227	120	Si	1396	-169	12%
195	4	31	1337	250	Si	1396	-58	4%
196	2	3	1606	276	No	1442	164	11%
196	3	29	1425	124	Si	1442	-17	1%
197	1	9	2507	398	Si	2847	-341	12%
197	2	21	2993	708	Si	2847	146	5%
198	1	29	1610	610	Si	1628	-18	1%
198	2	2	1887	247	No	1628	258	16%
199	2	29	3930	626	Si	3930	0	0%
200	2	28	2448	384	Si	2448	0	0%
201	1	2	2509	53	Si	2501	7	0%
201	2	28	2501	236	Si	2501	-1	0%

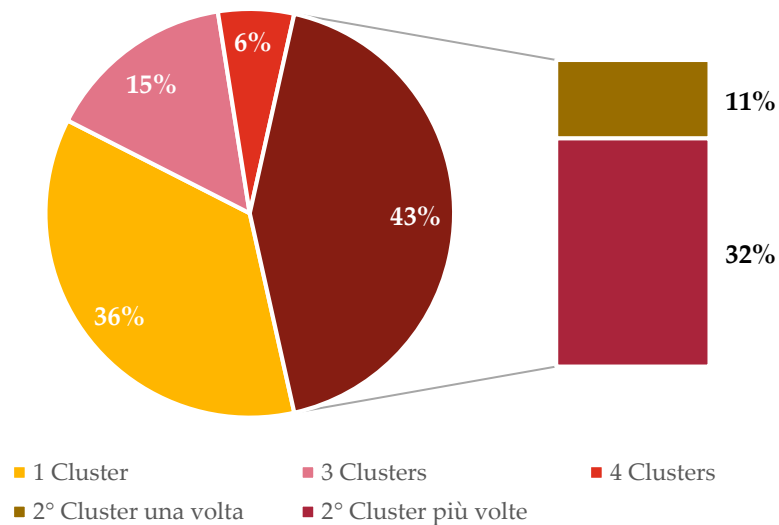
ID Coppia OD	Cluster	Numero di viaggi	Tempo di viaggio avg del cluster (s)	Dev. Standard media del tempo di viaggio (s)	Precisione del 10% con livello di confidenza 80%	Tempo di viaggio avg O-D (s)	Differenza del tempo di viaggio media (s)	Differenza del tempo di viaggio media (%)
	<i>Sum</i>	9809	-	-	400	-	-	-
	<i>Average</i>	21	2224	423	-	2197	27	12%

Nonostante i viaggi effettuati dagli utenti siano distribuiti su un numero significativo di percorsi e siano raccolte poche osservazioni sugli itinerari meno utilizzati, è stata ottenuta una significatività statistica dei risultati con un livello di confidenza dell'80% su 400 dei 467 cluster delle O-D oggetto di analisi.

La differenza tra il tempo di viaggio medio del cluster e il tempo di viaggio medio della coppia O-D è del 12% e del 9% se si prendono in considerazione soltanto i dati statisticamente significativi.

#### 4.2.2 Analisi della dinamica giornaliera

La clusterizzazione dei percorsi analizzati, ha permesso di effettuare una analisi relativa al processo di scelta del percorso abituale da parte degli utenti e alle condizioni che producono possibili cambiamenti nella scelta di tale percorso (ovvero nella scelta di un percorso appartenente ad un altro cluster rispetto a quello del percorso abituale).



*Figura 23 – Variazioni nella scelta del percorso (analisi della dinamica giornaliera)*

I risultati mostrano che il 36% degli utenti stradali conferma la scelta abituale (utilizza sempre un percorso dello stesso cluster).

Il 64% circa dei conducenti, invece, verifica altri percorsi durante il periodo di osservazione.

Di questi ultimi, tuttavia, solo il 21% degli utenti sceglie più di un percorso alternativo rispetto al solito, mentre il 43% un solo percorso alternativo e nel 32% dei casi lo seleziona più volte (sembra avere, pertanto, due percorsi “abituali”).



### *Variazioni di percorso a seguito di un aumento del tempo di percorrenza*

A questo punto, è stata verificata l'eventuale esistenza di una correlazione tra la scelta di cambiare il percorso e i possibili ritardi nei tempi di percorrenza riscontrati dai conducenti durante il viaggio precedente al cambio.

È stata stimata la differenza tra il tempo medio di percorrenza sperimentato dall'utente sul percorso nei giorni precedenti e il valore registrato durante il viaggio precedente il cambio di percorso:

$$27) \quad \Delta_t = t_{x-1} - \frac{\sum_{i=1}^{x-2} t_i}{N}$$

A seguito dell'incremento o della riduzione del tempo di viaggio, è stata investigata la propensione al cambio di percorso o al cambio di orario di partenza nel viaggio successivo.

I risultati ottenuti sono sintetizzati nella tabella successiva e mostrano un sostanziale equilibrio tra i casi in cui la partenza venga ritardata o anticipata, sia tra utenti che cambiano percorso (totale casi 2242), sia tra utenti che confermano la precedente scelta (totale casi 7285).

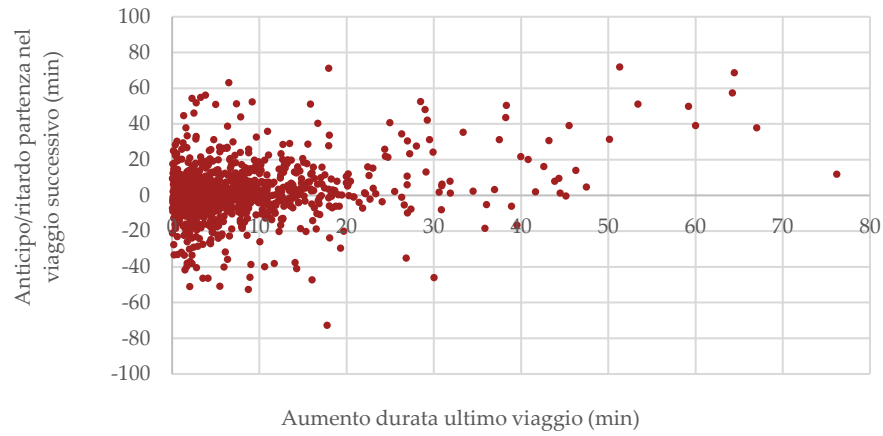
*Tabella 11 – Propensione al cambio di percorso o al cambio di orario di partenza a seguito di un incremento o una riduzione del tempo di viaggio sperimentato*

Viaggio x-1	Viaggio x		Numero casi	% Casi
Aumento* ( $\Delta t > 0$ )	Cambio percorso	Anticipo orario di partenza	489	45%
Aumento* ( $\Delta t > 0$ )	Cambio percorso	Ritardo orario di partenza	582	54%
Aumento* ( $\Delta t > 0$ )	Cambio percorso	Non cambio orario	6	1%
Aumento* ( $\Delta t > 0$ )	Non cambio percorso	Anticipo orario di partenza	1671	48%
Aumento* ( $\Delta t > 0$ )	Non cambio percorso	Ritardo orario di partenza	1794	51%
Aumento* ( $\Delta t > 0$ )	Non cambio percorso	Non cambio orario	22	1%
Riduzione* ( $\Delta t < 0$ )	Cambio percorso	Anticipo orario di partenza	591	51%
Riduzione* ( $\Delta t < 0$ )	Cambio percorso	Ritardo orario di partenza	569	49%
Riduzione* ( $\Delta t < 0$ )	Cambio percorso	Non cambio orario	5	0%
Riduzione* ( $\Delta t < 0$ )	Non cambio percorso	Anticipo orario di partenza	1891	50%
Riduzione* ( $\Delta t < 0$ )	Non cambio percorso	Ritardo orario di partenza	1893	50%
Riduzione* ( $\Delta t < 0$ )	Non cambio percorso	Non cambio orario	14	0%

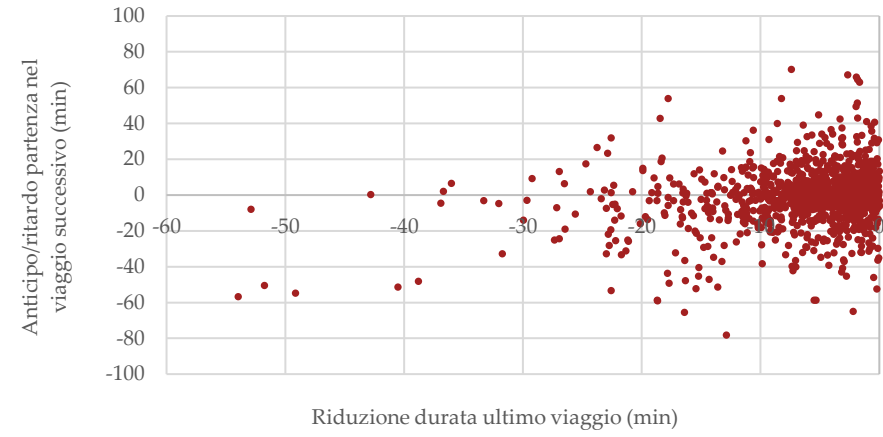
\*del tempo di viaggio rispetto al viaggi precedenti.

Nelle immagini seguenti sono rappresentate graficamente le casistiche sintetizzate nella tabella precedente.

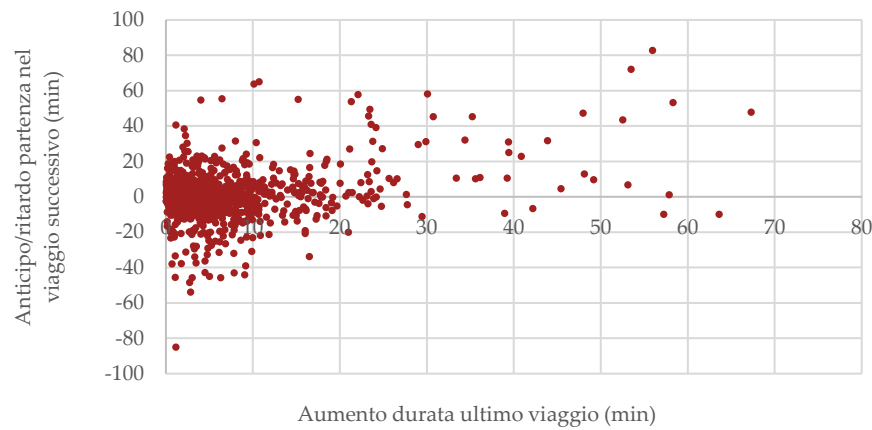
$\Delta_t > 0$  (Aumento) e cambio di percorso



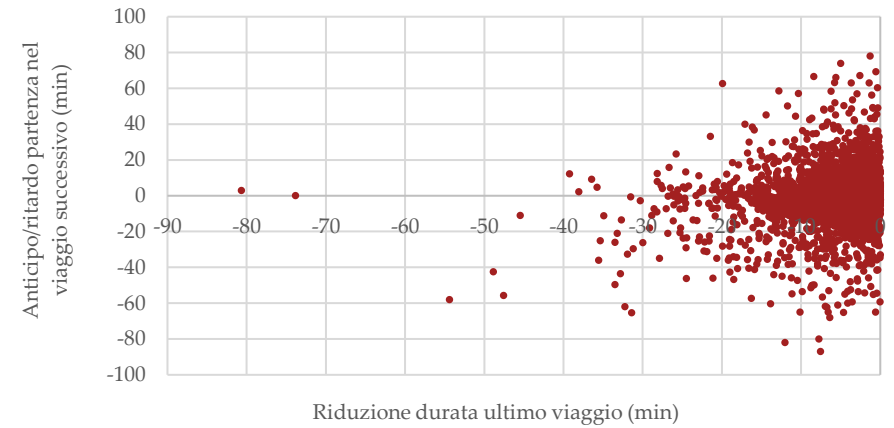
$\Delta_t < 0$  (Riduzione) e cambio di percorso



