IRPET Istituto Regionale Programmazione Economica della Toscana

Grandi attrattori infrastrutturali e stima dei bacini di domanda degli scali aeroportuali

2014 2020 POR CreO PROGRAMMA OPERATIVO REGIONALE CRESCITA e OCCUPAZIONE
Fondo europeo di sviluppo regionale (Fesr)
Unione europea Fondo Europeo di Sviluppo Regionale
REPUBBLICA ITALIANA
Regione Toscana

Firenze, dicembre 2016

RICONOSCIMENTI

La ricerca è stata curata da Daniele Fadda (ISTII/CNR) con la collaborazione di Leonardo Piccini, coordinamento di Patrizia Lattarulo, dirigente dell'area Economia pubblica e territorio dell'IRPET.

Abstract

Il lavoro ha lo scopo di valutare l'impatto della presenza di grandi attrattori di mobilità in ambito urbano e la relazione funzionale che questi stabiliscono con i territori corcostanti in termini di flussi di mobilità, utilizzando metodi e fondi dati innovative di ambito Big Data. Come studio pilota è stata scelta l'analisi della scelta aeroportuale in una regione multi-aeroporto come la Toscana. Sono statu utilizzati due dataset di fonte GPS e GSM incrocoati con dati più tradizionali di tipo socioeconomico. La rilevanza del tema della programmazione dei servizi aeroportuali è testimoniata dall'attenzione ricevuta dalle strategie di sviluppo del sistema aeroportuale toscano negli ultimi anni. Gli strumenti di pianificazione nazionale del settore (masterplan ENAC) evidenziano la necessità di una gestione armonica dei due scali al fine di massimizzarne le sinergie ed evitare la sovrapposizione delle funzioni. Da ciò deriva la necessità di analizzare il tipo di rapporto che tali grandi attrattori instaurano con i territori circostanti.

The aim of the paper is to analyze the impact of the presence of big mobility generators within urban areas. As a case study we chose airport choice in a multi-airport region (specifically, Tuscany). We use two large sets of data (one containing GPS data, the other GSM data) cross-matched with more traditional socio-economic data. The relevance of airport catchment area analysis is testified by the attention reserved, in recent years, to the diversification and growth strategies implemented within the airport system of Tuscany. Even the national strategic documents (ENAC masterplan) acknowledges that, in future years, the two main regional airports (Florence and Pisa), need to carefully plan their offer to maximize their synergies through an integrated and non-overlapping specialization. From this situation comes the need for a detailed understanding of the relations that these two relevant infrastructures have established with their territories.

DANIELE FADDA

l'IMPATTO SULLA MOBILITÀ

dei

GRANDI POLI ATTRATTORI

UNO STUDIO SUGLI
AEROPORTI TOSCANI

SOMMARIO

L' IMPATTO SULLA MOBILITÀ DEI GRANDI POLI ATTRATTORI

1 **Introduzione** — 6

2 **Partner di progetto** — 8
- ISTI
- IRPET

3 **Dataset disponibili e integrazione dei dati** — 10
- Sorgenti dei dati di mobilità
- Sorgenti dati socio-demografiche
- Identificazione delle geometrie

4 **Una panoramica sugli aeroporti toscani** — 18
- l'aeroporto "Amerigo Vespucci" di Firenze
- l'aeroporto "Galileo Galilei" di Pisa

5 **L' impatto sulla mobilità generato dagli aeroporti internazionali della Toscana** — 24
- L'analisi delle soste: un'applicazione del modello di mobilità basato sulle traiettorie
- L'identificazione degli access point: un'applicazione del mobility data-mining per descrivere comportamenti collettivi di mobilità
- La mobilità dei residenti in Toscana: definizione della matrice OD: Residenza – Attrattore e il concetto di most frequent location

6 **Conclusioni e possibili sviluppi** — 34

7 **APPENDICI** — 38
- 1 - Strumenti analitici
- 2 - Planning e scheduling del progetto

8 **Riferimenti** — 39

3. DATASET DISPONIBILI E INTEGRAZIONE DEI DATI

Negli ultimi anni, lo studio della mobilità è cambiato notevolmente e le tradizionali fonti dati quali survey, censimenti e rilevamenti socio-demografici e statistici sono stati affiancati dai così detti Big Data. Questi sono dati di grandi dimensioni prodotti dall'utilizzo di dispositivi mobili o servizi da parte delle persone. I telefoni cellulari, i dispositivi di navigazione satellitare, Facebook o Twitter sono i principali esempi di sorgenti dati che hanno rivoluzionato lo studio empirico della mobilità. Ognuna di queste tipologie di dati permette di osservare fenomeni della mobilità a diverse granularità spaziali e temporali e abilita la creazione di diversi modelli analitici.
I dati utilizzati all'interno dello studio provengono da fonti eterogenee: da una parte i dati GPS forniti da Octo Telematics e i dati GSM forniti da WIND, dall'altra i dati di tipo socio-demografico, tra cui indagini campionarie, censimenti e più in generale il patrimonio informativo della regione Toscana. Mentre GPS e GSM sono stati messi a disposizione dal KDD Lab, gli altri sono patrimonio informativo e di studio per IRPET. A queste due tipologie di dati sono state affiancate una serie di sorgenti derivate principalmente da open data e API di servizi informativi disponibili online.
In questo paragrafo viene presentata una breve descrizione dei dati disponibili, con un focus specifico su quelli effettivamente utilizzati nel corso delle analisi e sugli step di data preparation ed understanding preliminari al processo di analisi.

3.1 Sorgenti dei dati di mobilità

Dati di posizionamento GPS (Global Positioning System)

I dati GPS consistono in una sequenza di rilevamenti spazio-temporali di veicoli sui quali è installato un dispositivo satellitare di posizionamento. Tale dispositivo si occupa di ricostruire la posizione geografica del veicolo in modo continuo, comunicandola regolarmente alla centrale che li raccoglie sotto forma di coordinate spazio-temporali (i punti GPS).

Fornitore: OctoTelematics (3) è una società leader a livello mondiale nel mercato dell'Insurance Telematics e si occupa del tracciamento di una flotta di veicoli tramite dispositivi GPS a bordo.

Granularità: Puntuale. Il dispositivo GPS rileva la posizione ogni 30 secondi e ha uno scarto massimo di 10 metri.

Copertura spaziale: Veicoli che hanno attraversato la regione Toscana nella finestra temporale considerata.

Finestra temporale: Fornitura 1: Maggio 2011. Fornitura 2: 26/01/2014 - 16/03/2014

Dati Disponibili: Sono disponibili due forniture di dati, la prima riguarda il rilevamento dei GPS di un campione di 156.436 veicoli privati circolanti in Toscana nel Maggio 2011. La seconda riguarda il rilevamento dei GPS di un campione di 250.239 veicoli. Le forniture non contengono taxi e flotte private (es: autonoleggi). Il criterio di selezione del campione è: persone che hanno attraversato la regione Toscana nella finestra temporale considerata. A differenza della prima fornitura l'area presa in considerazione nella seconda fornitura è leggermente più ampia e comprende anche la città di Bologna.

156.436
NUMERO VEICOLI

05-2011
PERIODO DI RIFERIMENTO

Dati di telefonia mobile GSM

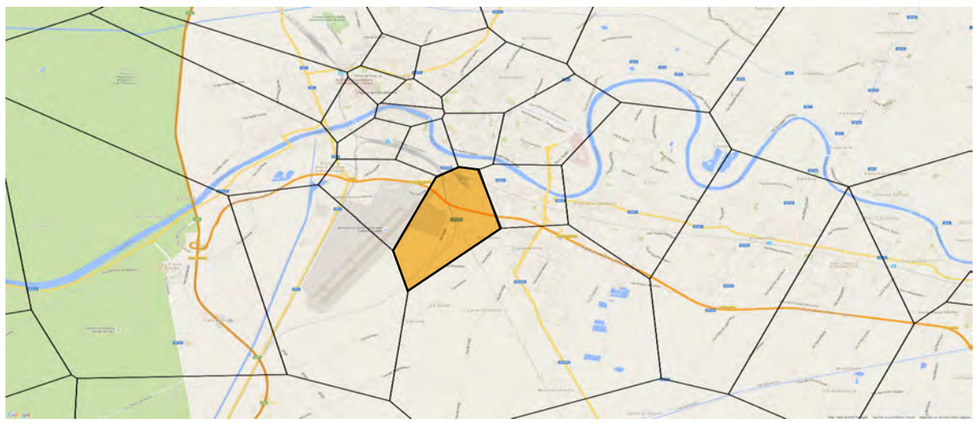
I dati GSM sono dati di telefonia mobile raccolti dall'operatore telefonico durante eventi generati dall'uso dei telefoni cellulari (chiamate, SMS, traffico dati). All'interno del progetto vengono utilizzati i Call Data Record (CDR). I CDR sono record registrati durante le attività di chiamata e contengono l'informazione temporale (timestamp), un identificatore anonimo dell'utente, la cella telefonica o l'antenna che ha servito la chiamata e la durata della chiamata.

Fornitore: WIND Telecomunicazioni (4) è un operatore leader nel settore delle telecomunicazioni in Italia. Con la sua infrastruttura di rete mobile copre potenzialmente il 99,88% della popolazione italiana. Vanta, come numero di clienti (SIM Card registrate), al 31 marzo 2015 circa 21,4 milioni di utenti e al 31 marzo 2014 circa 22 milioni. Detiene, come quota stimata di mercato italiano, al 31 marzo 2015 il 24,6% della popolazione mobile e al 31 marzo 2014 il 24,3%. Escludendo gli Operatori Mobili Virtuali si colloca come terzo operatore dopo Tim Telecom Italia che ha una quota di mercato fra il 34-35%, Vodafone fra il 29-31% e, seguito da H3G fra il 10-11% della popolazione mobile. Il mercato italiano rappresenta il quarto maggior mercato europeo, seguendo il Regno Unito, la Francia e la Germania. In Italia, al 31 dicembre 2014, si contano 86,8 milioni di SIM registrate con tasso di penetrazione di circa 143% della popolazione italiana (dato dovuto alla presenza di più SIM registrate per utente).

Granularità: Antenna Telefonica (allegare mappa Voronoi).

Copertura spaziale: Provincia di Livorno, Pisa, Lucca e Firenze.

