

CONTRIBUTO DI RICERCA 284/2019

Profili territoriali dell'incidentalità
stradale in Piemonte
2019

Sylvie Ocelli, Alessandro Sciuolo

L'**IRES PIEMONTE** è un ente di ricerca della Regione Piemonte disciplinato dalla Legge Regionale 43/91 e s.m.i. Pubblica una relazione annuale sull'andamento socioeconomico e territoriale della regione ed effettua analisi, sia congiunturali che di scenario, dei principali fenomeni socioeconomici e territoriali del Piemonte.

CONSIGLIO DI AMMINISTRAZIONE

Mario Viano, Presidente
Luca Angelantoni, Vicepresidente
Gianluca Aimaretti, Antonio Amoroso, Lia Fubini

COLLEGIO DEI REVISORI

Maurizio Cortese, Presidente
Paola Dall'Oco e Sara Rolando, Membri effettivi
Annamaria Mangiapelo e Pierangelo Reale, Membri supplenti

COMITATO SCIENTIFICO

Nerina Dirindin, Presidente
Gabriella Agnoletti, Andrea Barasolo, Sergio Conti, Fabrizio Faggiano, Ludovico Monforte, Stefania Ravazzi

DIRETTORE

Marco Sisti

STAFF

Luciano Abburrà, Marco Adamo, Stefano Aimone, Enrico Allasino, Loredana Annaloro, Cristina Aruga, Maria Teresa Avato, Davide Barella, Cristina Bargerò, Stefania Bellelli, Giorgio Bertolla, Marco Carpinelli, Marco Cartocci, Pasquale Cirillo, Renato Cogno, Alessandro Cunsolo, Luisa Donato, Elena Donati, Carlo Alberto Dondona, Fiorenzo Ferlino, Vittorio Ferrero, Claudia Galletto, Anna Gallice, Filomena Gallo, Lorenzo Giordano, Martino Grande, Simone Landini, Federica Laudisa, Sara Macgano, Eugenia Madonia, Maurizio Maggi, Maria Cristina Migliore, Giuseppe Mosso, Daniela Musto, Carla Nanni, Daniela Nepote, Sylvie Occelli, Gianfranco Pomatto, Giovanna Perino, Santino Piazza, Sonia Pizzuto, Elena Poggio, Francesca Silvia Rota, Chiara Rivoiro, Valeria Romano, Martina Sabbadini, Lucrezia Scalzotto, Bibiana Scelfo, Luisa Sileno, Alberto Stanchi, Filomena Tallarico, Guido Tresalli, Stefania Tron, Roberta Valetti, Giorgio Vernoni.

COLLABORANO

Filomena Berardi, Stefano Cavaletto, Elisabetta Cibiniel, Salvatore Cominu, Simone Contu, Giovanni Cuttica, Lorenzo Fruttero, Silvia Genetti, Gabriella Gianoglio, Enrico Gottero, Giulia Henry, Veronica Ivanov, Ludovica Lella, Luigi Nava, Serena Pecchio, Ilaria Perino, Stefano Piperno, Samuele Poy, Francesca Prunotto, Alessandro Sciallo, Paolo Saracco, Antonio Soggia, Anda Tarbuna, Nicoletta Torchio, Elisa Tursi, Silvia Venturelli, Paola Versino, Gabriella Vberti, Augusto Vino, Paolo Zeppetella.

Il documento in formato PDF è scaricabile dal sito www.ires.piemonte.it

La riproduzione parziale o totale di questo documento è consentita per scopi didattici, purché senza fine di lucro e con esplicita e integrale citazione della fonte.

© 2019 IRES – Istituto di Ricerche Economico-Sociali del Piemonte
via Nizza 18 – 10125 Torino
www.ires.piemonte.it