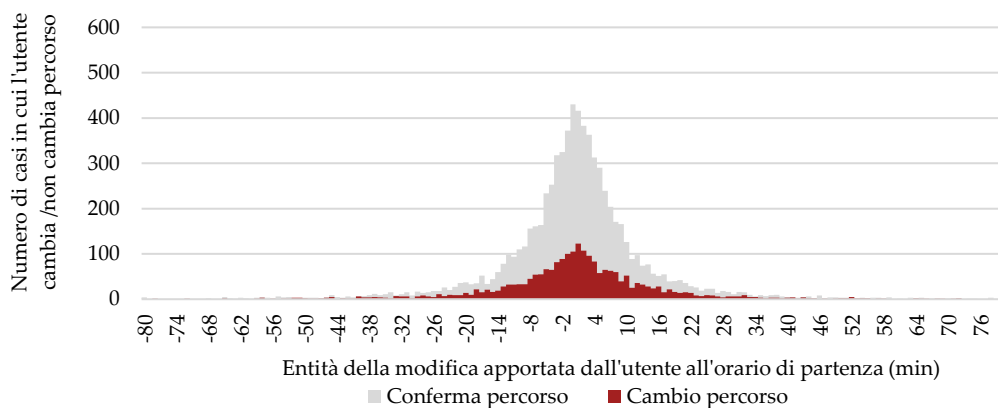
$\Delta_t > 0$  (Aumento) e nessun cambio di percorso



$\Delta_t < 0$  (Riduzione) e nessun cambio di percorso

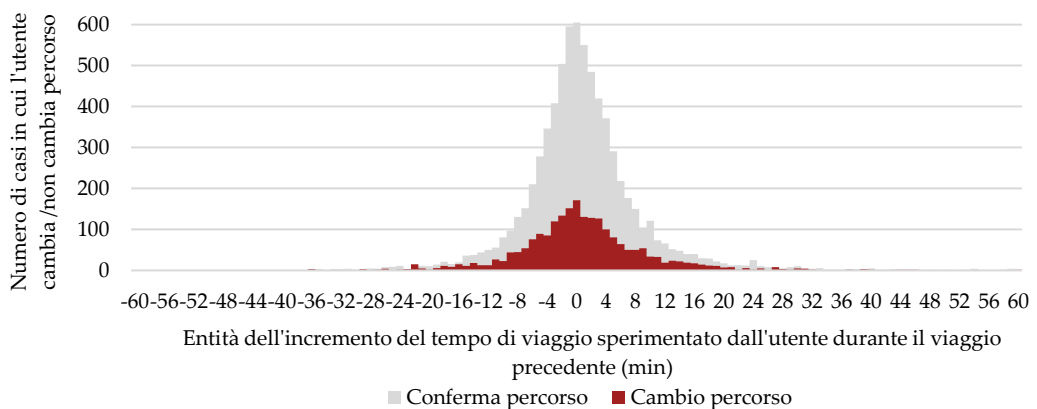


Si riporta di seguito l'analisi della distribuzione delle modifiche apportate all'orario di partenza nel caso di utenti che cambiano / non cambiano percorso (sia in caso di anticipo, che di riduzione del tempo di viaggio sperimentato durante il viaggio precedente): è stato ottenuto un valore medio totale (tra chi conferma e chi cambia il percorso) di ritardo dell'orario di circa 9 minuti e di anticipo pari a quasi 10 minuti.



*Figura 24 – Distribuzione delle modifiche apportate all'orario di partenza nel caso di utenti che confermano o cambiano percorso*

Gli incrementi / riduzioni dei tempi di viaggio sperimentati dagli utenti sono generalmente molto ridotti: il valore medio totale (tra chi conferma e chi cambia il percorso) dell'aumento del tempo di viaggio è pari a circa 6 minuti, mentre il valore medio totale della riduzione a circa 5 minuti.



*Figura 25 - Distribuzione degli incrementi di tempo sperimentati da utenti che confermano o cambiano percorso*

Analizzando la propensione al solo cambio di orario di partenza a seguito di un aumento / riduzione del tempo di percorrenza sperimentato durante il viaggio precedente si ottiene la stessa situazione di sostanziale equilibrio.

*Tabella 12 – Propensione al solo cambio di orario di partenza a seguito di una variazione sperimentata nel tempo di percorrenza*

Ultimo viaggio effettuato	Viaggio successivo	% casi
Aumento* ( $\Delta t > 0$ )	Anticipo orario di partenza	47%
	Ritardo orario di partenza	52%
	Non cambio orario	1%
Riduzione* ( $\Delta t < 0$ )	Anticipo orario di partenza	50%
	Ritardo orario di partenza	50%
	Non cambio orario	0%

*\*del tempo di viaggio rispetto ai viaggi precedenti.*

Analizzando, invece, la propensione al solo cambio di percorso, è emersa una tendenza a confermare il percorso scelto durante il precedente viaggio, a prescindere dall'aumento o dalla riduzione del tempo di percorrenza sperimentato.

*Tabella 13 – Propensione al solo cambio di percorso a seguito di una variazione sperimentata nel tempo di percorrenza*

Ultimo viaggio effettuato	Viaggio successivo	% casi
Aumento* ( $\Delta t > 0$ )	Cambio percorso	24%
	Non cambio percorso	76%
Riduzione* ( $\Delta t < 0$ )	Cambio percorso	23%
	Non cambio percorso	77%

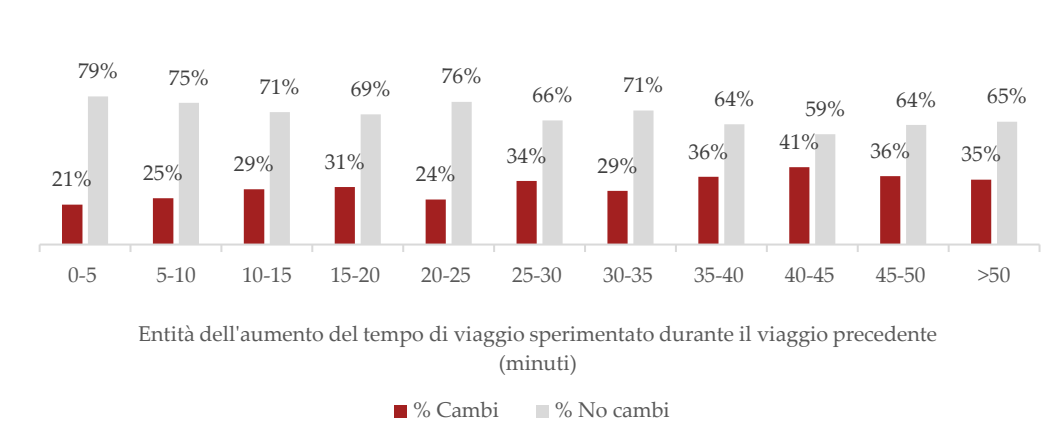
*\*del tempo di viaggio rispetto ai viaggi precedenti.*

Di seguito è rappresentata la distribuzione della propensione a cambiare / confermare la scelta del percorso al variare dell'entità dell'aumento / riduzione del tempo di viaggio sperimentato.

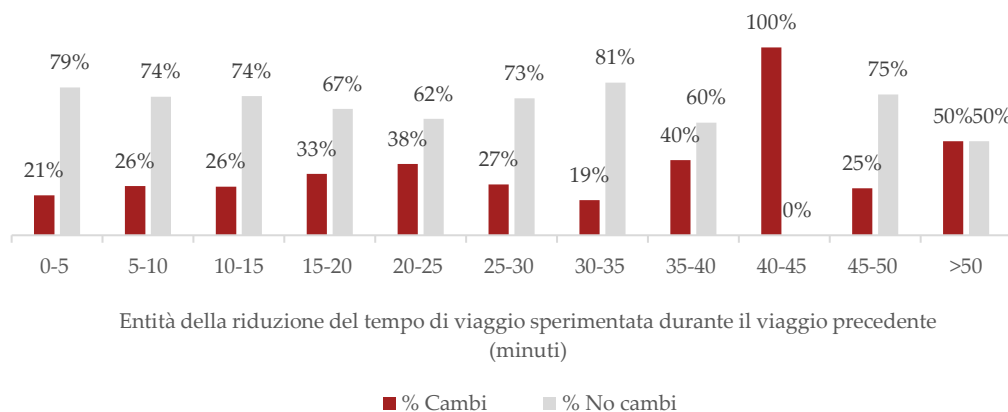
Come si può notare, in corrispondenza di entità diverse di aumento / riduzione del tempo di percorrenza, gli utenti tendono comunque a confermare il percorso abituale.

L'unico caso in cui vi è una preponderanza di utenti che cambia percorso a seguito di una riduzione di circa 40 minuti, è rappresentato nel secondo

grafico. Si tratta, tuttavia, di due soli casi di utenti che cambiano a fronte di nessun caso di utenti che cambiano percorso.



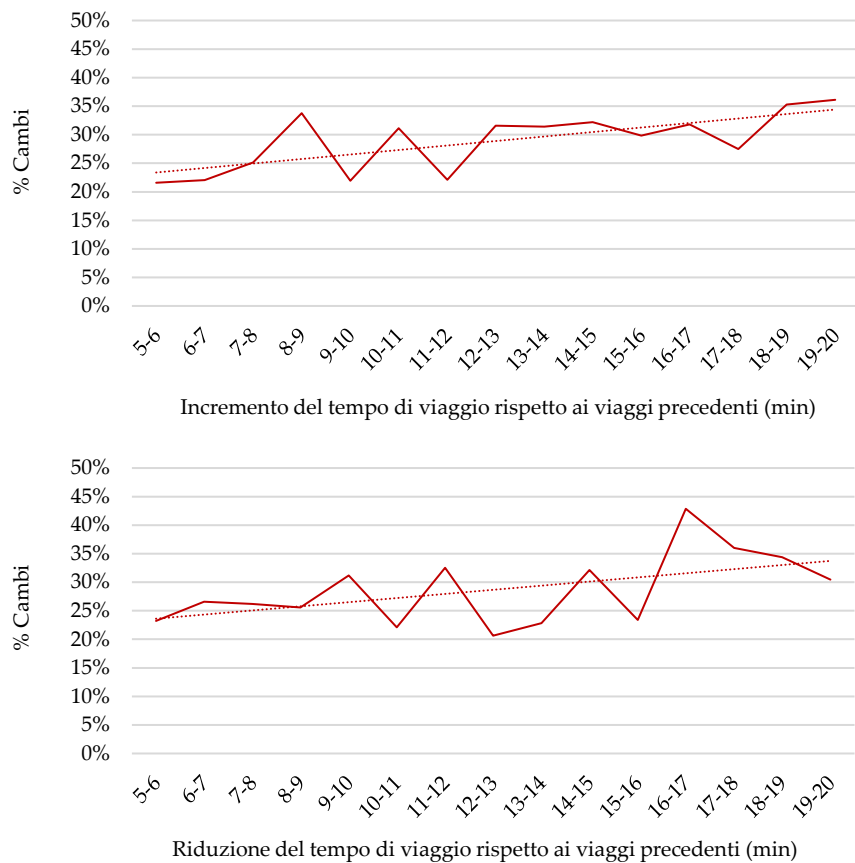
**Figura 26 – Percentuale di cambi di percorso e non, in funzione dell'entità dell'aumento di durata sperimentato durante l'ultimo viaggio**



**Figura 27 - Percentuale di cambi di percorso e non, in funzione dell'entità della riduzione di durata sperimentata durante l'ultimo viaggio**

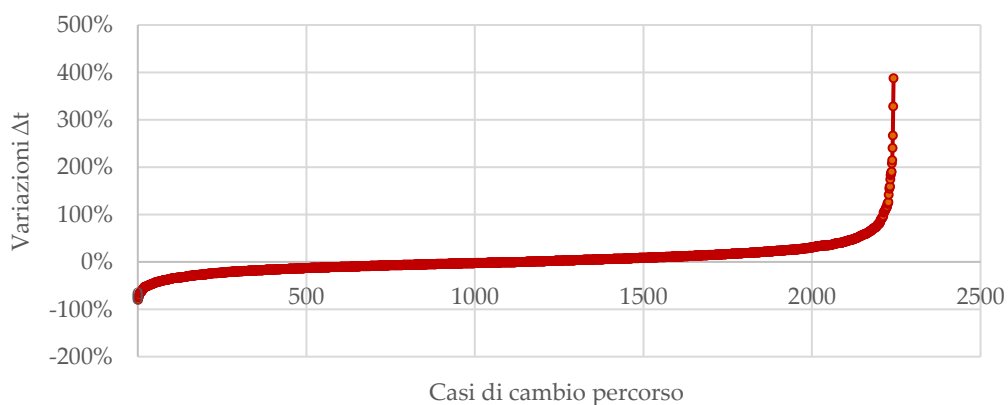
Dalle figure precedenti si nota, inoltre, come incrementi o riduzioni dei tempi di viaggio superiori ai 20 minuti, non influiscano sulla scelta del percorso da parte degli utenti.