Finestra temporale: 16 febbraio 2014 - 23 marzo 2014.
Dati Disponibili: I dati descrivono le attività di chiamata degli utenti di telefonia mobile all'interno delle province di Firenze, Livorno, Lucca e Pisa. L'infrastruttura è costituita da una rete di antenne dislocate sul territorio italiano che coprono idealmente delle aree organizzate in un diagramma di tipo Voronoi. Le attività descritte all'interno della rete, di base GSM (Global System for Mobile Communications), sono monitorabili attraverso i CDR (Call Data Record). I CDR sono registri di attività di chiamate in cui vengono raccolti dati relativi al codice attribuito all'utente, all'antenna telefonica interessata e al tempo d'inizio di un evento chiamata con la sua durata. Le attività riportate nei CDR sono riferite solo a chiamate telefoniche mentre sono escluse tutte le attività di traffico dati inerenti alla connessione Internet. I dati raccolti sono riferiti solo alle attività del singolo utente, e non è presente l'identificatore dell'utente ricevente la chiamata. I CDR sono raccolti dall'operatore per finalità di fatturazione.
Il dataset ha una finestra temporale di 2 mesi (febbraio e marzo 2014) e contiene 1146897 utenti distinti e 51155663 eventi.



tassellazione di Voronoi che corrisponde alle antenne dell'area di Pisa

Matrici Origine/Destinazione dei tempi Comune – Attrattore

La matrice origine destinazione è uno degli strumenti più utilizzati per rappresentare la domanda di mobilità in riferimento ad una rete viaria. Si tratta di una matrice dove sulle righe sono rappresentate le origini (O) e sulle colonne le destinazioni (D). I valori della tabella indicano generalmente tempi di percorrenza o distanze.
Le API Google Maps Distance Matrix fornite da Google, sono state utilizzate per calcolare le matrici Origine Destinazione dei tempi di percorrenza usando sia i mezzi privati (auto) che il trasporto pubblico locale (TPL). Come punto di destinazione sono state utilizzate le coordinate dell'attrattore analizzato, mentre come origine le coordinate del Municipi dei comuni toscani.
Attraverso una chiamata di questo tipo:

```
distanza= gmaps.distance_matrix(
      origins=[10.1012,440793],
      destinations='Aeroporto Amerigo Vespucci, Firenze, FI',
      mode='driving',
      departure_time=1453741798,
      traffic_model='best_guess')
```

è stato un file Json con queste caratteristiche:

```
{u'destination_addresses': [u'Florence Airport, Peretola, Via del Termine,
11, 50127 Firenze, Italy'],
 u'origin_addresses': [u'Carrara MS, Italy'],
 u'rows': [{u'elements': [{u'distance': {u'text': u'123 km', u'value':
123343},
     u'duration': {u'text': u'1 hour 21 mins', u'value': 4878},
     u'duration_in_traffic': {u'text': u'1 hour 23 mins', u'value': 4954},
     u'status': u'OK'}]}],
 u'status': u'OK'}
```

Cambiando il parametro mode='driving' in mode='transit' è possibile ottenere lo stesso risultato calcolato per il Trasporto Pubblico Locale, che viene considerato come la migliore combinazione di mezzi pubblici, treno, e spostamento a piedi per raggiungere il punto di interesse.
Le API utilizzate restituiscono inoltre una stima del traffico in condizioni normali (parametro traffic_model='best_guess') ed è stato osservato un giorno infrasettimanale tipo, un lunedì mattina: departure_time = 1453741798, per non influenzare l'algoritmo che stima il traffico con parametri particolari.
L'uso di questa matrice delle distanze è stato preferito a quella inizialmente fornita da IRPET in quanto tiene conto della stima del traffico e restituisce risultati molto più realistici rispetto a quelli calcolati sulla rete scarica. È stata poi calcolata una matrice OD utilizzando i dati forniti da OCTO TELEMATICS. Nonostante questo tipo di matrice restituisca l'effettivo tempo medio di percorrenza da comune ad attrattore derivandolo dall'osservazione diretta dei dati e non sia frutto di una predizione, nel caso dei comuni con pochi veicoli si è riscontrata un'osservazione difficilmente riconducibile al comportamento dell'utente medio. Si è pertanto deciso di non usare questa matrice come strumento di analisi per tutto il progetto a causa di una scarsa rappresentatività del campione, in particolare per quanto riguarda i dati del 2011. Per alcuni comuni di piccole dimensioni infatti le stime ottenute sono talvolta risultate non del tutto attendibili. Si è preferito quindi affidarsi allo strumento di Google che sulle zone considerate più critiche da questo punto di vista ha una disponibilità di dati sicuramente più ampia per fare delle stime.

3.2 Sorgenti dati socio-demografiche

Dati sociodemografici Censimento 2011

Fornitore: Istituto Nazionale di Statistica – ISTAT (5).
Granularità minima: Zona censuaria.
Copertura spaziale: Italia.
Finestra temporale: Raccolta dati riferita al 9 ottobre 2011.
Dati Disponibili: I dati sono relativi al censimento 2011. Le informazioni raccolte per il 15° Censimento generale si riferiscono al 9 ottobre 2011 e sono raccolte sul data warehouse dell'Istat e accessibili tramite il sito web. Il censimento descrive la popolazione e le abitazioni e contiene informazioni, disaggregate fino a livello subcomunale (zona censuaria), sulla struttura demografica e sociale della popolazione con dimora abituale in Italia e sul patrimonio abitativo italiano. In particolare, sono stati presi in considerazione i dati relativi al numero di abitanti e i dati geografici sui comuni toscani (shape file) dai quali è stata ricavata, con analisi GIS, l'estensione e varie misure derivate per descrivere il territorio, come ad esempio la densità abitativa.

Parco veicolare per categoria e comune

Fornitore: Automobile Club d'Italia – ACI (6). L'ACI è oggi in Italia la più grande libera associazione di cittadini, dei quali si propone come rappresentante e portavoce presso le istituzioni nazionali ed internazionali riguardo alle tematiche del turismo, della mobilità e dell'ambiente. Ente Pubblico, l'ACI gestisce una serie di servizi tra cui la gestione del Pubblico Registro Automobilistico.

Granularità: scala comunale.
Dati Disponibili: L'ACI fornisce sul suo portale una serie di pubblicazioni statistiche relative al fenomeno della motorizzazione nel suo complesso con dati aggiornati dal 2002 al 2014. In particolare è stato utilizzato il dataset del parco veicolare per categoria e comune per riportare il campione fornito da Octo Telematics alla parco circolante della Toscana.

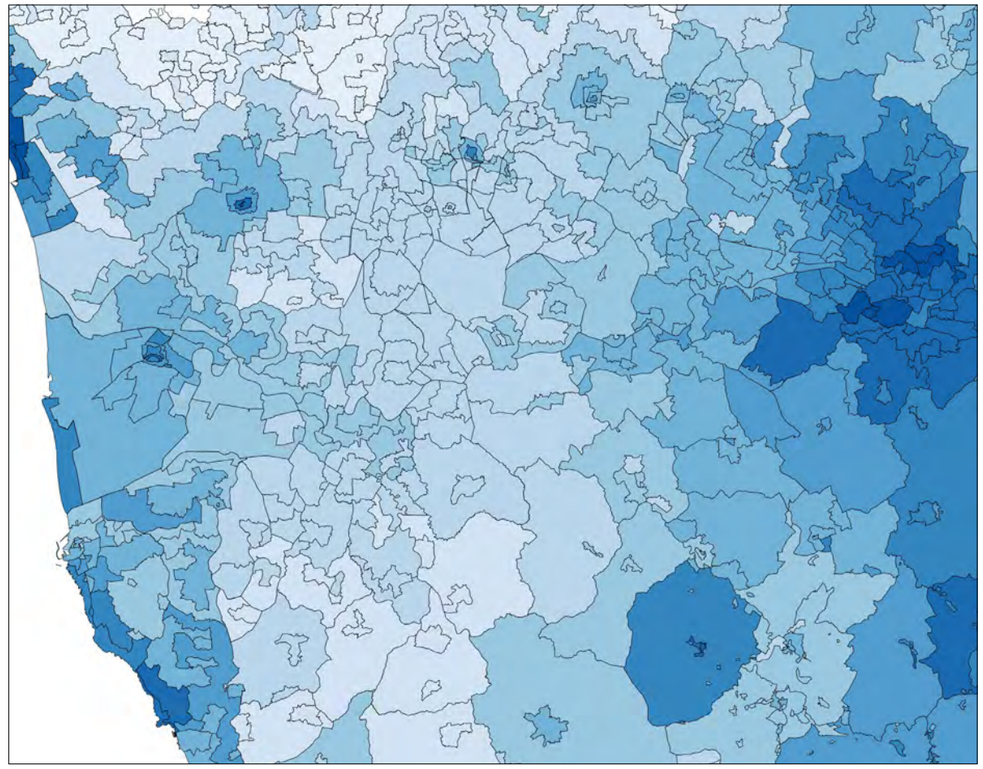
Dati di traffico aereo

Fornitore: L'Ente Nazionale per l'Aviazione Civile - Enac (7) è l'autorità che si occupa della regolazione dell'aviazione civile, del controllo e vigilanza sull'applicazione delle norme adottate, della disciplina degli aspetti amministrativo-economici del sistema del trasporto aereo. Tra le varie attività si occupa dell'esame e la valutazione dei piani regolatori e dei programmi di interventi, investimenti e sviluppo in ambito aeroportuale; Per questo motive effettua un monitoraggio costante sullo stato degli aeroporti e I loro flussi di traffico.

Granularità: Mensile.
Copertura spaziale: Italia
Finestra temporale: annuale dal 2002 al 2014
Dati Disponibili: I dati dell'Enac disponibili sono relativi all'andamento mensile del traffico passeggeri, alla suddivisione del traffico per principali destinazioni e alla tipologia di volo (vettori tradizionali o low-cost).

Aree Omi

Fornitore: Agenzia delle Entrate (fornitura dati acquistata da IRPET)
L'Osservatorio del Mercato Immobiliare che fa capo all'Agenzia delle Entrate cura la rilevazione e l'elaborazione delle informazioni di carattere tecnico-economico relative ai valori immobiliari, al mercato degli affitti e ai tassi di rendita e la pubblicazione di studi ed elaborazioni e la valorizzazione statistica degli archivi dell'Agenzia.
Granularità: aree Omi (dimensione subcomunale).
Dati Disponibili: la fornitura messa a disposizione da IRPET è relative alle zone omogenee OMI del 2010-2011, in particolare sono stati considerati i valori medi per la destinazione residenziale di ciascuna zona OMI. L'ipotesi di base dell'OMI è che il fattore posizionale sia quello maggiormente esplicativo delle differenze di valore tra le varie unità immobiliari, in particolare di quelle a destinazione residenziale. Si ritiene pertanto che al fine di ridurre la variabilità campionaria è utile suddividere il territorio in porzioni che esprimano livelli omogenei di mercato, in funzione delle caratteristiche comuni (urbanistiche, socio-economiche, di dotazioni di servizi, ecc.).
Pertanto ogni territorio comunale è segmentato in una o più zone omogenee. La zona omogenea OMI riflette un comparto omogeneo del mercato immobiliare locale, nel quale si registra una sostanziale uniformità di apprezzamento per condizioni economiche e socio-ambientali.
Al fine di avere un quadro uniforme del mercato immobiliare sull'intero territorio nazionale, le zone OMI, infine, sono state raggruppate in fasce, le quali individuano aree territoriali con precisa collocazione geografica nel Comune e rispecchiano, in generale, una collocazione urbanistica consolidata (Centrale, Semicentrale, Periferica, Suburbana, Rurale).