PROFILI TERRITORIALI DELL'INCIDENTALITÀ STRADALE IN PIEMONTE

© 2019 IRES
Istituto di Ricerche Economico Sociali del Piemonte
Via Nizza 18 - 10125 Torino

www.ires.piemonte.it

GLI AUTORI

Sylvie Occelli, Alessandro Sciuolo

INDICE

GLI AUTORI	IV
SINTESI	7
1. MOTIVAZIONI DELLO STUDIO	10
1.1 Studiare le determinanti dell'incidentalità per una mobilità sicura e sostenibile	10
1.2 Perché la prospettiva territoriale è rilevante	10
1.3 Domande di ricerca e organizzazione del lavoro	13
1.3.1 Costruzione di una base informativa integrata.....	13
1.3.2 Una lettura del fenomeno incidentale attraverso la costruzione di profili territoriali	14
2. IL PERCORSO METODOLOGICO	18
2.1 Costruire una base informativa integrata	18
2.1.1 Le basi dati	18
2.1.2 Integrazione delle basi dati	19
2.2 Valorizzazione dei segmenti del DB strade	19
2.3 Collegamento dei database e funzioni di Geoprocessing	20
2.3.1 Unione dei DB incidenti con il DB sezioni censimento	21
2.3.2 Unione dei DB incidenti con il DB Strade	22
2.4 La Cluster Analysis per identificare i profili territoriali.....	24
2.4.1 Che cos'è la cluster analysis	24
BOX 1. L'algoritmo di clustering	25
2.4.2 L'applicazione della cluster analysis all'analisi dell'incidentalità	25
3. COSA HA PRODOTTO LO STUDIO	27
3.1 Una base dati integrata per la sicurezza stradale (DBSS)	27
3.2 I profili territoriali dell'incidentalità.....	30
3.2.1 La tipologia dei comuni	30
3.2.2 Il database degli archi stradali incidentati: un esempio applicativo del dbss	35
3.2.3 Casi d'uso.....	38
Caso 1 – Quanti sono i tratti di strada ad alta incidentalità (con più di 4 incidenti ? Come si distribuiscono tra i profili comunali individuati dalla Cluster Analysis?	38
Caso 2 – Come si caratterizzano gli eventi incidentali sui segmenti di strada ad alta incidentalità, dal punto di vista della natura dell'evento e del luogo in cui avvengono?	38
Caso 3 – Come può essere utilizzato il DB per esaminare le criticità incidentali a livello sub-comunale?....	39
4.PROSPETTIVE DI APPROFONDIMENTO	42
4.1 Ottimizzare il metodo di integrazione basi dati	42
4.2 Approfondire l'analisi	43
BIBLIOGRAFIA	44

SINTESI

Questo studio prende le mosse da un precedente lavoro dell'Ires Piemonte nel quale era stata condotta un'investigazione del rischio incidentale (numero di incidenti in rapporto al livello di traffico) sulla rete stradale di interesse regionale di I e II livello.

Qui si arricchisce il percorso di analisi seguito in quel lavoro e se ne approfondiscono alcuni aspetti, cercando di dare risposta alle domande che hanno guidato la realizzazione del lavoro:

- come e in che misura l'incidentalità stradale di un'area è legata al suo profilo insediativo e socioeconomico?
- è possibile individuare dei profili territoriali dell'incidentalità e cioè qualificare le sub aree regionali in relazione alla presenza del fenomeno incidentale?

Per rispondere a queste domande lo studio:

- costruisce una base informativa, inedita, che integra alcune informazioni attinenti il fenomeno incidentale, esistenti per il Piemonte ma fino ad oggi utilizzate in modo separato;
- propone una lettura delle situazioni incidentali del territorio piemontese, realizzata attraverso la costruzione di una tipologia di *profili incidentali*.

Una base informativa integrata

Le principali basi informative utilizzate sono:

1. il Database (d'ora in avanti Db) Incidenti che raccoglie i dati sul singolo evento incidentale, per il periodo 2013-16 (Istat, Regione Piemonte);
2. Db Sez. Censimento (al 2011) che contiene una selezione di variabili socioeconomiche e riguardanti l'edificato a livello di sezione censuaria (Istat);
3. Db Strade che contiene i dati sulle caratteristiche geometriche, sui flussi di traffico e le velocità, per arco della rete di viabilità regionale (Regione Piemonte, TOC);
4. alcuni Db comunali riguardanti la mobilità, il consumo di suolo, ecc., che provengono da varie fonti (Istat, CSI, Regione Piemonte, ARPA, Ministero dei Trasporti).