Analizzando, invece, incrementi o riduzioni dei tempi di viaggio tra i 5 e i 20 minuti, è emersa una correlazione tra la percentuale di utenti che cambia il percorso e l'entità dell'incremento o della riduzione sperimentata durante l'ultimo viaggio effettuato.



**Figura 28 – Correlazione tra incremento (in alto) o riduzione (in basso) del tempo di viaggio e percentuale di utenti che cambia percorso**

Sono stati poi analizzati i 2242 casi di utenti che cambiano percorso, a prescindere dall'orario di partenza: il grafico seguente ordina i casi in cui il percorso è stato modificato a seguito di un aumento o di una riduzione del tempo di viaggio riscontrato.



**Figura 29 – Numero di casi in cui gli utenti cambiano percorso a seguito di un aumento o una riduzione del tempo di viaggio riscontrato**

Nel 48% dei casi l'utente ha effettuato un cambio di percorso a seguito di un incremento del tempo di percorrenza sperimentato durante il viaggio precedente, nel 52% a seguito di una riduzione.

*Variazioni di percorso a seguito di un ritardo rispetto all'orario di arrivo a destinazione*

Per verificare l'esistenza di una eventuale correlazione tra un ritardo o un anticipo sperimentato dagli utenti rispetto all'orario medio di arrivo a destinazione (anziché una variazione del tempo di viaggio) e la scelta di cambiare o confermare il percorso durante il viaggio successivo, sono stati selezionati i soli utenti che abbiano compiuto almeno 20 viaggi durante il periodo di osservazione

I risultati ottenuti sono molto simili al caso della variazione del tempo di viaggio: è emersa una tendenza a confermare il percorso a prescindere dall'orario di arrivo sperimentato durante l'ultimo viaggio.

*Tabella 14 – Utenti che cambiano e non il percorso a fronte di un anticipo o un ritardo sperimentato durante il viaggio precedente rispetto all'orario di arrivo medio*

Viaggio x-1	Viaggio x	Numero casi	% Casi
Anticipo*	Cambio percorso	1032	22%
	Non cambio percorso	3605	78%
Ritardo*	Cambio percorso	1111	25%
	Non cambio percorso	3359	75%

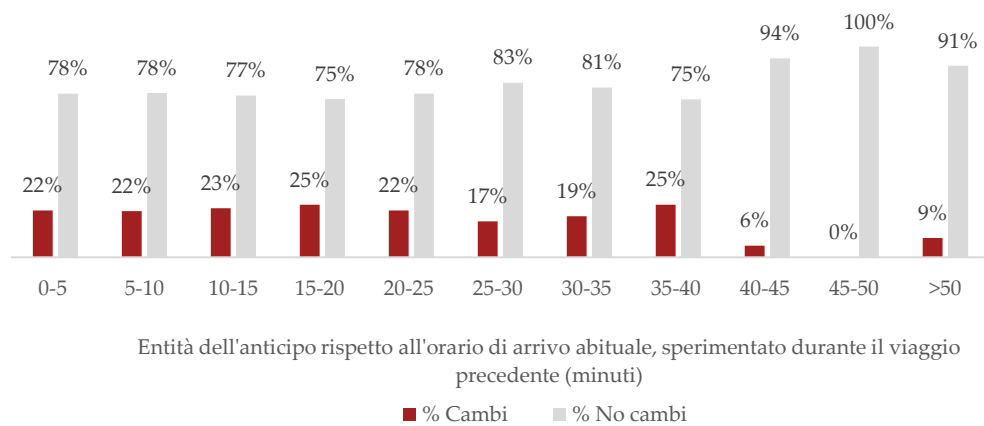
\* Sperimentato durante il viaggio precedente rispetto all'orario di arrivo medio.

Di seguito è rappresentata la distribuzione della propensione a cambiare / confermare la scelta del percorso al variare dell'entità dell'anticipo / ritardo rispetto all'orario di arrivo abituale, sperimentato il viaggio precedente.

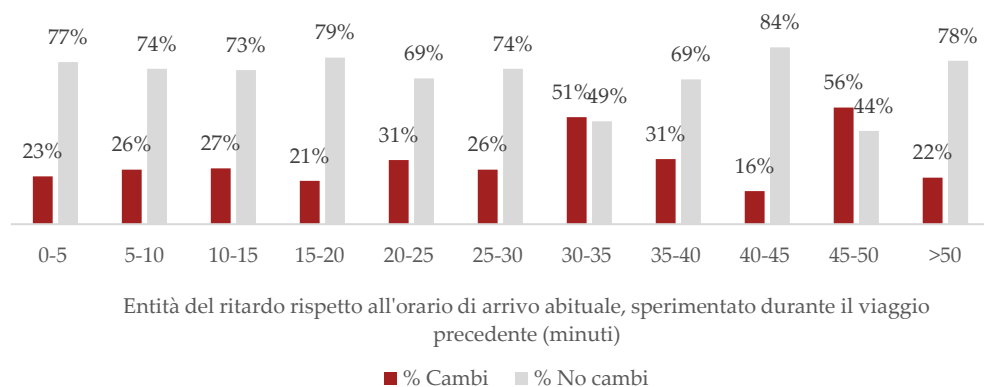
Come si può notare, in corrispondenza di entità diverse di aumento / riduzione del tempo di percorrenza, gli utenti tendono comunque a confermare il percorso abituale.

Di nuovo, l'unico caso in cui vi è una preponderanza di utenti che cambia percorso a seguito di un ritardo nell'orario di arrivo di circa 40 minuti, è rappresentato nel secondo grafico.

Si tratta, tuttavia, di cinque casi di utenti che cambiano a fronte di quattro casi di utenti che non cambiano percorso.



**Figura 30 – Distribuzione del numero di cambi e non di percorso in funzione dell'entità dell'anticipo rispetto all'orario di arrivo abituale**



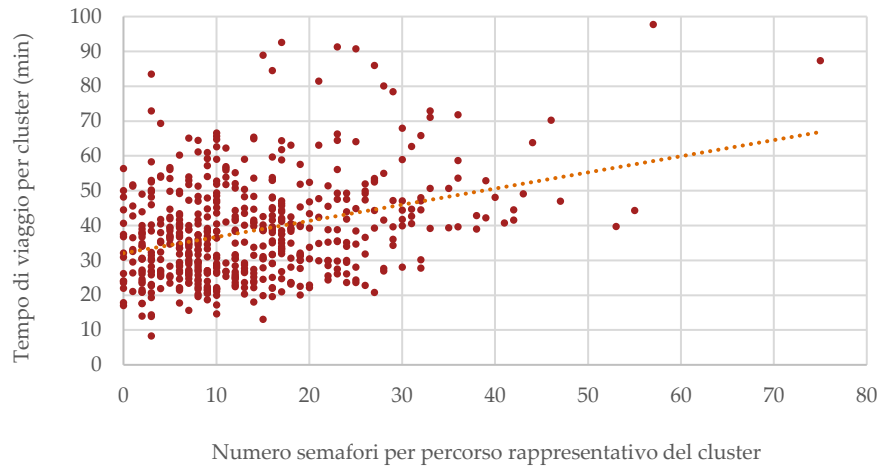
**Figura 31 – Distribuzione del numero di cambi e non di percorso in funzione dell'entità del ritardo rispetto all'orario di arrivo abituale**

*Variazioni di percorso in favore di un minore tempo medio di percorrenza o di un minor numero di semafori*

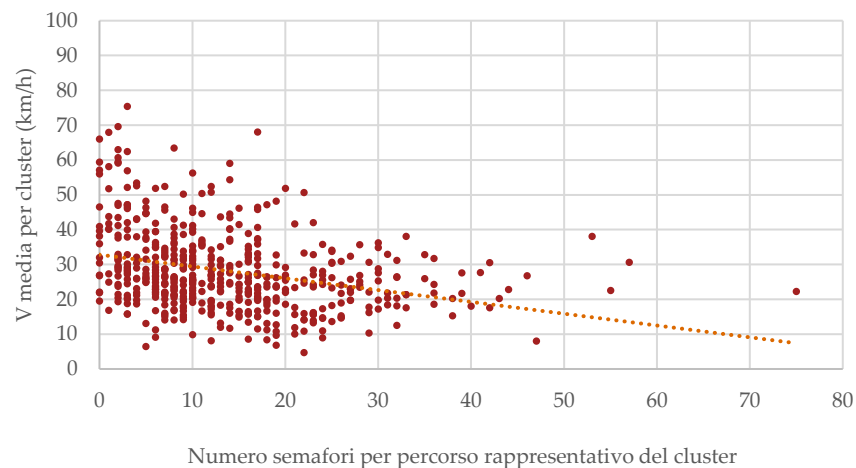
È stato, infine, calcolato il numero di semafori localizzati su ciascun percorso rappresentativo e, per tutti i percorsi appartenenti allo stesso cluster, è stato assunto lo stesso numero di semafori del percorso rappresentativo del cluster.



Nelle figure seguenti sono riportati i risultati dell'analisi di correlazione – relativa alla totalità dei viaggi analizzati – condotta tra il tempo di percorrenza e il numero semafori e la velocità commerciale dei veicoli e il numero di semafori, per ciascuna coppia O-D.



*Figura 32 – Correlazione tra tempo di percorrenza e il numero semafori*



*Figura 33 – Correlazione tra velocità commerciale media e il numero semafori*

Una volta associato a ciascun percorso il numero di semafori, è stata condotta una analisi relativa ai 2242 casi totali in cui gli utenti hanno cambiato percorso.

In particolare, è stata indagata la propensione a cambiare in favore di:

- un percorso appartenente ad un cluster il cui tempo medio di percorrenza è minore rispetto a quello del cluster selezionato durante il viaggio precedente (“condizione A verificata”);

- un percorso caratterizzato da un numero di semafori minore rispetto a quello del percorso selezionato durante il viaggio precedente (“condizione B verificata”).

È stato ottenuto il seguente risultato: il 69% degli utenti cambia percorso in corrispondenza della sola condizione A verificata, della sola condizione B verificata o entrambe le condizioni verificate.

Il 31% cambia percorso in corrispondenza sia di condizione A che di condizione B non verificata, ovvero per altre ragioni rispetto al tempo di viaggio o al numero di semafori che incontra lungo il percorso.

*Tabella 15 – Utenti che cambiano percorso a fronte di un anticipo o un ritardo sperimentato durante il viaggio precedente rispetto all'orario di arrivo medio*

Condizione A	Condizione B	Numero casi	%casi
<b>Verificata</b> (Minore travel time)	<b>Verificata</b> (Minor numero semafori)	414	18%
<b>Verificata</b> (Minore travel time)	<b>Non verificata</b> (Maggior numero semafori)	714	32%
<b>Non verificato</b> (Maggiore travel time)	<b>Verificata</b> (Minor numero semafori)	426	19%
<b>Non verificato</b> (Maggiore travel time)	<b>Non verificata</b> (Maggior numero semafori)	688	31%

#### *Guadagno ottenuto a seguito del cambio di percorso*

Sono stati, infine, indagati i casi in cui gli utenti hanno ottenuto un risparmio o un incremento del tempo di viaggio, a seguito del cambio di percorso, ovvero i casi in cui gli utenti hanno ottenuto o meno un guadagno a seguito del cambiamento.