aree OMI della zonda compresa tra Pisa e Firenze

3.3 Identificazione delle geometrie

Per l'identificazione degli attrattori si è consultato il sito *dati.toscana.it*, repository di dati open fornito dalla regione Toscana. Dal momento che il database della Regione fornisce solamente una geolocalizzazione puntuale delle strutture si è proceduto al reperimento delle geometrie più complesse utilizzando Open Street Map. In ultima istanza è stato fatto un controllo delle aree selezionate osservando i punti di sosta dei veicoli GPS intorno alle strutture considerate e osservando l'uso del territorio con le Ortofotocarte. Avvalendosi del parere di esperti di dominio e del supporto di IRPET, sono state corrette le geometrie includendo aree di influenza dell'attrattore non considerate da Open Street Map.
Attraverso Qgis e i plugin OSMDownloader e QuickOSM è stato possibile interrogare il database di Open Street Map per arricchire i dataset utilizzati con shapefile dettagliati. Per alcuni punti di interesse è stato necessario procedere con un aggiustamento delle geometrie degli shape file scaricati al fine di includere aree che formalmente non appartengono agli attrattori studiati ma che sono effettivamente utilizzate per la sosta. È stato fatto anche un tentativo per capire quanto la geometria di un attrattore potesse essere assimilabile alla cella censuaria di Istat su cui insiste. Tale operazione avrebbe permesso di riportare le analisi ad una tassellazione del territorio nota e per cui sono disponibili molte statistiche. Tuttavia, osservando le modalità in cui i veicoli sostano nelle vicinanze dell'attrattore, si identifica un sostanziale "rumore" dovuto a quelli che di fatto non hanno come destinazione finale l'aeroporto. Per tale motivo l'intervento manuale è stato volto a definire in modo più specifico le aree di parcheggio tipicamente utilizzate da chi si reca all'aeroporto, diminuendo di fatto il rumore generato dal traffico cittadino.

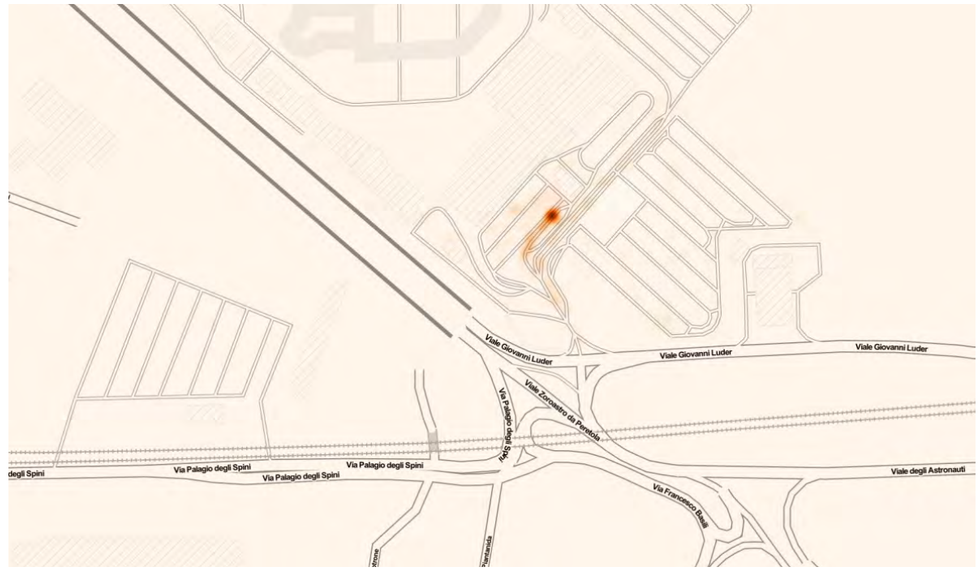
Visualizzazione delle soste intorno all'area dell'aeroporto Galilei

Fotografie Georeferenziate

Fonte: Regione Toscana – Geoscopio
Le foto sono state osservate tramite il servizio online della regione Toscana: Geoscopio – macchina temporale. Sono un utile supporto per la comprensione dell'uso del territorio nel tempo, in particolare, nell'analisi dei dati di mobilità vengono usate per comprendere come cambia la dislocazione delle possibili aree di sosta.
Granularità: Fotografie aeree con alto livello di dettaglio.
Dati disponibili: Ortofotocarte scala 1:2000 (anni 2010-2013)

Open Street Map

Fonte: Open Street Map – OSM
È un progetto collaborativo finalizzato a creare mappe del mondo a contenuto libero. Il progetto punta ad una raccolta di dati geografici, con scopo principale la creazione di mappe e cartografie. Le mappe sono create usando come riferimento i dati registrati da dispositivi GPS portatili, fotografie aeree ed altre fonti libere.
Dati utilizzati: Shape file dei punti di interesse.
OpenStreetMap è un progetto collaborativo finalizzato a creare mappe a contenuto libero del mondo. Il progetto punta ad una raccolta mondiale di dati geografici, con scopo principale la creazione di mappe e cartografie. Una particolarità di OSM, che con ogni probabilità verrà sfruttata per l'arricchimento semantico nelle fasi successive di progetto, è la classificazione molto accurata di tutti gli elementi scaricabili dal dataset.



Visualizzazione delle soste intorno all'area dell'aeroporto Vespucci

4.

UNA PANORAMICA SUGLI AEROPORTI DELLA REGIONE TOSCANA

L'obiettivo del lavoro di tirocinio è lo studio dell'impatto sulla mobilità di grandi infrastrutture della Regione.
Nel caso specifico sono stati analizzati i due aeroporti principali della Toscana oggetto della recente e discussa fusione. Tra i fattori che influenzano lo studio vi è la distanza geografica nella scelta dell'aeroporto, la concorrenza di altri scali e la specializzazione dell'offerta dei due aeroporti.
I paragrafi seguenti riportano alcuni dati pubblicati da Enac sull'uso degli scali nazionali e lo Studio sullo sviluppo degli aeroporti italiani (8) effettuato dall'ente sui due aeroporti. Tale panoramica è stata un'utile guida per orientarsi nell'analisi dei dati.

4.1 L’aeroporto internazionale “Amerigo Vespucci” di Firenze-Peretola

Posizionamento e ruolo

L’aeroporto di Firenze ha registrato nel 2008 più di 1,9 milioni di passeggeri annui, con un traffico quasi esclusivamente di linea, caratterizzato da una forte internazionalità (73%), con collegamento verso 15 città europee.
Le tratte internazionali sono molto differenziate, mentre quelle nazionali sono maggiormente concentrate sui collegamenti con Malpensa (29%), Fiumicino (26%) e Catania (17%), che insieme rappresentano il 72% del totale nazionale.
Insieme con l’aeroporto di Pisa, con il quale si sommano oltre 5 milioni di passeggeri, lo scalo di Firenze ha risposto con una crescita consistente nell’ultimo decennio alla domanda di trasporto aereo di un bacino denso e strategico per l’intero Paese che si estende anche oltre la Toscana, incidendo significativamente sullo sviluppo economico ed in particolare sul turismo.
Proprio la lettura della crescita in parallelo dei due scali e la caratterizzazione molto diversificata dei rispettivi flussi, deve orientare verso il superamento della competizione in chiave di rivalità per concentrare invece le attenzioni sul consolidamento e lo sviluppo delle infrastrutture e dei servizi sulla base degli importanti progetti già approntati. Tale approccio, per poter gestire ulteriori fasi di crescita e rafforzare i ruoli nel sistema anche a scala nazionale.

Inquadramento territoriale

L’aeroporto è situato a nord-ovest di Firenze, al limite nord dell’area densamente abitata di Peretola, e ricade nei Comuni di Firenze e Sesto Fiorentino. Lo scalo dista circa 4 Km da Firenze, 92 km da Pisa, 95 km da Arezzo, ed è relativamente vicino anche a capoluoghi di altre regioni come Bologna (100 km), La Spezia (154 km) e Perugia (170Km). Lo scalo nasce negli anni trenta come campo di volo e apre al traffico commerciale negli anni ’40 con l’inizio dei primi voli di passeggeri.
Il sedime aeroportuale ha un’estensione di 120 Ha ed è delimitato ad est dalla A11, a nord dal Canale di Cinta Orientale, verso est ed ovest da aree agricole sottoposte a salvaguardia ambientale.

Accessibilità e sosta

L’aeroporto è collegato alla viabilità ordinaria in corrispondenza dell’importante svincolo rotatorio adducente verso nord-ovest all’autostrada A11 Firenze - Mare e quindi all’autostrada A1 Bologna-Roma, nonché alle strade statali Pratese e Pistoiese.
Relativamente all’accessibilità urbana si individuano due assi principali: il viale Guidoni che collega al centro cittadino e viale XI Agosto che raggiunge la Piana di Castello, Rifredi e Sesto Fiorentino.
Lo scalo non è collegato con la linea ferroviaria. Attualmente l’aeroporto è collegato alla stazione ferroviaria di Firenze da autobus urbani (tempo di percorrenza 25 minuti, frequenza ogni 30 minuti) ed alla stazione di Prato da autobus extra urbani (tempo di percorrenza 60 minuti, frequenza ogni 30 minuti), ma la percentuale di utenti che utilizzano il bus per raggiungere l’aeroporto rimane ancora bassa (13%) per uno scalo inserito nel tessuto urbano della città.
La viabilità all’interno dell’area terminale è resa complessa dalla sua particolare morfologia e dalla limitatezza degli spazi. I parcheggi, divisi in due aree, dispongono complessivamente di circa 900 posti auto, ai quali si aggiungono circa 390 posti auto per il personale e le auto a noleggio.

Criticità e progetti in corso

L’aeroporto presenta una molteplicità di problematiche, legate al suo inserimento in un contesto urbanizzato difficile, a cui si aggiungono ridotte dimensioni e limiti delle infrastrutture di volo, nonché obsolescenza e scarsa funzionalità del terminal passeggeri.
Le ipotesi di allungamento della pista si scontrano con una molteplicità di costruzioni disseminate nelle immediate vicinanze del sedime e aree sottoposte a vincolo paesaggistico, ciononostante l’indirizzo politico degli ultimi anni sembra essere fortemente orientato all’ipotesi di ampliamento dell’infrastruttura.

4.2 l’aeroporto internazionale “Galileo Galilei” di Pisa

Posizionamento e ruolo

L’Aeroporto di Pisa ha costruito il proprio ruolo di porta principale d’accesso alla Regione Toscana, capace di servire un network che comprende anche destinazioni intercontinentali e connessioni di superficie in un’ottica di sviluppo intermodale.
La forte crescita del traffico passeggeri dall’anno 2000 ad oggi, dovuta all’avvento delle compagnie aeree Low Cost, ha di fatto raddoppiato il traffico complessivo richiedendo significativi potenziamenti di infrastrutture e servizi, che hanno consentito al gestore di consolidare il ruolo di scalo principale nel bacino di traffico.