*Tabella 16 – Utenti che ottengono un guadagno dal cambio di percorso*

	Risparmio di tempo	% casi
<b>Cambio a seguito di un aumento</b>	SI	31%
<b>Cambio a seguito di un aumento</b>	NO	69%
<b>Cambio a seguito di una riduzione</b>	SI	67%
<b>Cambio a seguito di una riduzione</b>	NO	33%

Per ciascun utente stradale di ciascuna coppia O-D è stata calcolata:

- la variazione del tempo di viaggio – rispetto al viaggio precedente - nel caso in cui sia stata anticipata/ritardata la partenza e si sia scelto un percorso diverso rispetto a quello abituale (2242 casi);

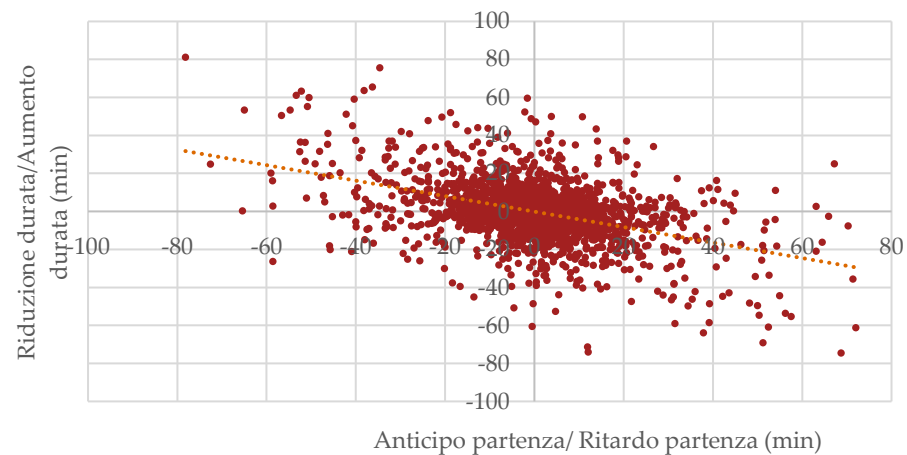
- la variazione del tempo di viaggio – rispetto al viaggio precedente - nel caso in cui sia stata anticipata/ritardata la partenza, a prescindere dal percorso seguito (9527 casi).

I risultati sono riportati di seguito.

In entrambe le ipotesi, è stata registrata una prevalenza di casi in cui:

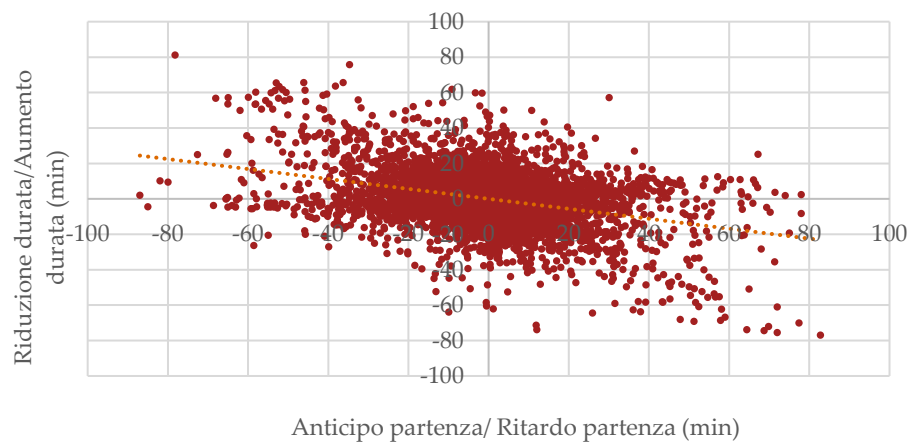
- l'utente ritarda l'orario di partenza e registra una diminuzione della durata del percorso, oppure
- anticipa l'orario di partenza e registra un aumento del tempo di viaggio.

**Variazione del tempo di viaggio – rispetto al viaggio precedente - nel caso in cui sia stata anticipata/ritardata la partenza e si sia scelto un percorso diverso rispetto a quello abituale (2242 casi)**



	1 quadr.	2 quadr.	3 quadr.	4 quadr.
<b>Casi</b>	473	646	434	689
<b>%</b>	21%	29%	19%	31%

**Variazione del tempo di viaggio – rispetto al viaggio precedente - nel caso in cui sia stata anticipata/ritardata la partenza, a prescindere dal percorso seguito (9527 casi)**



	1 quadr.	2 quadr.	3 quadr.	4 quadr.
<b>Casi</b>	2260	2486	2156	2625
<b>%</b>	24%	26%	23%	27%

#### 4.2.3 Calibrazione di un modello di scelta del percorso

Il database, così elaborato, dei 9882 viaggi è stato utilizzato per la calibrazione di un modello logit di scelta del percorso, attraverso il software Biogeme (Bierlaire, M., 2020).

L'input del modello è l'intero set di coppie O-D analizzate: per ciascuna relazione, l'utente ha a disposizione due, tre o quattro alternative di viaggio (in funzione dei risultati ottenuti attraverso la routine di clusterizzazione dei percorsi).

Per rappresentare all'interno del modello questo aspetto, è stato inserito nel database un "indicatore di disponibilità", pari a 1 nel caso in cui l'alternativa è disponibile e a 0 nell'altro caso (ovvero, per le coppie O-D in cui non sono disponibili la terza o la quarta alternativa di viaggio, l'indicatore è pari a 0).

Sono state, poi, definite alcune forme funzionali per l'utilità sistematica di scelta ed è stata effettuata la calibrazione del modello per ciascuna di esse.

Per le seguenti forme funzionali sono stati ottenuti risultati particolarmente interessanti:

1.  $V = \beta_l * l + \beta_v * v$
2.  $V = \beta_s * n_s + \beta_l * l + \beta_v * v$
3.  $V = \beta_A * \%_A + \beta_B * \%_B + \beta_l * l + \beta_v * v$
4.  $V = \beta_l * l + \beta_t * t$
5.  $V = \beta_s * n_s + \beta_l * l + \beta_t * t$
6.  $V = \beta_A * \%_A + \beta_B * \%_B + \beta_l * l + \beta_t * t$

In cui:

- $l$  rappresenta la lunghezza del percorso normalizzata rispetto al valore massimo di lunghezza tra le alternative di viaggio a disposizione su una specifica coppia O-D;
- $v$  rappresenta la velocità del percorso in km/h;

- $n_s$  sono il numero di semafori incontrati dall'utente lungo il percorso;
- $\%_{categ.A}$  è la percentuale – rispetto al totale degli archi che compongono ciascun percorso – di strade di maggiore importanza (autostrade, strade statali e strade urbane di scorrimento, etc.);
- $\%_{categ.B}$  è la percentuale di strade di minore importanza (strade locali) - rispetto al totale degli archi che compongono ciascun percorso. Sono, quindi, state escluse dall'analisi le strade secondarie;
- $t$  è il tempo medio di percorrenza del percorso in secondi.

Per i primi tre modelli, sono stati ottenuti i risultati sintetizzati nella tabella seguente. Come si può notare, il modello è stato calibrato sia imponendo almeno un coefficiente pari a 1, sia non assegnando valore 1 ad alcun coefficiente, in modo da verificare che in entrambi i test sui coefficienti abbiano esito positivo.

Tabella 17 – Risultati della calibrazione dei modelli 1-2-3

<b>Modello 1)</b>							
Risultati calibrazione ( $\beta_v = 1$ )							
Rho-square=0.0203							
Rho-square-bar=0.0202							
Akaike Information Criterion=18853.24							
	<b>Value</b>	<b>Std err</b>	<b>t-test</b>	<b>p-value</b>	<b>Rob. Std err</b>	<b>Rob. t-test</b>	<b>Rob. p-value</b>
$\beta_l$	-1.69979	0.089997	-18.8872	0	0.091526	-18.5716	0
Risultati calibrazione (no $\beta = 1$ )							
Rho-square=0.0239							
Rho-square-bar=0.0237							
Akaike Information Criterion=18786.01							
	<b>Value</b>	<b>Std err</b>	<b>t-test</b>	<b>p-value</b>	<b>Rob. Std err</b>	<b>Rob. t-test</b>	<b>Rob. p-value</b>
$\beta_l$	-1.85547	0.091789	-20.2145	0	0.091341	-20.3135	0
$\beta_v$	0.016936	0.002043	8.289192	2.22E-16	0.002151	7.875334	3.33E-15
<b>Modello 2)</b>							
Risultati calibrazione ( $\beta_v = 1$ )							
Rho-square=0.0458							
Rho-square-bar=0.0456							
Akaike Information Criterion=18364.22							
	<b>Value</b>	<b>Std err</b>	<b>t-test</b>	<b>p-value</b>	<b>Rob. Std err</b>	<b>Rob. t-test</b>	<b>Rob. p-value</b>
$\beta_l$	-1.52946	0.09402	-16.2674	0	0.097665	-15.6603	0
$\beta_s$	-0.06237	0.003056	-20.4124	0	0.003463	-18.0122	0

Risultati calibrazione ( $no \beta = 1$ )							
Rho-square=0.0459							
Rho-square-bar=0.0456							
Akaike Information Criterion=18364.07							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_l$	-1.56429	0.096886	-16.1456	0	0.101222	-15.4541	0
$\beta_s$	-0.06111	0.003176	-19.2452	0	0.003591	-17.0192	0
$\beta_v$	0.003212	0.00219	1.466871	0.142411	0.002246	1.430177	0.152666
Modello 3) <sup>2</sup>							
Risultati calibrazione ( $\beta_v = 1$ )							
Rho-square=0.0397							
Rho-square-bar=0.0394							
Akaike Information Criterion=18378.01							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_A$	0.758197	0.097955	7.740287	9.99E-15	0.095021	7.979253	1.55E-15
$\beta_B$	-2.0673	0.162445	-12.7262	0	0.155722	-13.2756	0
$\beta_l$	-1.62369	0.089862	-18.0687	0	0.089305	-18.1815	0
Risultati calibrazione ( $no \beta = 1$ )							
Rho-square=0.0401							
Rho-square-bar=0.0397							
Akaike Information Criterion=18372.09							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_A$	0.693315	0.100657	6.887919	5.66E-12	0.098232	7.057907	1.69E-12
$\beta_B$	-2.03155	0.162712	-12.4855	0	0.155277	-13.0834	0
$\beta_l$	-1.67777	0.09191	-18.2544	0	0.090608	-18.5167	0
$\beta_v$	0.00605	0.002152	2.810807	0.004942	0.002224	2.719669	0.006535

Per quanto riguarda, invece, i modelli in cui la velocità è stata sostituita con il tempo di percorrenza, sono stati ottenuti i risultati riportati nella tabella seguente. Anche in questo caso, i modelli sono stati calibrati sia imponendo almeno un coefficiente pari a 1, sia non assegnando valore 1 ad alcun coefficiente.

Tabella 18 – Risultati della calibrazione dei modelli 4-5-6

Modello 4)							
Risultati calibrazione ( $\beta_l = 1$ )							
Rho-square=0.0184							
Rho-square-bar=0.0182							
Akaike Information Criterion=18890.25							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_t$	-0.00036	1.99E-05	-17.8878	0	2.03E-05	-17.5259	0
Risultati calibrazione ( $no \beta = 1$ )							
Rho-square=0.0228							
Rho-square-bar=0.0226							
Akaike Information Criterion=18806.71							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_t$	-0.00026	2.27E-05	-11.3097	0	2.46E-05	-10.4287	0
$\beta_l$	-3.2E-05	3.65E-06	-8.81519	0	4.03E-06	-7.98252	1.33E-15

<sup>2</sup> Il modello 3 è stato calibrato su un numero minore di viaggi (9 832). È stata, infatti, esclusa dalle analisi la coppia O-D 392-345 poiché una delle alternative di viaggio è costituita solo da strade di categoria A e B e non da strade secondarie.