L’aeroporto Galileo Galilei è ubicato nella zona sud del territorio comunale di Pisa, in località San Giusto, in

Inquadramento territoriale

adiacenza al tessuto urbano. Di proprietà in gran parte dell’Aeronautica Militare, viene costruito intorno agli anni 30. Negli anni 60-70 viene costruita l’attuale aerostazione ad uso civile con trasferimento al Demanio Trasporti di parte del sedime dello scalo.
L’aeroporto dista 2 km dal centro città di Pisa, 20 km da Livorno e 80 km da Firenze. Il sedime aeroportuale occupa un’area di 378 ettari con ampie zone libere da ostacoli ed è ubicato in prossimità di nuclei residenziali sull’asse est-ovest e delimitato a nord dalla strada a scorrimento veloce Firenze-Pisa-Livorno e dalla ferrovia e, a sud, dall’aeroporto militare.

Accessibilità e sosta

L’aeroporto è accessibile da autostrade, superstrade. È direttamente accessibile dalla Superstrada Firenze - Pisa (uscita Pisa Centro-Aeroporto), dall’Autostrada A12 Genova - Rosignano (uscita Pisa Centro), dall’Autostrada A11Firenze - Pisa Nord (attraverso lo svincolo di Pisa Nord con la A12) e dalla strada statale SS1 Aurelia. Con i mezzi su gomma si accede direttamente al piazzale, dove il curb presenta spazi adeguati per auto private, auto in servizio pubblico (taxi), auto a noleggio e autobus di linea urbana ed extraurbana. Il fronte dell’aerostazione è dotato di un’ampia pensilina di protezione.
Per quanto riguarda il collegamento ferroviario, la linea dirette che collegano l’Aeroporto di Pisa con la stazione di Pisa Centrale è stata soppressa nel 2010.
Sono disponibili ampi parcheggi per la sosta breve, a pagamento o a disco orario per le auto private. Nell’area ovest del land side è presente un parcheggio multipiano a tre livelli di 1.015 posti auto e, a est, il parcheggio per le auto a noleggio con ben 1.500 posti auto. Per contratti di autonoleggio lo scalo di Pisa dispone di 17 piazzole di sosta ed è uno tra i primi in Italia. È disponibile una stazione dei bus che collegano l’aeroporto a vari centri della Regione.
L’aeroporto Galilei ha visto una costante crescita del numero di passeggeri negli anni e conferma della sua

Criticità e progetti in corso

importanza a livello nazionale grazie all’offerta low cost variegata e ad una ottima accessibilità a livello infrastrutturale. La chiusura della linea ferroviaria nel 2010 è stata compensata con un potenziamento del trasporto pubblico locale verso l’aeroporto, tuttavia il servizio non raggiunge gli stessi standard di efficienza. Per questo motivo è stato ideato il progetto di People Mover, più adeguato alla rapida connessione con la stazione FS e che dovrebbe vedere la luce nel 2017.

L' IMPATTO SULLA MOBILITÀ GENERATO DAGLI AEROPORTI INTERNAZIONALI DELLA TOSCANA

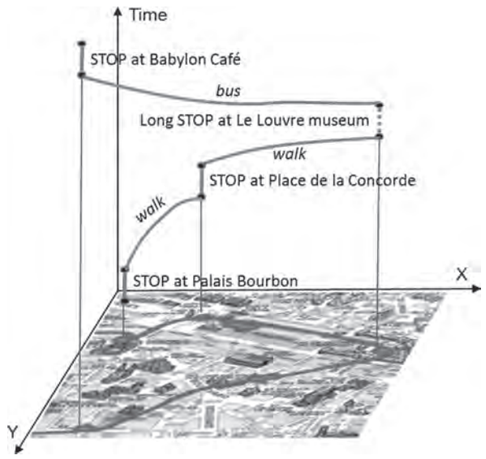
5

La disponibilità delle tracce GPS ottenute tramite i dispositivi di OCTO TELEMATICS sono un indicatore della mobilità degli individui, un microscopio sociale che permette di scoprire pattern nascosti e modelli di comportamento che caratterizzano l'uomo. Attraverso lo studio di questi dati è possibile fornire un utile supporto all'ingegneria dei trasporti ma più in generale alla progettazione urbana e l'analisi socio economica.
Tipicamente le analisi di mobilità vengono fatte con due tipi di fonti di dati: GPS e GSM. Durante il tirocinio ci si è concentrati sull'analisi dei dati GPS. A parità di area coperta dai due tipi di dati disponibili bisogna considerare nella valutazione dei risultati che i GPS di OCTO TELEMATICS sono un campione della popolazione veicolare di circa il 2% delle automobili (dataset 2011) mentre WIND ha una penetrazione di mercato intorno al 30% della popolazione italiana. Il dato GPS tuttavia ha una granularità molto più fina e nel caso vengano studiati attrattori inseriti in un contesto urbano a densità abitativa media o medio alta offre un dettaglio nell'osservazione molto più alto del GSM. Ancora è bene ricordare che i GPS utilizzati osservano un solo tipo di mobilità, quella veicolare privata. È pertanto opportuno sottolineare ad esempio che nel caso di un'analisi comparativa tra due attrattori la prevalenza di presenze date dalle tracce GPS non descrive la predilezione in assoluto verso un attrattore bensì la predilezione a recarvisi con il mezzo privato.

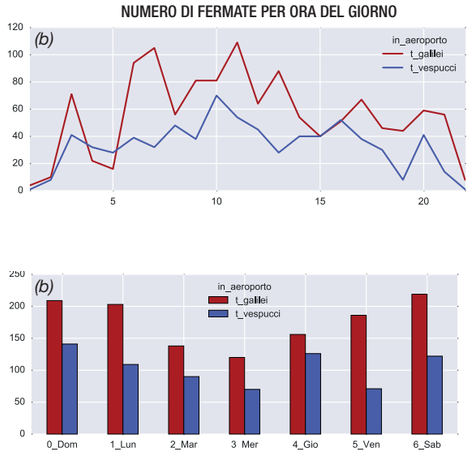
Nei paragrafi successivi vengono descritte brevemente alcune tecniche e modelli di analisi utlizzati: nel primo vengono analizzate le soste nei pressi degli attrattori da un punto di vista temporale (arrivo nei luoghi di sosta e durata delle soste); nel secondo paragrafo si è osservato come il traffico generato dai due attrattori influenza la mobilità del territorio circostante; nel terzo paragrafo viene descritto il metodo di costruzione delle matrici Origine Destinazione e viene descritta l'influenza che i due aeroporti presi in esame esercitano sulla regione Toscana.

5.1 L’analisi delle soste: un’applicazione del modello di mobilità basato sulle traiettorie.

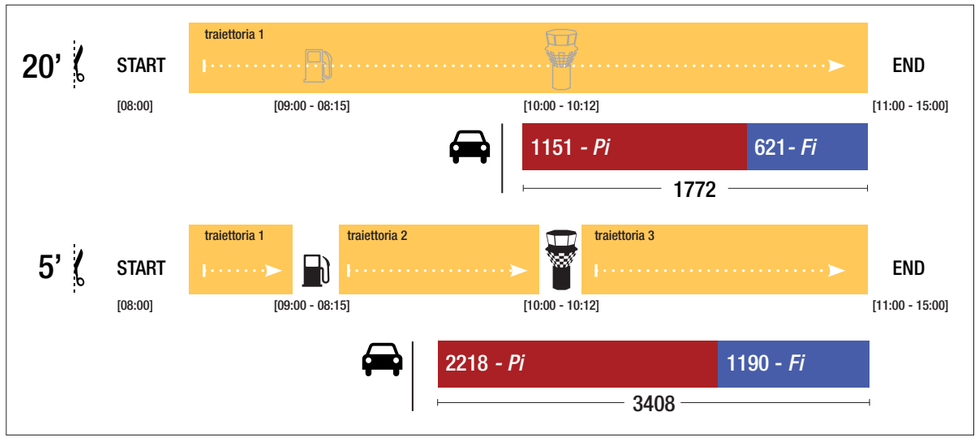
Ogni segnale emesso da un rilevatore GPS registra un punto con coordinate lat lon e un timestamp. La sequenza di punti trasmessa dai rilevatori di posizione GPS di ogni veicolo, se unita in una linea, crea un itinerario o movement track. Dal momento che la movement track può essere molto lunga (nel caso in esame un mese di dati) essa deve essere divisa in segmenti chiamati traiettorie per poter effettuare analisi significative. Il taglio che viene effettuato sull’intera sequenza è determinato dalla tipologia di oggetto osservata, dal contesto in cui si muove l’oggetto e dallo scopo dell’osservazione.
La traiettoria di un veicolo viene definita utilizzando la sosta come discriminante per il taglio della movement track. Tipicamente viene tagliata una traiettoria di un auto che si muove in un contesto urbano quando la distanza tra due punti nell’arco di 20 minuti è inferiore a 50 metri. Il primo punto è definito come Start Point e l’ultimo come End Point. Nonostante la letteratura (CIT) definisca 20 minuti come la soglia che ben identifica un veicolo che effettua una sosta, nel caso degli aeroporti tale tempo risulta essere troppo lungo. Molti guidatori effettuano brevissime fermate nei pressi dell’aeroporto per permettere una rapida discesa dei passeggeri e ripartire immediatamente. Tale comportamento è incentivato dal fatto che la sosta gratuita per i veicoli nel parcheggio antistante gli aeroporti è consentita per un periodo di tempo inferiore ai 15 minuti.
È stato quindi fatto un esperimento considerando tagli delle traiettorie con soglia di fermata di 20 minuti ed



Rappresentazione delle traiettorie di un viaggio

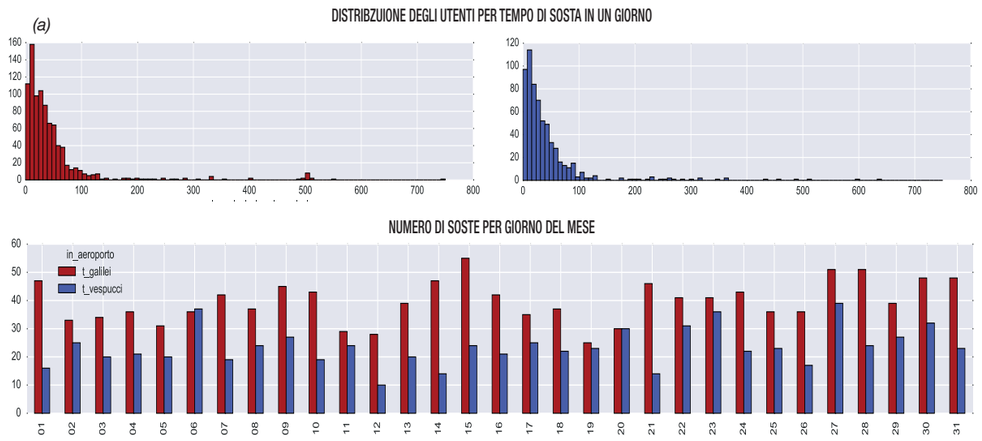


(b) Numero di soste per ora del giorno
(c) Numero di soste per giorno della settimana