**Modello 5)**

Risultati calibrazione ( $\beta_l = 1$ )							
Rho-square=0.0458							
Rho-square-bar=0.0456							
Akaike Information Criterion=18364.22							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_t$	-0.00022	2.15E-05	-10.1445	0	2.31E-05	-9.42738	0
$\beta_s$	-0.05637	0.003237	-17.4124	0	0.003862	-14.5938	0
Risultati calibrazione ( $no \beta = 1$ )							
Rho-square=0.0452							
Rho-square-bar=0.0448							
Akaike Information Criterion=18378.40							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_t$	-6.6E-05	2.47E-05	-2.67318	0.007514	2.57E-05	-2.56907	0.010197
$\beta_l$	-4.8E-05	4.01E-06	-11.9174	0	4.3E-06	-11.1118	0
$\beta_s$	-0.06333	0.003239	-19.5509	0	0.0036	-17.5897	0

**Modello 6)<sup>3</sup>**

Risultati calibrazione ( $\beta_l = 1$ )							
Rho-square=0.0344							
Rho-square-bar=0.0341							
Akaike Information Criterion=18478.44							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_t$	-0.0003	1.98E-05	-15.2536	0	2E-05	-15.1546	0
$\beta_A$	0.463622	0.097577	4.751351	2.02E-06	0.093879	4.938485	7.87E-07
$\beta_B$	-2.10596	0.159147	-13.2328	0	0.149464	-14.0901	0
Risultati calibrazione ( $no \beta = 1$ )							
Rho-square=0.0409							
Rho-square-bar=0.0405							
Akaike Information Criterion=18357.08							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_t$	-0.00017	2.35E-05	-7.13411	9.74E-13	2.46E-05	-6.80601	1E-11
$\beta_A$	0.67699	0.10022	6.755034	1.43E-11	0.096691	7.001554	2.53E-12
$\beta_B$	-2.12031	0.161317	-13.1437	0	0.152471	-13.9063	0
$\beta_l$	-4E-05	3.77E-06	-10.4878	0	3.92E-06	-10.1084	0

Sono, invece, state scartate le seguenti forme funzionali, a causa dei risultati ottenuti:

$$7. V = \beta_i * n_i + \beta_l * l + \beta_v * v$$

$$8. V = \beta_A * \%_A + \beta_B * \%_B + \beta_s * n_s + \beta_i * n_i + \beta_l * l + \beta_v * v$$

$$9. V = \beta_A * \%_A + \beta_B * \%_B + \beta_s * n_s + \beta_l * l + \beta_v * v$$

$$10. V = \beta_A * \%_A + \beta_B * \%_B + \beta_i * n_i + \beta_l * l + \beta_v * v$$

$$11. V = \beta_A * \%_A + \beta_B * \%_B + \beta_s * n_s + \beta_i * n_i + \beta_l * l + \beta_t * t$$

$$12. V = \beta_A * \%_A + \beta_B * \%_B + \beta_s * n_s + \beta_l * l + \beta_t * t$$

In cui  $n_i$  rappresenta il numero di intersezioni presenti lungo il percorso.

Nella tabella seguente sono riportati i risultati ottenuti in corrispondenza delle funzioni di utilità sopra elencate (7-12).

Come si può notare:

<sup>3</sup> Si veda nota n.2.



- in alcuni casi (evidenziati in rosso) non risultano soddisfatti i test informali sui coefficienti ( $\beta_l > 0$ ,  $\beta_v < 0$ , etc.):
- in altri il t-test ( $|t| < 1.96$ ).

Tabella 19 – Risultati della calibrazione dei modelli 7-8-9-10-11-12

<b>Modello 7)</b>							
Risultati calibrazione ( $\beta_v = 1$ )							
Rho-square=0.0668							
Rho-square-bar=0.0666							
Akaike Information Criterion=17959.32							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_i$	-0.01341	0.000504	-26.6223	0	0.000546	-24.5463	0
$\beta_l$	0.483943	0.11712	4.132038	3.6E-05	0.121229	3.991983	6.55E-05
<b>Modello 8)<sup>4</sup></b>							
Risultati calibrazione ( $\beta_v = 1$ )							
Rho-square=0.0725							
Rho-square-bar=0.072							
Akaike Information Criterion=17754.62							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_A$	0.215017	0.106077	2.02699	0.042663	0.101714	2.113947	0.03452
$\beta_B$	-1.25988	0.168617	-7.47186	7.9E-14	0.161877	-7.78296	7.11E-15
$\beta_i$	-0.00993	0.000583	-17.0186	0	0.000606	-16.3872	0
$\beta_l$	0.047245	0.126141	0.374538	0.708004	0.130283	0.36263	0.716881
$\beta_s$	-0.02177	0.003567	-6.10221	1.05E-09	0.003503	-6.21309	5.2E-10
<b>Modello 9)<sup>5</sup></b>							
Risultati calibrazione ( $\beta_v = 1$ )							
Rho-square=0.0556							
Rho-square-bar=0.0552							
Akaike Information Criterion=18075.19							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_A$	0.290504	0.102996	2.820542	0.004794	0.099112	2.931062	0.003378
$\beta_B$	-1.9194	0.163331	-11.7516	0	0.155963	-12.3068	0
$\beta_l$	-1.46466	0.09392	-15.5948	0	0.09683	-15.1261	0
$\beta_s$	-0.05262	0.003202	-16.4322	0	0.003502	-15.0259	0
Risultati calibrazione (no $\beta = 1$ )							
Rho-square=0.0558							
Rho-square-bar=0.0552							
Akaike Information Criterion=18074.89							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_A$	0.314688	0.104302	3.017078	0.002552	0.101359	3.104696	0.001905
$\beta_B$	-1.93974	0.163902	-11.8348	0	0.156293	-12.4109	0
$\beta_l$	-1.42992	0.096729	-14.7827	0	0.100246	-14.2641	0
$\beta_s$	-0.05371	0.003282	-16.3671	0	0.003602	-14.9127	0
$\beta_v$	-0.00343	0.002262	-1.5146	0.129875	0.00234	-1.46428	0.143118
<b>Modello 10)<sup>6</sup></b>							
Risultati calibrazione ( $\beta_v = 1$ )							
Rho-square=0.0705							
Rho-square-bar=0.0701							
Akaike Information Criterion=17790.44							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_A$	0.350485	0.10366	3.381093	0.000722	0.099294	3.529756	0.000416
$\beta_B$	-1.18579	0.168077	-7.05502	1.73E-12	0.16232	-7.30527	2.77E-13
$\beta_i$	-0.01154	0.000525	-21.9997	0	0.000557	-20.7225	0
$\beta_l$	0.269657	0.120329	2.24101	0.025025	0.124138	2.172245	0.029837

<sup>4</sup> Si veda nota n.2.

<sup>5</sup> Si veda nota n.2.

<sup>6</sup> Si veda nota n.2.

**Modello 11)<sup>7</sup>**

Risultati calibrazione ( $\beta_l = 1$ )							
Rho-square=0.0729							
Rho-square-bar=0.0723							
Akaike Information Criterion=17747.58							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_t$	6.72E-05	2.5E-05	2.682777	0.007301	2.52E-05	2.664677	0.007706
$\beta_A$	0.230385	0.105776	2.178055	0.029402	0.101597	2.267633	0.023352
$\beta_B$	-1.26065	0.167898	-7.50842	6E-14	0.16079	-7.84033	4.44E-15
$\beta_i$	-0.01038	0.000508	-20.4212	0	0.00053	-19.6043	0
$\beta_s$	-0.02284	0.003428	-6.66377	2.67E-11	0.003406	-6.70777	1.98E-11

**Modello 12)<sup>8</sup>**

Risultati calibrazione ( $\beta_l = 1$ )							
Rho-square=0.0467							
Rho-square-bar=0.0463							
Akaike Information Criterion=18245.34							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_t$	-0.0002	2.14E-05	-9.14108	0	2.29E-05	-8.54558	0
$\beta_A$	0.12487	0.101686	1.227994	0.219449	0.098243	1.271031	0.203717
$\beta_B$	-1.99621	0.160268	-12.4555	0	0.150521	-13.262	0
$\beta_s$	-0.04874	0.003362	-14.4955	0	0.003897	-12.5067	0

Tornando ai modelli 1-5, dalla Tabella 17 e dalla Tabella 18 si può notare come i test formali e informali sui risultati dei valori dei coefficienti ottenuti risultino soddisfatti.

Tuttavia, il  $\rho^2$  ottenuto è molto basso.

A tal proposito, sono stati meglio analizzati i risultati ottenuti per il modello 2), caratterizzato da uno dei valori di  $\rho^2$  maggiori ottenuti (0.046). Come si può notare dalla Tabella 17, la verifica della “ragionevolezza” dei segni dei coefficienti è soddisfatta: il coefficiente relativo alla lunghezza del percorso è di segno negativo e lo stesso vale per il coefficiente relativo al numero di semafori che il guidatore incontra su ciascun percorso.

Per quanto riguarda, invece, i test formali eseguiti sui risultati ottenuti è stato ottenuto quanto segue.

Il t-test – finalizzato a rigettare l’ipotesi che i valori stimati dei coefficienti siano statisticamente uguali a zero – risulta soddisfatto per entrambi i coefficienti.

<sup>7</sup> Si veda nota n.2.

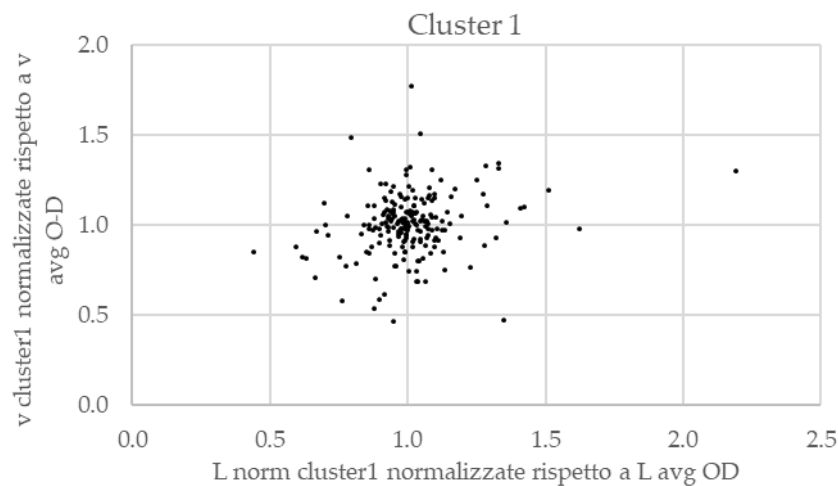
<sup>8</sup> Si veda nota n.2.

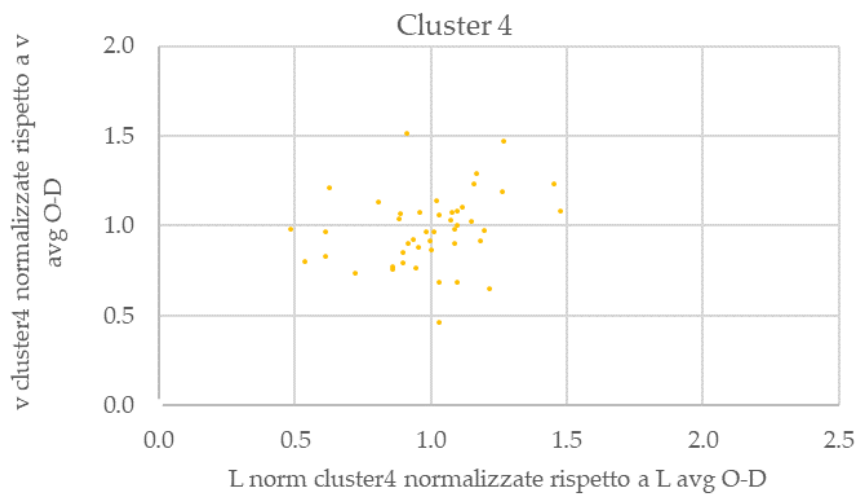
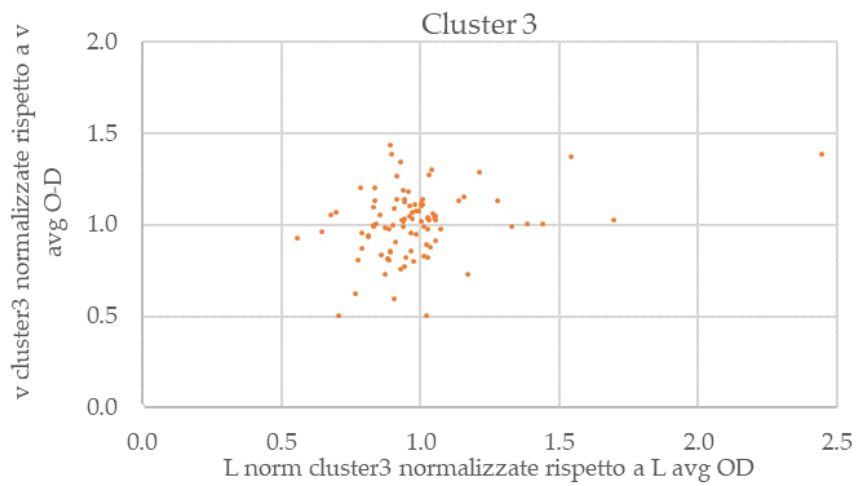
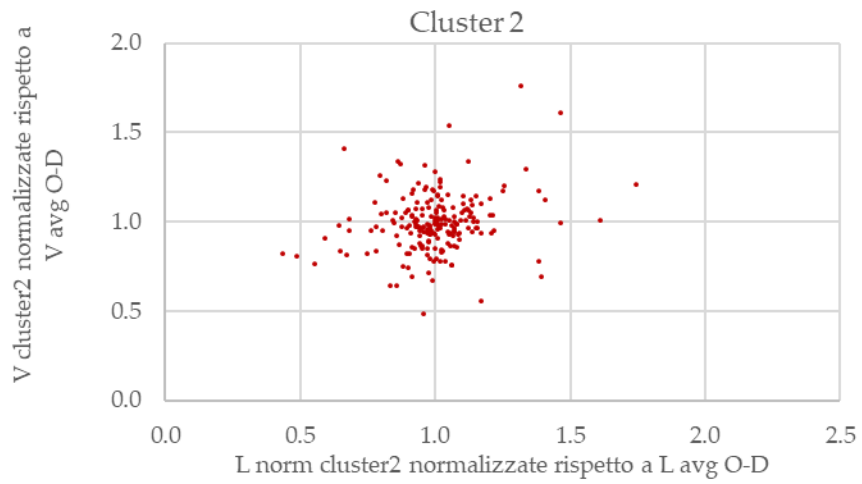
In entrambi i casi, infatti,  $|t| > 1.96$ , ovvero l'ipotesi può essere rigettata con una probabilità di errore del 5%, e quindi il valore stimato è considerato significativo.

Poiché il  $\rho^2$  ottenuto indica una bassa capacità riproduttiva del modello, sono stati analizzati nel dettaglio i risultati ottenuti calcolando - per ciascun viaggio - la probabilità di scelta di ciascun cluster a disposizione e analizzando il numero di casi in cui l'utente ha effettivamente selezionato il cluster a cui corrisponde massima probabilità di scelta: solo in 4186 casi sui 9882 totali, ovvero nel 42% dei casi, l'utente seleziona il cluster di probabilità massima.

Sono state, a questo punto, calcolate per ciascuna coppia OD e per ciascun cluster le lunghezze e le velocità, normalizzate rispetto al valore medio della coppia O-D.

I risultati sono stati poi rappresentati graficamente al fine di identificare eventuali casi particolarmente anomali ("outliers") in termini di lunghezze e velocità dei percorsi.





*Figura 34 – Rappresentazione di velocità e lunghezza medie dei percorsi di ciascuna O-D normalizzate rispetto al valore medio sulla relazione*

Calibrando il modello soltanto sulle coppie O-D per cui non vi sono valori della lunghezza e della velocità normalizzate minori di 0.75 o maggiori di 1.25 (per un totale di 6567 viaggi), sono stati ottenuti i risultati riportati nella tabella seguente.

*Tabella 20 – Risultati della calibrazione del modello 2, eliminando le anomalie*

Modello 4)							
Risultati calibrazione ( $\beta_v = 1$ )							
Rho-square=0.0557							
Rho-square-bar=0.0553							
Akaike Information Criterion=10740.67							
	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
$\beta_l$	-2.20166	0.158063	-13.9291	0	0.165726	-13.2849	0
$\beta_s$	-0.09581	0.005325	-17.994	0	0.005574	-17.1902	0

Calcolando, per ciascun viaggio analizzato, la probabilità di scelta di ciascun cluster a disposizione, è stata ottenuta una percentuale di casi in cui l'utente seleziona il cluster a cui corrisponde massima probabilità di scelta pari al 45%, invece che del 42% come nel caso precedente.

Il valore del  $\rho^2$  ottenuto è aumentato, ma ancora basso.

Per tale ragione, è stata effettuata una analisi finalizzata a verificare la possibilità che il valore del  $\rho^2$  ottenuto sia dovuto al fatto che gli utenti percepiscano come indifferenti le alternative di viaggio. L'ipotesi appare preliminarmente confermata, poiché calcolando – sempre per il modello 2) - la differenza di utilità sistematica tra il percorso dalla massima probabilità di scelta e il percorso effettivamente scelto si ottiene un valore medio su tutte le O-D del 14% e una deviazione standard rispetto alla media del 19%.

Si tratterebbe, in sostanza una conferma del I principio di Wardrop: gli utenti in viaggio su una stessa coppia O-D stanno effettuando la selezione del percorso rispetto ad una serie di alternative di costo minimo e, se non uguale, almeno simile.

Il modello probabilistico, in questo caso, non risulterebbe adatto a prevedere le scelte effettuate dagli utenti, ma potrebbe essere validamente utilizzato per stimare le probabilità di scelta delle diverse alternative.

Approfondendo questo aspetto, sono stati ordinati i cluster di ciascuna coppia O-D per frequenza di scelta, ovvero considerando come cluster 1 il cluster selezionato più volte dagli utenti in viaggio sulla relazione, come cluster 2 il secondo cluster selezionato più volte e così via e associando poi a ciascun cluster la probabilità di scelta calcolata dal modello.

Per ciascun cluster delle 201 coppie O-D in analisi, è stata calcolata la differenza assoluta tra le probabilità di scelta del percorso e la corrispondente frequenza relativa di scelta.

A questo punto sono stati calcolati, per cluster e sul totale delle coppie O-D:

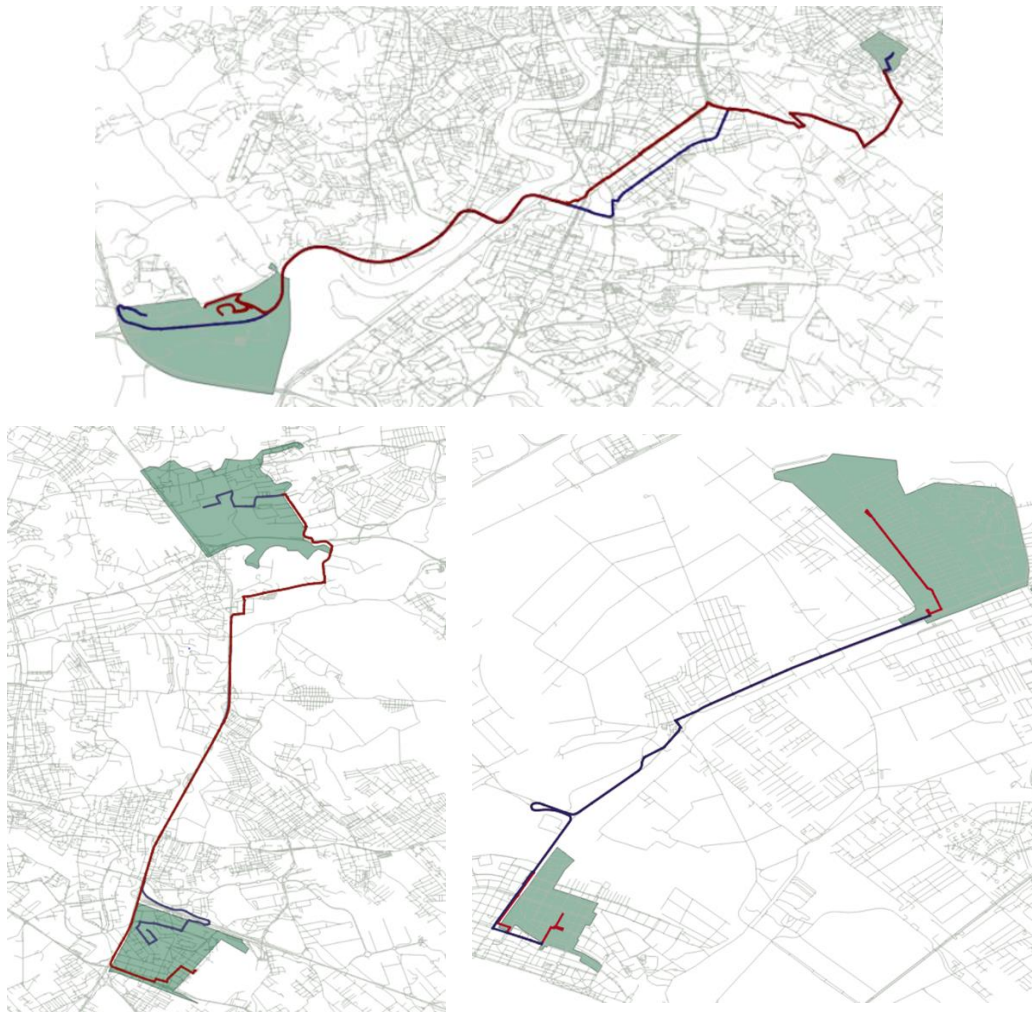
- il valore medio dell'indicatore, ovvero la media sulle 201 coppie O-D in analisi delle differenze assolute tra le probabilità di scelta dei percorsi disponibili e le corrispondenti frequenze relative di scelta in modulo;
- la deviazione standard delle differenze assolute tra le probabilità di scelta dei percorsi e le corrispondenti frequenze di scelta rispetto al valore medio.

*Tabella 21 – Analisi delle differenze assolute tra le probabilità di scelta dei percorsi e le corrispondenti frequenze relative di scelta*

	Cluster 1 (più scelto)	Cluster 2	Cluster 3	Cluster 4 (meno scelto)	Totale
Valore medio	37.22%	30.44%	25.94%	30.26%	32.17%
Dev. Standard	25.77%	25.60%	31.27%	34.68%	27.88%

Analizzando alcune delle coppie O-D caratterizzate da valori elevati della differenza tra probabilità e frequenza di scelta (valore massimo sui quattro cluster superiore allo 0.5%), è emerso come tali relazioni siano caratterizzate da alternative di viaggio che differiscono sostanzialmente

per la porzione di percorso interna alla zona di origine o destinazione (in questi casi generalmente piuttosto estese).



*Figura 35 – Esempi di alternative di viaggio che differiscono sostanzialmente per la porzione di percorso interna alla zona di origine o destinazione*

Un'ultima verifica è stata effettuata per tenere in considerazione l'attendibilità della stima dei tempi di percorrenza: a tal proposito, è stato confrontato l'errore delle probabilità di scelta (per ciascun cluster) con l'ampiezza dell'intervallo di confidenza della stima del tempo di percorrenza del percorso (ottenuto come rapporto tra lo scarto standard e la radice quadrata del numero di osservazioni).

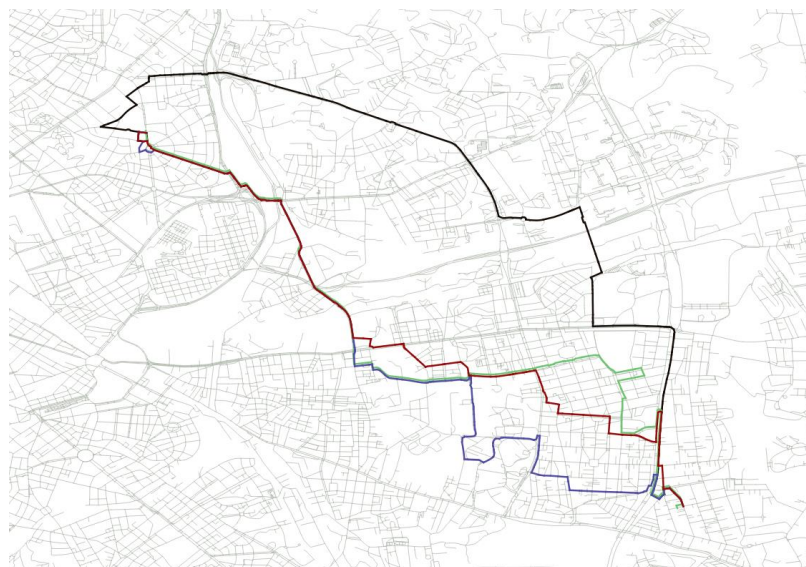
Anche in questo caso i valori delle differenze assolute ottenuti sono significativi (valore medio su tutte le O-D pari a circa 30%).