Differenze nella scelta della soglia per il taglio delle traiettorie

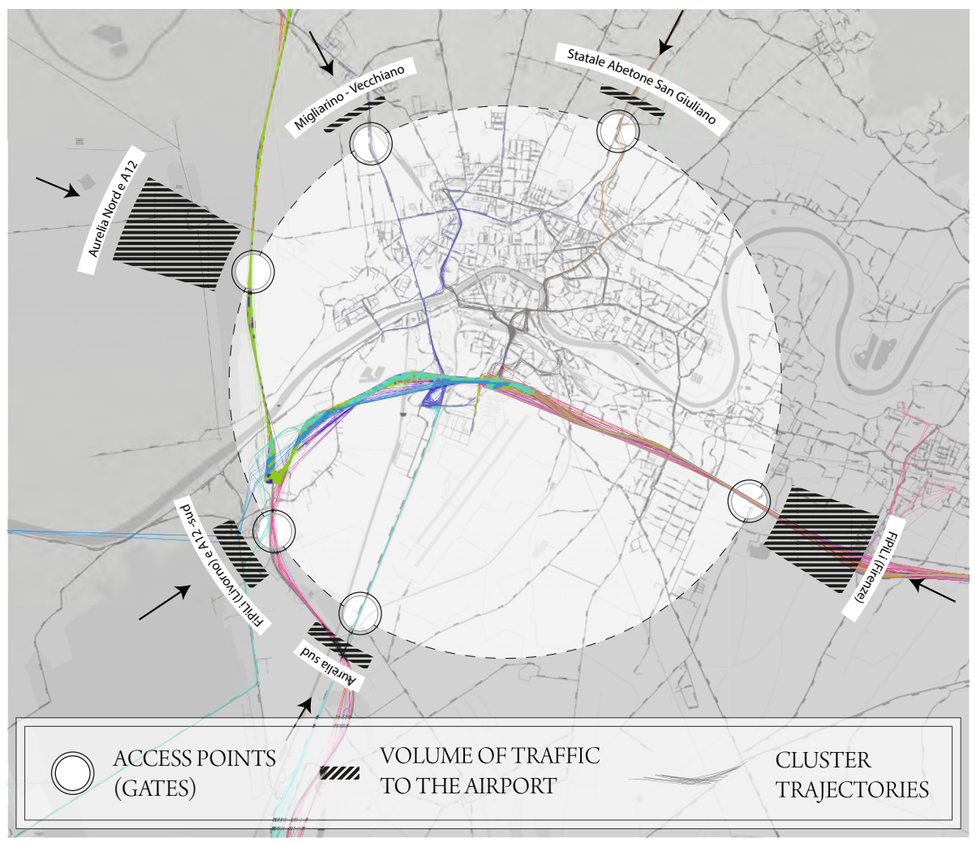
è stato ripetuto con soglie di 5 minuti. Come previsto il numero di veicoli osservati è aumentato significativamente passando da 1772 a 3408 veicoli osservati, +108%.
Tale dato è confermato dall’osservazione della distribuzione delle soste durante l’arco di una giornata. Si può notare infatti che le soste hanno durata generalente inferiore ai 30 minuti, con una concentrazione particolare di auto che sostano per meno di un quarto d’ora. Con le traiettorie così selezionate è stato possibile produrre diverse statistiche sui tempi di sosta da cui si evince che la distribuzione delle soste nei giorni della settimana nei due aeroporti è abbastanza differente e probabilmente è legata alla diversa offerta commerciale dei due aeroporti: la maggioranza di voli lowcost al Galilei tende a concentrare l’affluenza nei weekend e nei giorni a ridosso dei medesimi mentre al Vespucci la vocazione più business della clientela tende a fruire dell’aeroporto in maniera significativa anche durante la settimana, come si evince dal picco di presenze registrato nei giovedì.

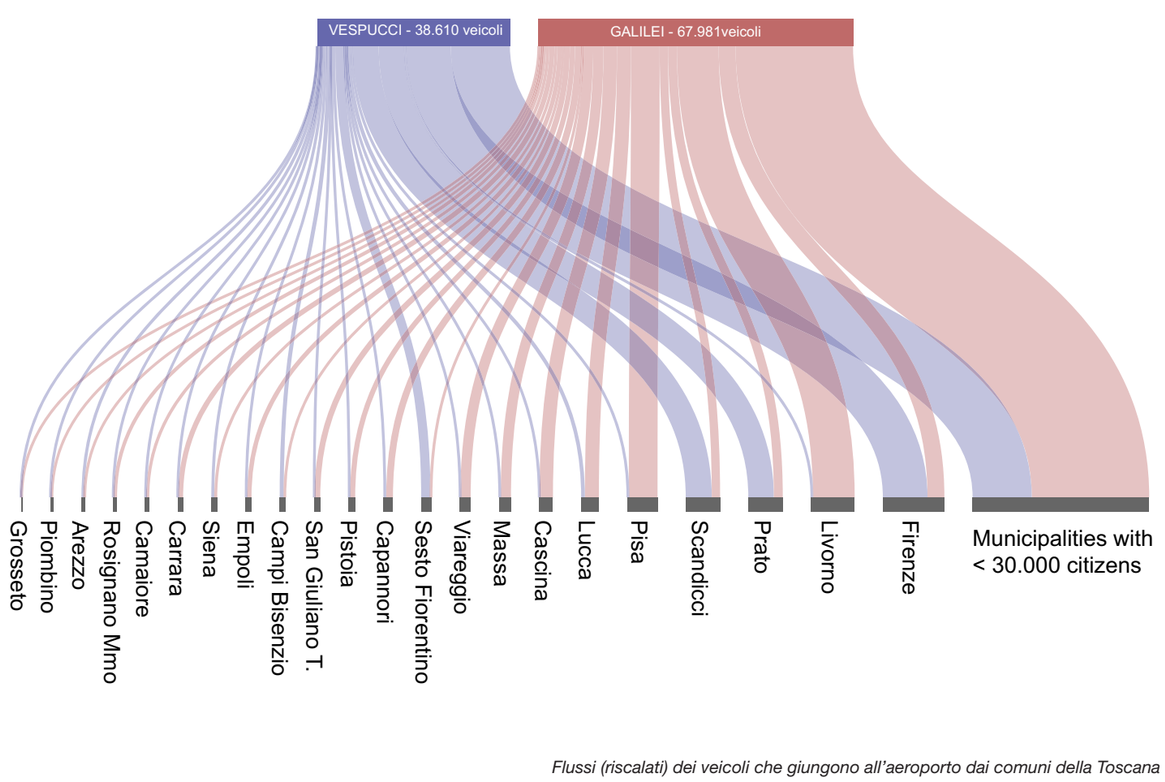


(d) Distribuzione delle soste all’aeroporto Galilei e Vespucci
(e) Distribuzione delle soste nei giorni della settimana

5.2 L’identificazione degli access point: un’applicazione del mobility datamining per descrivere comportamenti collettivi di mobilità.

Grazie alla disponibilità dei dati GPS è possibile analizzare l’impatto della mobilità a livello locale in maniera molto dettagliata. L’insieme delle tracce che giungono all’aeroporto sono state osservate attraverso gli strumenti messi a disposizione da M-Atlas. M-Atlas (acronimo di Mobility Atlas) è un software sviluppato dal KDD Lab per analizzare con tecniche di datamining comportamenti di mobilità del territorio. In particolare è stato utilizzato per osservare come il traffico veicolare che giunge all’attrattore impatta sulla città.
In prima istanza è stato creato un buffer di 5 km intorno all’area dei due aeroporti. Si è andati quindi ad identificare sul perimetro della geometria disegnata quali sono le arterie stradali più interessate dai veicoli che si recano agli aeroporti, individuando i principali access-point. A questo punto, grazie all’algoritmo di route similarity di M-Atlas, è stato possibile caratterizzare i flussi passanti per ciascun access point in più cluster, etichettandoli poi per area di provenienza.
L’algoritmo di Route Similarity è un algoritmo di cluster che raggruppa le traiettorie in base alla similarità dei percorsi. Volendo concentrare l’analisi sull’individuazione degli access point è stata scelta la versione dell’algoritmo che non considera il fattore temporale.
Grazie ai cluster calcolati possiamo osservare il contributo di ciascuna traiettoria appartenente ad un cluster e passante per un access point. Nell’immagine si può vedere la ripartizione per ogni access point del volume di traffico che termina all’aeroporto durante il periodo considerato.
Il traffico generato dalle traiettorie che giungono all’aeroporto Galilei è circa l’1.3% del traffico totale passante per gli access point considerati. Valore che con buona approssimazione può essere raddoppiato se consideriamo anche i viaggi di ritorno dagli aeroporti.
A questo punto grazie al clustering per Route Similarity sono state raggruppate, ed etichettate, le traiettorie che condividono la maggior parte del percorso. Si può così osservare più nel dettaglio che ad esempio il flusso di veicoli che giunge al Galilei passando dall’access point dall’Aurelia è composto da automobili che provengono sia dalla Versilia che dall’autostrada A12 ed è possibile quantificarlo.
Per quanto rigaurda l’aeroporto Vespucci l’individuazione degli access point è più complessa. Sono stati individuati quattro access point rilevanti per volume di traffico 1) l’Autostrada A1- Siena 2) La superstrada Fi-Pi-Li 3) L’autostrada Firenze-Mare 4) l’A1 in direzione Bologna. Tuttavia bisogna notare come una buona parte del volume di traffico è generato da automobili che attraversano il tessuto urbano della città e che non sono altrettanto ben distinguibili in un flusso veicolare omogeneo. Dal momento che il volume di traffico endogeno della città di Firenze è così elevato si può ipotizzare che un intervento infrastrutturale adeguato sul trasporto pubblico potrebbe alleggerire sensibilmente il traffico urbano.
Nel diagramma di Sankey è proposto un dettaglio dei flussi veicolari che giungono ai due aeroporti da ciascun comune. Da notare il fatto che una parte importante del traffico veicolare verso l’aeroporto Galilei viene generato da automobili provenienti dai comuni più piccoli della Toscana . I flussi rappresentati sono riscalati utilizzando i dati dell’ACI riferiti ai veicoli circolanti nella regione.
Per effettuare la riscalatura e riportare il campione dei veicoli osservati nel dataset OCTO al parco veicolare della Toscana è stato usato un indice di penetrazione Pi calcolato come il rapporto tra il numero di auto “residenti” in ciascun comune secondo il dataset OCTO (TOTocto) e il numero di auto circolanti in ciascun comune fornito dall’ACI (TOTaci). Ne consegue che il numero effettivo di auto provenienti da un comune (Aaci) e osservato nei pressi di un aeroporto sarà uguale al numero di auto osservate (Aocto) fratto l’indice di penetrazione (Pi).