È importante sottolineare, tuttavia, come questi risultati siano affetti da un margine di errore dovuto al fatto che le osservazioni dei tempi di viaggio non sono simultanee e ad eventuali errori di misura oltre che a ulteriori variabili di scelta non osservabili.

Infine, l'analisi condotta offre numerosi spunti di ricerca da approfondire, tra i quali ad esempio:

- la percezione delle grandezze e la previsione del tempo di percorrenza da parte dell'utente, subito prima di effettuare lo spostamento;
- i casi in cui l'utente sembra selezionare il percorso in funzione della necessità di effettuare deviazioni (nel caso di "viaggi a catena", ad esempio) piuttosto che sulla base dei costi associati alle alternative di viaggio.

Nella figura seguente, ad esempio, il percorso evidenziato in nero è significativamente diverso dalle altre alternative di viaggio e potrebbe essere stato selezionato sulla base della necessità di effettuare una sosta in un POI ad esempio, piuttosto che sulla base della lunghezza o della velocità.



*Figura 36 – Rappresentazione esemplificativa della casistica in cui sulla stessa coppia O-D, è presente una alternativa di viaggio significativamente diversa*



## 5 Conclusioni

Nella presente tesi di dottorato è definita una metodologia per ricostruire i percorsi seguiti dagli utenti stradali a partire da una serie di dati posizionali sparsi (FCD) e analizzare i comportamenti di viaggio dei guidatori rispetto alla scelta del percorso e dell'orario di partenza. Sono, inoltre, presentati i risultati ottenuti dall'applicazione di tale metodologia a un set di FCD registrati nella città di Roma nel 2015.

L'analisi preliminare di tali dati – svincolata dall'informazione del grafo stradale – non ha evidenziato un aumento significativo della percentuale di utenti che tende ad anticipare la partenza del viaggio successivo all'aumentare del tempo di viaggio sperimentato nel viaggio precedente. Ovvero non ha reso immediatamente identificabile la c.d. “banda di indifferenza” rispetto a miglioramenti o peggioramenti del tempo di percorrenza, superata la quale, gli utenti sono spinti ad anticipare l'orario di partenza.

Questo potrebbe essere dovuto all'entità modesta degli incrementi di durata dei tempi di viaggio sperimentati dai guidatori (nella metà dei casi circa, si tratta di variazioni entro i 5 minuti rispetto alla durata media del viaggio di ciascun utente).

Al fine di analizzare le variazioni apportate dagli utenti all'orario di partenza anche tenendo conto del percorso selezionato, è stata appunto applicata ai FCD una metodologia (descritta nel dettaglio nel capitolo 3), che prevede:

- la ricostruzione del percorso seguito dall'utente stradale a partire dalle posizioni sparse registrate (FCD);
- la generazione di un set ristretto di alternative di viaggio per ciascuna coppia O-D, basata su una analisi di “dissimilarità” dei percorsi selezionati dagli utenti stradali.

I risultati hanno evidenziato l'efficacia della procedura di clusterizzazione dei percorsi, nel suddividere alternative differenti e aggregare itinerari simili (sono stati, infatti, ottenuti un valore medio di "dissimilarità interna" pari al 16% e un valore medio di "dissimilarità esterna" del 50%), in particolare per analizzare le reti di grandi dimensioni e ad elevata complessità.

Per quanto riguarda la distribuzione dei tempi di viaggio di percorsi differenti, circa il 75% dei percorsi scelti dagli utenti presenta tempi di viaggio che superano il valore minimo di meno del 35%, rispetto ad uno scarto standard dei tempi di percorrenza nel periodo di osservazione dell'ordine del 24% su tutte le coppie O-D e del 19% su tutti i cluster. D'altro canto, circa l'80% dei percorsi differisce dal valore medio di meno del 25% e solo il 5% di più del 50%.

Limitando l'analisi ai tempi di viaggio medi dei cluster selezionati più volte dagli utenti, il 90% dei tempi di percorrenza differisce dal percorso medio di meno del 10%.

Una volta calcolate le differenze dei tempi di percorrenza dei percorsi selezionati, è stata effettuata una analisi della dinamica giornaliera delle scelte in funzione delle prestazioni della rete, i cui risultati hanno evidenziato come il 35% degli utenti stradali confermi la scelta del percorso abituale (ovvero tenda ad utilizzare sempre un percorso appartenente allo stesso cluster), mentre il 64% circa dei conducenti verifici anche altri percorsi durante il periodo di osservazione.

Di questi ultimi, tuttavia, solo il 21% degli utenti sceglie più di un percorso alternativo rispetto al solito, mentre il 43% un solo percorso alternativo, che seleziona più volte (sembra avere, pertanto, due percorsi "abituati").

Calcolando il delta medio tra il tempo di percorrenza minimo e medio dei percorsi utilizzati dallo stesso utente, si ottiene un valore pari al 26% (e corrispondente a 9 minuti).

È stata, poi approfondita la propensione degli utenti stradali al cambio di percorso o al cambio di orario di partenza a seguito di un incremento o una riduzione del tempo di viaggio sperimentato.

Mentre nel caso del cambio di orario di partenza è stato ottenuto un sostanziale equilibrio tra gli utenti che decidono di anticipare e quelli che ritardano l'orario di partenza, nel caso del cambio di percorso è stata evidenziata una tendenza spiccata degli utenti a confermare il percorso abituale.

Tuttavia, sebbene incrementi o riduzioni dei tempi di viaggio sperimentati superiori ai 20 minuti non influiscano sul meccanismo di cambio del percorso, è invece emersa una correlazione tra la percentuale di utenti che cambia il percorso e incrementi o riduzioni dei tempi di viaggio compresi tra i 5 e i 20 minuti ("banda di indifferenza" contenuta entro i 5 minuti).

Tra gli utenti che decidono di cambiare percorso, infine, il 69% cambia in favore di un percorso dal tempo di percorrenza minore (in media di meno del 15%), oppure da un numero di semafori minore, oppure ancora da un percorso dal tempo di percorrenza minore e caratterizzato da un numero di semafori minore.

Sulla base dei risultati ottenuti, è stata approfondita la possibilità di modellare i comportamenti di scelta del percorso mediante modelli di scelta discreta.

Sono state, a tal fine, calibrate diverse forme funzionali di utilità sistematica di un modello logit di scelta del percorso. Gli attributi di scelta analizzati sono i seguenti:

- $l$ , lunghezza del percorso normalizzata rispetto al valore massimo di lunghezza tra le alternative di viaggio a disposizione su una specifica coppia O-D;
- $v$ , velocità del percorso in km/h;
- $n_s$ , numero di semafori incontrati dall'utente lungo il percorso;
- $\%_{categ.A}$ , percentuale - rispetto al totale degli archi che compongono ciascun percorso - di strade di maggiore importanza (autostrade, strade statali e strade urbane di scorrimento, etc.);
- $\%_{categ.B}$ , percentuale di strade di minore importanza (strade locali) - rispetto al totale degli archi che compongono ciascun percorso;
- $t$ , tempo medio di percorrenza del percorso in secondi;
- $n_i$ , numero di intersezioni presenti lungo il percorso.

Le stime dei coefficienti ottenute sono statisticamente significative: in particolare, i risultati hanno confermato che lunghezza, tempo di viaggio e numero di semafori rappresentano tre attributi che con probabilità del 95% incidono sulla scelta del percorso da parte degli utenti.

I valori di  $\rho^2$  ottenuti possono essere spiegati assumendo che gli utenti percepiscano i percorsi alternativi come indifferenti (ulteriore conferma del principio di Wardrop: gli utenti in viaggio su una stessa coppia O-D effettuano la selezione del percorso rispetto ad una serie di alternative di costo minimo e, se non uguale, almeno simile). Infatti, calcolando la differenza di utilità sistematica tra il percorso dalla massima probabilità di scelta e il percorso effettivamente scelto si ottiene un valore medio su tutte le O-D del 14% e una deviazione standard rispetto alla media del 19%.

Inoltre, dalle analisi sulla variabilità dei tempi viaggio per cluster e per O-D (si veda Tabella 10) è emerso come i tempi di percorrenza di cluster appartenenti alla stessa coppia O-D risultino piuttosto simili tra loro.

Questo risultato non è scontato, in quanto l'operazione di clusterizzazione delle alternative di viaggio non è effettuata sulla base dei tempi di percorrenza, ma solo sulla lunghezza dei percorsi.

Anche i risultati della analisi della dinamica giornaliera appaiono coerenti con quanto sopra:

- da un lato vi sono utenti che sembrano seguire prevalentemente percorsi "abituati";
- dall'altro utenti che cambiano percorso e non registrano un miglioramento in termini di tempo di percorrenza.

Tuttavia, ulteriori analisi potrebbero essere utili per indagare i casi in cui l'utente sembra selezionare il percorso in funzione della necessità di effettuare deviazioni (nel caso di "viaggi a catena", ad esempio) piuttosto che sulla base dei costi associati alle alternative, o ancora per approfondire la percezione delle grandezze e la previsione del tempo di percorrenza da parte dell'utente, subito prima dello spostamento.

## Bibliografia

- Ajmar, A., Arco, E., Boccardo, P., Perez, F. (2019). Floating Car Data (FCD) for mobility applications. *International archives of the photogrammetry, remote sensing and spatial information sciences*, 42(2), 1517-1523.
- Akgün, V., Erkut, E., Batta, R. (2000). On finding dissimilar paths. *European Journal of Operational Research*, 121(2), 232-246.
- Arentze, T., Timmermans, H. (2003). Modelling learning and adaptation processes in activity-travel choice. *Transportation*, 30, 37–62.
- Astarita, V., Giofrè, V. P., Guido, G., Vitale, A. (2017). The Use of Adaptive Traffic Signal Systems Based on Floating Car Data. *Wireless Communications and Mobile Computing*.
- Avineri, E., Prashker, J.N. (2005). Sensitivity to travel time variability: travelers' learning perspective. *Transportation Research Part C: Emerging Technologies*, 13, 157-183.
- Batista, S. F. A., Zhao, C.-L., Leclercq, L. (2018). Effects of users' bounded rationality on a traffic network performance: a simulation study. *Journal of Advanced Transportation*, 1-20.
- Bekhor, S. B.-A. (2002). Adaptation of logit kernel to route choice situation. *Transportation Research Record*, 1805(1), 78-85.
- Bierlaire, M. (2020). A short introduction to PandasBiogeme. *Technical report TRANSP-OR 200605. Transport and Mobility Laboratory, ENAC, EPFL*.
- Bierlaire, M., Chen, J., Newman, J. (2013). A probabilistic map matching method for smartphone GPS data. *Transport. Research Part C: Emerging Technol.*, 26, 78–98.
- Bogers, E. A. (2005). Joint Modeling of Advanced Travel Information Service, Habit, and Learning Impacts on Route Choice by Laboratory Simulator Experiments. *Transportation Research Record*, 1926(1), 189-197.
- Cascetta, E., Nuzzolo, A., Russo, F., Vitetta, A. (1996). *A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks*. Lyon, France: Transportation and traffic theory: Proceedings of the 13th International Symposium on Transportation and Traffic Theory. Pergamon Press.