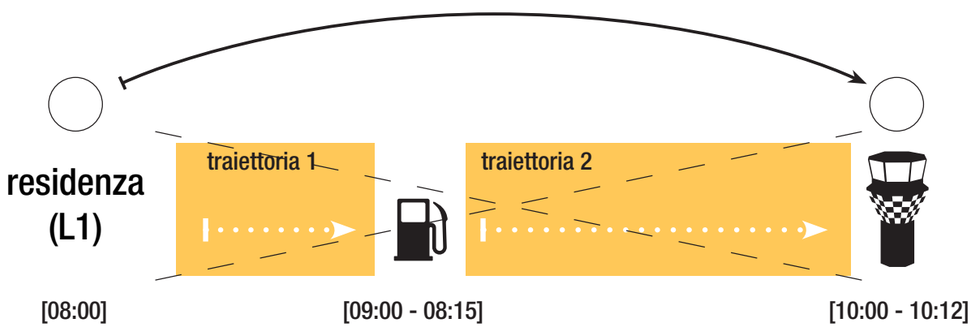


Flussi (riscalati) dei veicoli che giungono all’aeroporto dai comuni della Toscana

5.3 La mobilità dei residenti in Toscana: definizione della matrice OD: Residenza – Attrattore e il concetto di most frequent location.

Tipicamente le matrici Origine-Destinazione utilizzate per l'analisi di mobilità partendo dai dati GPS sono costruite utilizzando Start ed End Point delle traiettorie. Tuttavia, volendo focalizzare lo studio sulla mobilità dei residenti in Toscana siamo andati a controllare se effettivamente i veicoli che sostano nell'area dell'attrattore hanno come origine del proprio viaggio una location che coincide con il proprio comune di residenza. Nel caso di traiettorie con tagli brevi,5 minuti, è altamente improbabile che lo start point della traiettoria coincida sempre con il comune di residenza dell'utente .
Per questo motivo nei successivi studi si è deciso di costruire le matrici origine destinazione considerando come origine il comune in cui viene osservata la first most frequent location di un veicolo. Per most frequent locations si intendono i punti dove è più probabile che il veicolo è osservato quando è fermo. Si presume che la first more frequento location (9) o L1 coincida con la residenza. Per calcolarla vengono osservate tutte le locazioni dove il veicolo si è fermato estraendo start point e end point senza considerare il tempo di sosta. Quindi viene applicato un algoritmo di clustering di tipo Bisecting K-means per raggruppare i punti per similarità. Il Bisecting K-Means è un'estensione del K-Means che suddivide l'insieme di punti in due cluster e quindi procede per divisioni successive fino a che il raggio del cluster non è uguale o minore ad una soglia determinata, nel caso specifico 250 metri. Il centroide del cluster con la più alta frequenza è scelto come L1.

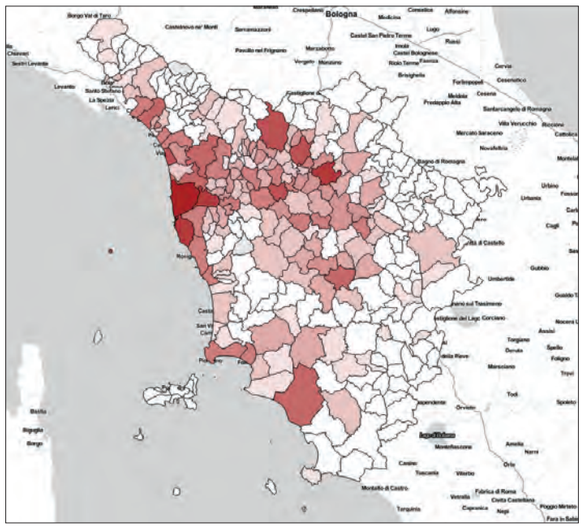
Definita quindi la matrice Origine/Destinazione come matrice Residenti/Attrattore è stata visualizzata l'attrattività per ciascuno dei due aeroporti espressa in numero di residenti che si recano all'aeroporto dal comune considerato. Grazie alla sola visualizzazione si può facilmente notare come l'aeroporto di Pisa sviluppi la sua influenza su un'area più estesa rispetto all'aeroporto di Firenze. Lo step di analisi successivo è stato quello di quantificare in maniera più precisa l'attrattività dei due scali con l'obiettivo di disegnare una carta della regione Toscana composta da due aree formate ciascuna dall'aggregazione dei comuni con una prevalenza di persone che scelgono uno dei due attrattori. È stata inoltre misurata l'accessibilità temporale dei due scali a partire da ciascun comune. Tale accessibilità è stata definita come il tempo necessario a raggiungere un attrattore in condizioni di traffico medie partendo dal municipio di ciascun comune e per calcolarla sono state usate le API Distance Matrix di Google. L'accessibilità temporale è stata calcolata sia per i mezzi privati che per i mezzi pubblici. Nel secondo caso viene considerata la migliore combinazione di autobus, treni



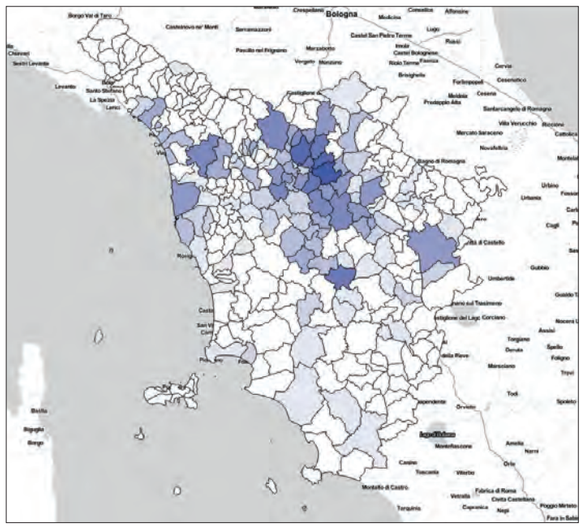
e percorsi pedonali utile per raggiungere l'attrattore considerato. Sebbene il dato più interessante per questa analisi sia stato quello dell'accessibilità con l'automobile, è stata calcolata anche la matrice per i mezzi pubblici. Tale matrice sarà utile per un successivo confronto con i dati GSM, dati più indicati per catturare le dinamiche delle persone che si spostano con questo tipo di mobilità.
Calcolate le due matrici di accessibilità per gli aeroporti, i tempi della matrice del Galilei sono stati sottratti alla matrice del Vespucci. La matrice risultante è stata definita come matrice dell'opportunità. I delta ottenuti possono essere considerati un indice di accessibilità che indica il tempo risparmiato per recarsi ad un aeroporto piuttosto che all'altro. Valori negativi indicano una guadagno di tempo nel recarsi al Galilei mentre valori positivi un guadagno di tempo nel recarsi al Vespucci.
MAPPA SENZA FRECCE DELL'ACCESSIBILITÀ.
Nella figura sono state indicate in bianco i comuni dove l'indice di accessibilità è compreso tra ±10 minuti, ovvero le zone dove l'accessibilità è pressoché ininfluente per la scelta. Tali zone sono di particolare interesse perché la propensione alla scelta non viene influenzata da questo parametro.
Alla stessa maniera delle matrici di accessibilità, anche le matrici Residenti/Attrattore sono state sottratte tra di loro andando ad ottenere una terza matrice che descrive la propensione alla scelta di un aeroporto per ciascun comune.
Sovrapponendo graficamente alla matrice di accessibilità quella della propensione alla scelta risulta subito evidente come le aree bianche dove l'accessibilità risulta essere non influente si orientino completamente verso il Galilei nella scelta. La correlazione tra vicinanza temporale e scelta dell'aeroporto si fa sentire abbastanza per i comuni che sono molto vicini all'attrattore mentre è tendenzialmente poco influente per i comuni più lontani. Da notare che i comuni a Nord dell'aeroporto fiorentino risentono della concorrenza dell'aeroporto di Bologna, non portando un contributo significativo all'aeroporto Vespucci. Lo stesso fenomeno potrebbe essere attribuito agli aeroporti romani per quanto riguarda i comuni nell'area sud della Toscana. Un altro fenomeno degno di nota è la propensione alla scelta del Galilei da parte di molti comuni a sud di Firenze nonostante abbiano un'accessibilità temporale nettamente migliore nei confronti del Vespucci.
Una ulteriore sintesi visiva del fenomeno si ottiene osservando il grafico a barre successivo dove il colore è mappato sulla variabile dell'accessibilità temporale mentre l'altezza degli elementi indica il grado di predilizione verso uno dei due aeroporti (valori positivi indicano, ancora una volta la predilizione per il Vespucci).

L'ultima visualizzazione proposta esemplifica in maniera geografica l'area di influenza. In bianco sono stati lasciati i comuni per cui la predilizione non è stata espressa in maniera netta o il campione di dati non era particolarmente significativo per trarre conclusioni nette. Si può notare per l'ennesima volta come la ripartizione è tendenzialmente frammentata, a volte discontinua.

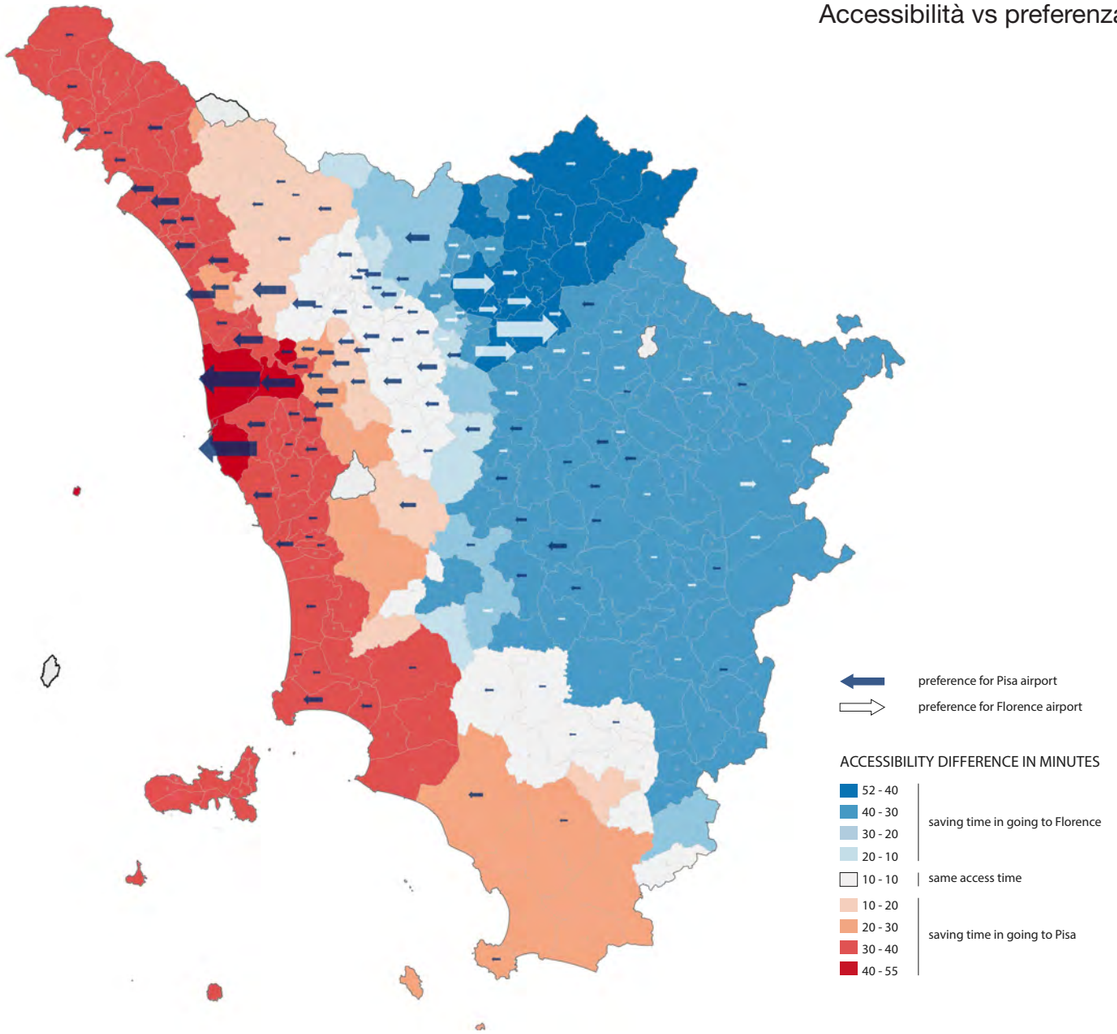
ATTRATTIVITÀ DELL'AEROPORTO GALILEI



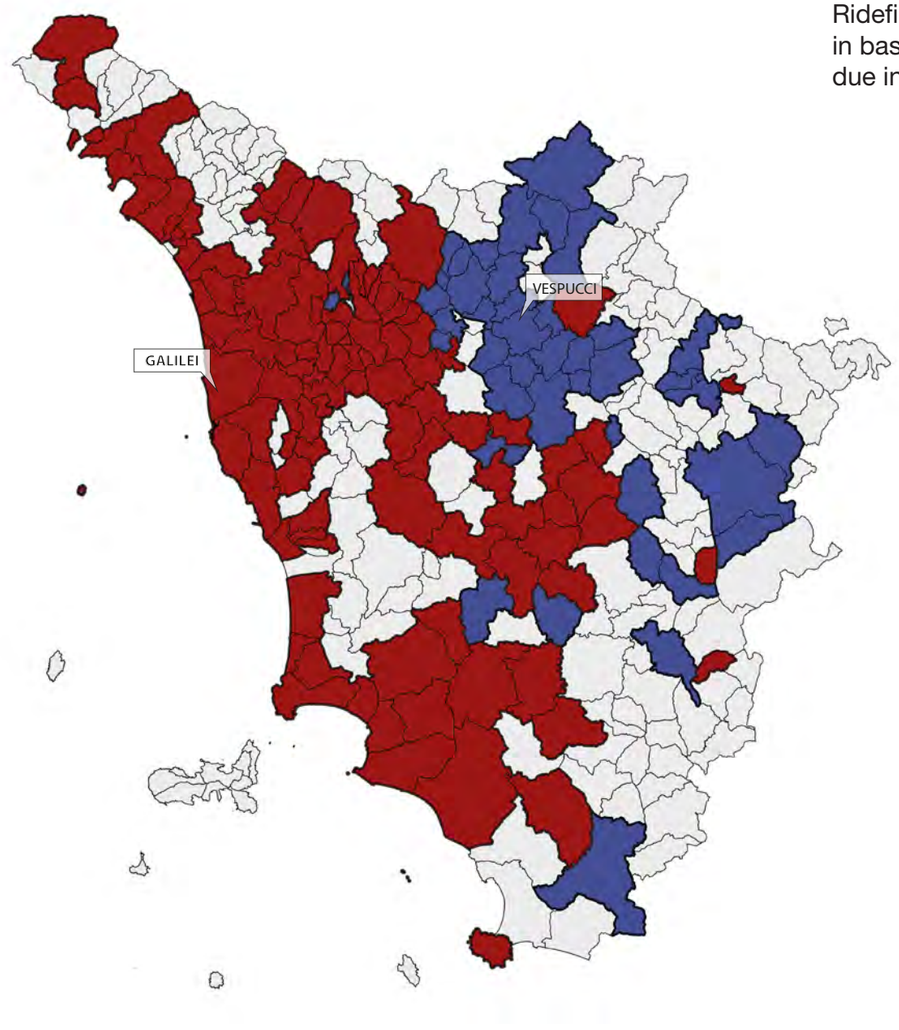
ATTRATTIVITÀ DELL'AEROPORTO VESPUCCI

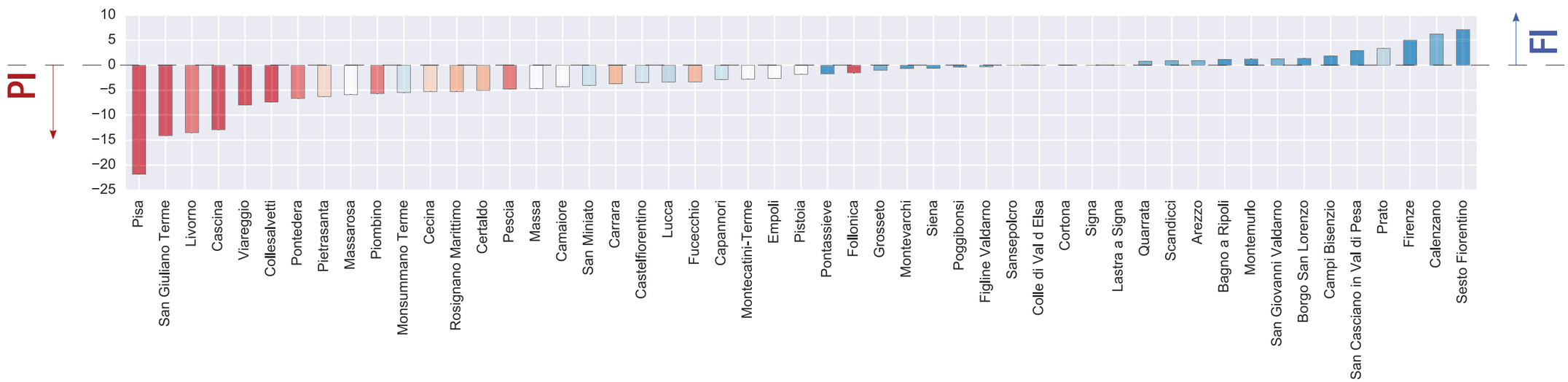


Accessibilità vs preferenza



Ridefinizione del territorio in base all'attrazione delle due infrastrutture





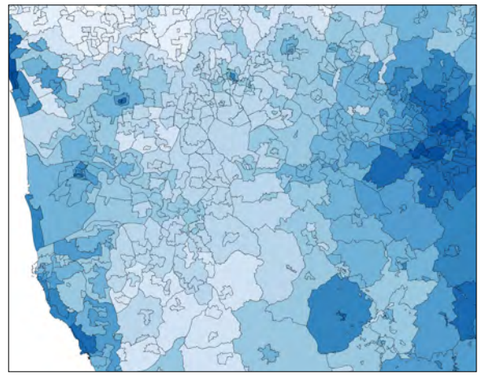
6 CONCLUSIONI E POSSIBILI SVILUPPI

Le analisi proposte finora sono un primo tentativo di verificare come i dati di mobilità possano essere usati per rispondere ad alcune domande che gli enti preposti alla pianificazione territoriale e infrastrutturale tipicamente si pongono. Lungi dal proporre soluzioni esaustive sul complesso tema della pianificazione aeroportuale le analisi finora proposte possono essere un supporto per la progettazione, andando ad integrare e in parte sostituire alcune metodologie usate tradizionalmente. Ad esempio le indagini campionarie sono uno strumento spesso molto costoso e non in grado di catturare i comportamenti occasionali di mobilità. Proprio i viaggi occasionali nel loro insieme risultano però essere una parte importante del traffico osservato ed è fondamentale introdurre questo dato nelle valutazioni. Inoltre i dati GPS, forniscono un livello di granularità molto fino, sono prodotti costantemente e quindi risultano più funzionali rispetto ai censimenti che hanno cadenza decennale e che finora sono stati utilizzati per realizzare alcune analisi di mobilità dei territori.

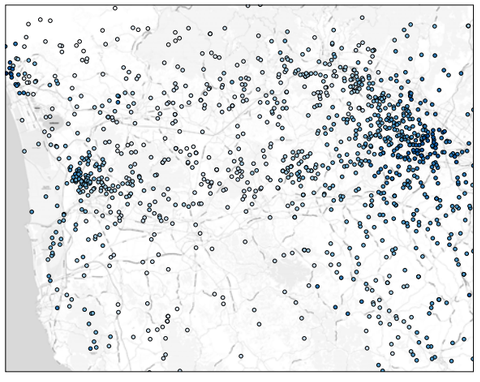
L'arricchimento semantico dei dati di mobilità

Durante il tirocinio, che si è inserito nell'ambito di una più ampia collaborazione tra IRPET e KDD Lab, sono state fatte alcune analisi propedeutiche al proseguimento del progetto di ricerca tra i due enti. Un obiettivo della collaborazione è quello di arricchire semanticamente i dati di mobilità utilizzando variabili socioeconomiche. Sull'esperienza di studi già intrapresi all'interno del KDD Lab (10) (11) si vuole tentare di comprendere se e in quali casi i comportamenti di mobilità possano essere un valido indicatore per descrivere tendenze economiche e comportamenti sociali. Capire quindi se sia possibile individuare dei profili di mobilità che possano essere associati a particolari dinamiche di uso del territorio.
Un primo tentativo in questo senso è stato iniziato realizzando una mappatura della L1 degli utenti del dataset 2011 di OCTO con la corrispondente area OMI (si veda la sezione 3.2 per una descrizione più dettagliata delle OMI). Dal momento che le aree OMI sono un buon indicatore della ricchezza di un territorio è possibile assegnare ad ogni utente un valore che sia rappresentativo della sua ricchezza. Nel fare questa operazione è stato considerato comunque che il valore non può essere strettamente correlato con il reddito degli individui in quanto è legato a dei beni immobili , tuttavia il dato è molto più significativo del valore catastale in quanto si riferisce alle compravendite immobiliari recenti.