- Cascetta, E., Papola, A. (2001). Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. *Transportation Research Part C*, 9(4), 249-263.
- Chen, C., Ma, J., Susilo, Y., Liu, Y., Wang, M. (2016). The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation Research Part C: Emerging Technologies*, 68, 285-299.
- Chen, F., Shen, M., Tang, Y. (2011). Local path searching based map matching algorithm for floating car data. *Procedia Environmental Sciences*, 10, 576-582 .
- Ciscal-Terry, W., Dell'Amico, M., Hadjidimitriou, N.S., Iori, M. (2016). An analysis of drivers route choice behaviour using GPS data and optimal alternatives. *J. Transp. Geogr.*, 51, 119-129.
- Croce, A.I., Musolino, G., Rindone, C., Vitetta, A. (2009). Transport system and Big Data: Zoning and Graph building with traditional surveys, FCD and GIS. *ISPRS Int. J. Geo-Inf.*, 8, 187.
- De Fabritiis, C., Ragona, R., Valenti, G. (2008). Traffic estimation and prediction based on real time floating car data. *2008 11th International IEEE Conference on Intelligent Transportation Systems, Beijing*, 197-203.
- Dell'Olmo, P., Gentili, M., Scozzari, A. (2005). On finding dissimilar Pareto-optimal paths. *European Journal of Operations Research*, 162, 70-82.
- Di, X., Liu, H.X. (2016). Boundedly rational route choice behavior: a review of models and methodologies. *Transp. Res. B*, 85, 142-179.
- Di, X., Liu, H.X., Zhu, S., Levinson, D.M. (2017). Indifference bands for boundedly rational route switching. *Transportation*, 44, 1169-1194.
- Diependaele, K., Riguelle, F., Temmerman, P. (2016). Speed behavior indicators based on floating car data: results of a pilot study in Belgium. *Transp Res Proc*, 14, 2074-2082.
- Eikenbroek, O.A., Still, G.J., van Berkum, E.C., Kern, W. (2018). The boundedly rational user equilibrium: a parametric analysis with application to the network design problem. *Transp. Res. Part B*, 107, 1-17.
- Frejinger, E., Bierlaire, M. (2007). Capturing correlation with subnetworks in route choice models. *Transportation Research Part B*, 41(3), 363-378.
- Fusco, G., Bracci, A., Caligiuri, T., Colombaroni, C., Isaenko, N. (2018). Experimental analyses and clustering of travel choice behaviours by floating

- car big data in a large urban area. *IET Intelligent Transport Systems*, 12 (4), 270–278.
- Fusco, G., Colombaroni, C., Comelli, L., Isaenko, N. (2015). Short term traffic predictions on large urban traffic networks: application of network-based machine learning models and dynamic traffic assignment models. *Int. Conf. Models Techn. Intell. Transp. Syst. (MT-ITS)*, 93-101.
- Fusco, G., Colombaroni, C., Comelli, L., Isaenko, N. (2015). Short-term traffic predictions on large urban traffic networks: applications of network-based machine learning models and dynamic traffic assignment models. *2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 93-101.
- Fusco, G., Colombaroni, C., Isaenko, N. (2016). Comparative analysis of implicit models for real time short-term prediction. *IET Intelligent Transport Systems*, 10(4), 270–278.
- Fusco, G., Colombaroni, C., Isaenko, N. (2016). Short-term speed predictions exploiting big data on large urban road networks. *Transportation Research Part C: Emerging Technologies*, 73, 183-201.
- Gigerenzer, G., Selten, R. Editors. (2002). *Bounded Rationality – The adaptive toolbox*. Cambridge: MIT Press.
- Han, K., Szeto, W.Y., Friesz, T.L. (2015). Formulation, existence, and computation of boundedly rational dynamic user equilibrium with fixed or endogenous user tolerance. *Transportation Research Part B*, 79, 16-49.
- Han, Q., Timmermans, H. (2006). Interactive Learning in Transportation Networks with Uncertainty. *Transportation Research Record*, 1964(1), 27-34.
- Hato, E., Asakura, Y. (2000). Incorporating bounded rationality concept into route choice model for transportation network analysis. *PTRC-Publications*, 1-12.
- Isaenko, N., Colombaroni, C., Fusco, G. (2017). Traffic dynamics estimation by using raw floating car data. *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 704-709.
- Jones, B. D. (1999). Bounded Rationality. *Annu. Rev. Polit. Sci.*, 2, 297-321.



- Jou, R., Lam, S., Liu, Y., Chen, K. (2005). Route switching behavior on freeways with the provision of different types of real-time traffic information. *Transportation Research Part A*, 39, 445-461.
- Kerner, B.S., Demir, C., Herrtwich, R.G., Klenov, S.L., Rehborn, H., Aleksic, M., Haug, A. (2005). Traffic State Detection with Floating Car Data in Road Networks. *Proceedings. 2005 IEEE Intelligent Transportation Systems*, 44-49.
- Kim, J., Mahmassani, H.S. (2015). Spatial and temporal characterization of travel patterns in a traffic network using vehicle trajectories. *Trans. Res. C*, 59, 375–390.
- Li, H., Guensler, R., Ogle, J., Wang, J. (2004). Using Global Positioning System Data to Understand Day-to-Day Dynamics of Morning Commute Behavior. *Transportation Research Record*, 1895(1), 78-84.
- Li, J., Xie, L., Lai, X. (2013). Route reconstruction from floating car data with low sampling rate based on feature matching. *Research Journal of Applied Sciences, Engineering and Technology*, 6(12), 2153-8.
- Li, Q., Zhang, T., Wang, H., Zeng, Z. (2011). Dynamic accessibility mapping using floating car data: a network constrained density estimation approach. *Journal of Transport Geography*, 19(3), 379-393.
- Li, Q., Zhang, T., Yu, Y. (2011). Using cloud computing to process intensive floating car data for urban traffic surveillance. *Int. J. Geog. Inf. Sci.*, 25(8), 1303-1322.
- Lima, A., Stanojevic, R., Papagiannaki, D., Rodriguez, P., González, M. C. (2016). Understanding individual routing behaviour. *J. R. Soc. Interface*, 13, 20160021.
- Liu, X., Liu, K., Li, M., Lu, F. (2016). A ST-CRF map-matching method for lowfrequency floating car data. *IEEE Transactions on Intelligent Transportation Systems*, 18 (5), 1241-1254.
- Mahmassani, H.S., Liu, Y.-H. (1999). Dynamics of commuting decision behaviour under advanced traveller information systems. *Transportation Research C*, 7, 91-107.
- Mannini, L., Carrese, S., Cipriani, E., Crisalli, U. (2015). On the short-term prediction of traffic state: an application on urban freeways in Rome. *Transportation Research Procedia*, 10, 176-185.

- Martí, R., Velarde, J.L., Duarte, A. (2009). Heuristics for the biobjective path dissimilarity problem. *Computers and Operations Research*, 36, 2905-2912.
- Meneguzzer, C., Olivieri, A. (2013). Day-to-day traffic dynamics: laboratory-like experiment on route choice and route switching in a simple network with limited feedback information. *Procedia - Social and Behavioral Sciences*, 87, 44-59.
- Nigro, M., Cipriani, E., Del Giudice, A. (2018). Exploiting floating car data for time-dependent Origin – Destination matrices estimation. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 22, 159-174.
- Protschky, V., Ruhhammer, C., Feit, S. (2015). Learning Traffic Light Parameters with Floating Car Data. *2015 IEEE 18th International Conference on Intelligent Transportation Systems, Las Palmas*, 2438-2443.
- Rahmani, M. K. (2010). Requirements and Potential of GPS-based Floating Car Data for Traffic Management: Stockholm Case Study. *13th International IEEE Conference on Intelligent Transportation Systems*, 730-735.
- Rahmani, M., Jenelius, E., Koutsopoulos, H. N. (2013). Route Travel Time Estimation Using Low-Frequency Floating Car Data. *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 2292-2297.
- Rahmani, M., Jenelius, E., Koutsopoulos, H. N. (2015). Non-parametric estimation of route travel time distribution from low-frequency floating car data. *Transportation Research Part C*, 58, 343-362.
- Rahmani, M., Koutsopoulos, H. (2013). Path inference from sparse floating car data for urban networks. *Transportation Research Part C*, 30, 41-54.
- Rambaldi, S., Marchioni, M., Bazzani, A., Giorgini, B. (2012). Traffic Global analysis on the whole Italian Road network. *2012 Proceedings of the 35th International Convention MIPRO*, 1678-1682.
- Selten, R. (1999). What is bounded rationality? *Paper prepared for the Dahlem Conference*, SFB Discussion Paper B-454.
- Selten, R., Chamura, T., Pitz, T., Kube, S., Schreckenberg, M. (2007). Commuters route choice behaviour. *Games and Economic Behavior*, 58, 394-406.

- Shang, W., Han, K., Ochieng, W., Angeloudis, P. (2017). Agent-based day-to-day traffic network model with information percolation. *Transp. A Transp. Sci.*, 13, 1-29.
- Simon, HA. (1972). "Theories of Bounded Rationality" in "Decision and Organization: A Volume in Honor of Jacob Marschak". Amsterdam: North Holland: Ed. C. B. McGuire, and Roy Radner, 161–76.
- Simon, HA. (1979). Rational Decision Making in Business Organizations. *American Economic Review*, 69, 493-513.
- Song, G., Yu, L. (2009). Estimation of Fuel Efficiency of Road Traffic by Characterization of Vehicle-Specific Power and Speed Based on Floating Car Data. *Transportation Research Record*, 2139(1), 11-20.
- Stopher, P., Clifford, E., Montes, M. (2008). Variability of Travel over Multiple Days - Analysis of Three Panel Waves. *Transportation Research Record*, 2054(1), 56-63.
- Stopher, PR., Kockelman, K., Greaves, SP., Clifford, E. (2008). Reducing Burden and Sample Sizes in Multi-day Household Travel Surveys. *Transportation Research Record*, 2064(1), 12-18.
- Sun, C., Li, M., Cheng, L., Zhu, S., Chu., Z. (2016). Boundedly rational user equilibrium with restricted unused routes. *Discrete Dynamics in Nature and Society*, 10.1155/2016/9848916.
- Sun, D. Z.-R. (2014). Urban travel behavior analyses and route prediction based on floating car data. *Transportation Letters*, 6(3), 118-125.
- Tang, W.Y., Cheng, L. (2016). Analyzing multiday route choice behavior of commuters using GPS data. *Advances in Mechanical Engineering*, 8, 1-11.
- Tawfik, A. M., Rakha, H. A., Miller, S. D. (2010). An experimental Exploration of Route choice: identifying drivers choices and choice patterns and capturing network evolution. *13th International IEEE Conference on Intelligent Transportation Systems*, 1005-1012.
- Tawfik, A. M., Rakha, H. A., Miller, S. D. (2010). Driver Route Choice Behavior: Experiences, Perceptions, and Choices. *2010 IEEE Intelligent Vehicles Symposium, San Diego, CA*, 1195-1200.
- Tulic, M., Bauer, D., Scherrer, W. (2014). Link and Route Travel Time Prediction Including the Corresponding Reliability in an Urban Network

- Based on Taxi Floating Car Data. *Transportation Research Record*, 2442(1), 140-149.
- Vacca, A. (2015). Analisi comportamentale della scelta del percorso attraverso l'utilizzo di nuove tecnologie di acquisizione delle informazioni. *Tesi di dottorato*.
- Vacca, A., Prato, C.G., Meloni, I. (2019). Should I stay or should I go? Investigating route switching behavior from revealed preferences data. *Transportation*, 46, 75-93.
- Waller, M. A. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34, 77-84.
- Wang, D., Liao, F., Gao, Z., Timmermans, H. (2019). Tolerance-based strategies for extending the column generation algorithm to the bounded rational dynamic user equilibrium problem. *Transp. Res. Part B*, 119, 102-121.
- Wardrop, J. G. (1952). Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 325-362.
- Watling, D.P., Rasmussen, T.K., Prato, C.G., Nielsen, O.A. (2018). Stochastic user equilibrium with a bounded choice model. *Transp. Res. B Methodol.*, 114, 254-280.
- Xu, L., Yue, Y., Li, Q. (2013). Identifying Urban Traffic Congestion Pattern from Historical Floating Car Data. *Procedia-Social and Behavioral Sciences*, 96, 2084-2095.
- Yao, E. J. (2013). Taxi driver's route choice behavior analysis based on Floating Car Data. *Applied Mechanics and Materials*, 361-363, 2036-2039.
- Yong-chuan, Z., Xiao-qing, Z., Zhen-ting, C. (2011). Traffic Congestion Detection Based On GPS Floating-Car Data. *Procedia Engineering*, 15, 5541-5546.
- Yu, Y., Han, K., Ochieng, W. (2020). Day-to-day dynamic traffic assignment with imperfect information, bounded rationality and information sharing. *Transportation Research Part C: Emerging Technologies*, 114, 59-83.
- Zhang, X., Guan, H., Zhu, H., Zhu, J. (2019). Analysis of Travel Mode Choice Behavior Considering the Indifference Threshold. *Sustainability*, 11, 5495.

Zhu, S., Levinson, D. (2015). Do people use the shortest path? An empirical test of Wardrop's first principle. *91th annual meeting of the Transportation Research Board, Washington, 8.*