Nella prima immagine si può osservare una porzione della regione Toscana ripartita in aree Omi, nella seconda il mapping degli utenti dell'aeroporto con il valore Omi associato alla loro L1. Oltre al mapping del valore OMI lo studio prevede l'assegnazione a tutti gli utenti del dataset di altri valori derivati dalle variabili censuarie disponibili sui portali dell'Istat (popolazione residente, popolazione straniera, famiglie, grado di istruzione, condizione professionale, pendolarismo e altre).



valori medi di compravendita immobiliare per aree OMI (ambito territoriale omogeneo)



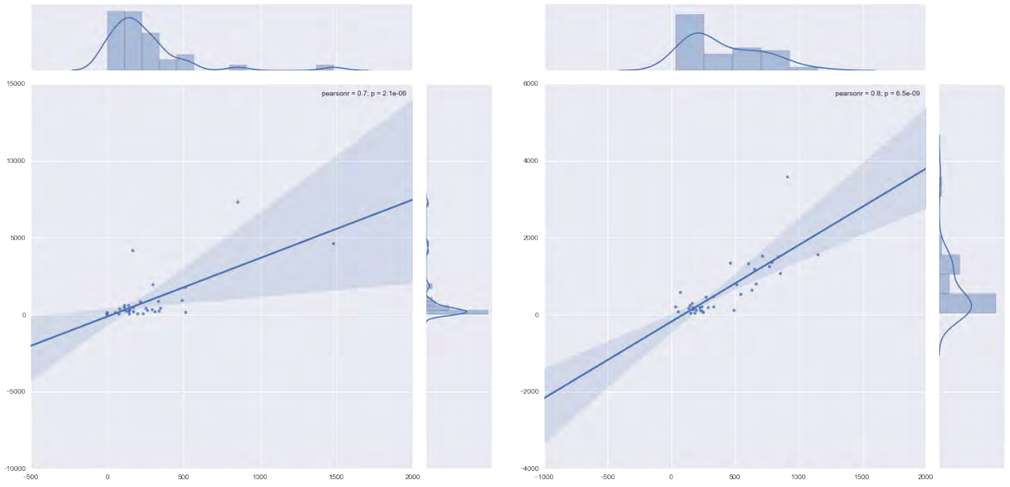
mapping valori OMI con most frequent location (residenza) degli utenti OCTO TELEMATICS

L'integrazione dello studio con i dati del 2014

Al fine di studiare l'evoluzione delle dinamiche osservate tutte le analisi compiute sul dataset del 2011 saranno rieseguite sui dati del 2014. L'area considerata nel 2014 è leggermente più estesa e comprende anche la città di Bologna, sarà possibile analizzare quindi l'influenza dell'aeroporto considerato il più diretto concorrente del Vespucci.

L'integrazione dello studio con i dati GSM

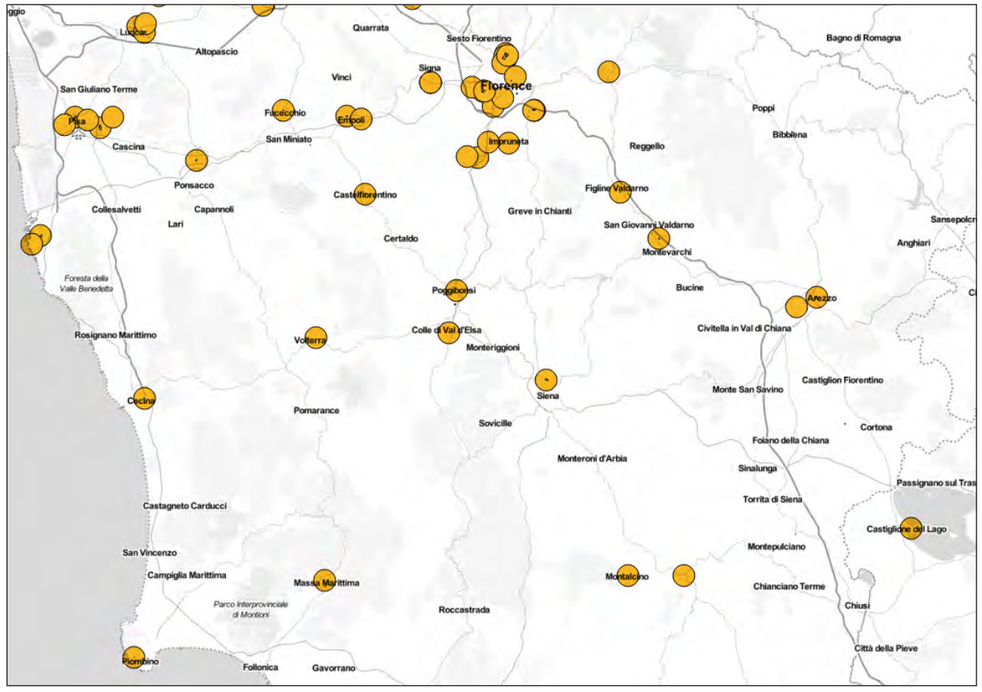
Analisi simili saranno quindi riproposte sui dati GSM forniti da WIND sia nell'anno 2011 che nell'anno 2014. Una prima comparazione delle due fonti di dati è stata fatta correlando il numero di utenti che si recano agli aeroporti da ciascun comune. Eliminando i comuni che sono molto vicini all'area dei due aeroporti è emersa una buona correlazione (0.7 per Firenze e 0.8 per Pisa) che lascia intuire la possibilità di poter confrontare, adottando le dovute cautele, le due fonti di dati.



Correlazione dati GSM - GPS

Nuove tipologie di attrattori

Il passo successivo sarà poi quello di replicare l'analisi su altri grandi attrattori. La prima tipologia presa in considerazione è quella degli ospedali. Grazie alla collaborazione dell'Agenzia Regionale di Sanità della Toscana (ARS) sono stati richiesti i dati di ospedalizzazione delle principali strutture della regione e in attesa della fornitura del dataset si è proceduto con la fase preliminare di identificazione delle geometrie per lo studio della mobilità veicolare intorno a questi nuovi attrattori.



mappatura degli ospedali oggetto del prossimo studio

Il lavoro realizzato finora è stato sottoposto al SEBD2016 (24th Italian Symposium on Advanced Database Systems) con un discussion paper realizzato in collaborazione tra KDD e IRPET dal titolo: Big Data and Public Administration: a case study.

7. APPENDICI

7.1 Strumenti analitici

La fase di data understanding e preparation dei dati GPS è stata effettuata utilizzando principalmente PostgreSQL. Per la costruzione delle traiettorie è stato usato M-ATLAS, il software progettato dal KDD Lab per l'analisi della mobilità basato sul concetto di traiettoria. Nel caso di query più complesse si è affiancato a PostgreSQL il linguaggio di programmazione Python, e in particolare la libreria orientata alla data analysis Pandas. Questa libreria python è stata usata in maniera consistente sia per le analisi esplorative sia nelle fasi successive, grazie alla possibilità di importare e analizzare i dati come dataframe. Il dataset è stato arricchito attraverso l'uso di API google e Open Street Map. Sono state utilizzate in particolare le Distance Matrix di Google Maps per la costruzione di matrici origine/destinazione sia per quanto riguarda il trasporto privato che quello pubblico. Per quanto riguarda invece Open Street Map sono stati usati i plugin OSMDownloader e QuickOSM presenti all'interno di QGis per ottenere le geometrie dei vari punti di interesse.
I risultati sono stati poi visualizzati attraverso QGis per quanto riguarda i dati di tipo geografico e attraverso librerie python per la data visualization quali matplotlib. Inoltre è stata utilizzata la web app RAW per creare visualizzazioni vector-based (SVG) dei flussi. Raw è costrutio sulla libreria D3.js ed è utilizzabile attraverso una semplice interfaccia Web.

7.2 Planning e scheduling del progetto

Il progetto di tirocinio si è inserito nella prima fase di collaborazione tra CNR e IRPET. Lo scopo è stato quello di identificare le fonti dei dati più interessanti e procedere quindi al loro reperimento. Sono stati effettuati diversi incontri per mettere in comune le conoscenze e per comprendere quali fonti di dati potessero essere integrate tra loro, confrontando inoltre le metodologie per lo studio della mobilità.

Data	Luogo	Argomento
5 Ottobre 2015	CNR – Pisa	Inizio Tirocinio
14 Ottobre 2015	IRPET – Firenze	Inserimento in IRPET
22 Ottobre 2015	IRPET – Firenze	Discussione su attività di ricerca formalizzazione collaborazione
27 Ottobre 2015	IRPET – Firenze	Inizio corso di formazione QGIS
9 Novembre 2015	CNR – Pisa	Formalizzazione collaborazione scientifica e prima pianificazione
17 Novembre 2015	IRPET – Firenze	Corso di formazione QGIS attività di ricerca
24 Novembre 2015	IRPET – Firenze	Corso di formazione QGIS attività di ricerca
14 Dicembre 2015	CNR – Pisa	Presentazione dei primi risultati
17 – 18 Dicembre 2015	IRPET – Firenze	Studio delle aree OMI
19 Gennaio 2016	IRPET – Firenze	Presentazione risultati sullo studio degli aeroporti.
29 Gennaio 2016	Master So Big Data	Presentazione del progetto durante la giornata di inaugurazione del Master
2 Marzo 2016	San Giuliano Terme	Workshop KDD Lab Presentazione progetto di collaborazione IRPET CNR
13 Febbraio 2016	CNR – Pisa	Sottomissione del paper a SEBD 2016 - Big Data and Public Administration: a case study

8. RIFERIMENTI

1. ISTI - CNR. http://www.isti.cnr.it/. [Online]
2. IRPET. http://www.irpet.it/. [Online]
3. Octo Telematics. Octo Telematics. [Online] http://www.octotelematics.com/it.
4. Wind Telecomunicazioni . Wind. [Online] http://www.wind.it.
5. Istituto Nazionale di Statistica. ISTAT. [Online] http://www.istat.it/.
6. Automobil Club d'Italia. ACI. [Online] http://www.aci.it/.
7. Ente Nazionale per l'Aviazione Civile. ENAC. [Online] https://www.enac.gov.it/.
8. ENAC. Studio sullo sviluppo degli aeroporti italiani. s.l. : Ministero delle Infrastrutture e dei Trasporti, 2010.
9. Understanding the patterns of car travel. Pappalardo, Rinzivillo, Qu, Pedreschi, Giannotti. Pisa : s.n., 2013.
10. ISTAT. Matrici del pendolarismo. [Online] http://www.istat.it/it/archivio/139381.
11. Big data analytics for smart mobility - a case study. Trasarti, Furletti, Gabrielli, Nanni, Pedreschi
12. Unveiling the complexity of human mobility by querying and mining massive trajectory data - Giannotti, Nanni, Pedreschi, Pinelli, Renso, Rinzivillo, Trasarti
13. Data Science and Simulation in Transportation Research - AA VV, IGI Global
14. Returners and explorers dichotomy in human mobility - Pappalardo, Simini, Rinzivillo, Pedreschi, Giannotti, Barabasi