Per integrare queste basi dati (ma questo vale per qualsiasi altra base dati) è necessario avere un identificativo comune che consenta di associare le informazioni ivi contenute. In questo studio, il collegamento è realizzato considerando il più basso livello di osservazione (diverso da quello comunale), adottato in ciascun Db. Si considerano pertanto il singolo incidente, nel Db incidente, la sezione censuaria nel Db socioeconomico e dell'edificato, l'arco stradale nel Db strade.

Nello specifico: ogni singolo incidente, è associato all'arco stradale in cui si è verificato e alla sezione di censimento in cui l'evento si è prodotto.

Questa operazione, semplice da un punto di vista concettuale, ha richiesto però una serie di elaborazioni piuttosto onerose dal punto di vista computazionale. Infatti, è stato necessario **valorizzare tutti i segmenti del DB strade e collegare i database attraverso i dati di georeferenziazione** presenti nelle basi dati.

A conclusione di queste elaborazioni è stato predisposto un Data-Base integrato, per la Sicurezza Stradale (DBSS) che collega il DB Incidenti con quello delle Strade e delle Sezioni censuarie.

Una lettura del fenomeno incidentale attraverso la costruzione di profili territoriali

Per identificare i profili territoriali dell'incidentalità, questo lavoro utilizza la Cluster Analysis, (in particolare una tecnica ben consolidata, la k-means). L'unità di osservazione è il comune, caratterizzato secondo un certo numero di variabili predisposte (per la quasi totalità) a partire dai dati individuali contenuti nel DBSS.

Le variabili riguardano:

- a. **il profilo incidentale**: cioè l'incidenza e le caratteristiche del fenomeno incidentale ;
- b. **il profilo insediativo**: l'utilizzo del suolo ;
- c. **il profilo infrastrutturale**: il livello e l'intensità del traffico sulla rete viaria e le modalità di mobilità .

L'applicazione della Cluster Analysis ha identificato cinque gruppi (tipi) di comuni. La loro distribuzione nel territorio regionale suggerisce che esiste un legame, tra profili incidentali e morfologia insediativa. I nomi dei cluster riflettono questo legame.

- **Periferico (11% dei comuni)**: situato in aree periferiche e/o interstiziali del territorio regionale, ha valori bassi di infrastrutturazione e di criticità incidentale, in termini sia di rischio, numero di infortunati (feriti più morti) rispetto alla popolazione, sia di livello di infortunio (numero feriti e di morti) per 100 incidenti. Molto rilevante la quota di infortunati in ambito urbano. La quota di mobilità che usa il TPL è superiore alla media regionale.
- **Periferico Rischio Elevato (RE) (32% dei comuni)**: situato in aree poco densamente abitate, ha una bassa infrastrutturazione ma un'alta criticità incidentale in termini sia di rischio sia di infortunio. Traffico e incidenti si concentrano soprattutto in ambito extra-urbano. Anche in questo gruppo, la quota di mobilità che usa il TPL è superiore alla media regionale. Pur concentrando poco meno del 5% degli incidenti che avvengono nella regione, la quota di morti è quasi tre volte più elevata (14%).
- **Hinterland (11% dei comuni)**: si colloca soprattutto lungo le direttrici più importanti. Ha una modesta infrastrutturazione, ma la presenza di strade a scorrimento veloce è maggiore che negli altri gruppi. Infortunati e volume di traffico si concentrano soprattutto in area extra-urbana.
- **Poli Urbanizzati (16% dei comuni)**: Include i centri regionali più popolosi. Ha valori elevati di infrastrutturazione e di densità dell'edificato. In questo gruppo si concentrano il 78% degli incidenti e il 51% dei morti. Il rischio di incidentalità e di infortunio hanno valori contenuti. Quasi tre infortuni su quattro sono causati da sinistri in ambito urbano.
- **Periurbani (30% dei comuni)**: situato in prossimità dei comuni maggiori, ha un profilo di infrastrutturazione simile a quello medio regionale, associato però a livello di rischio (numero di infortunati rispetto alla popolazione) superiore. La quota di infortunati è quasi egualmente distribuita tra ambito urbano ed extra-urbano.

A livello sub-regionale, molti comuni delle province di Torino, Novara, Biella e VCO appartengono ai gruppi **Poli Urbanizzati** e **Periurbani**; insieme questi tipi caratterizzano oltre il 50% dei loro comuni. Il gruppo **Periferico a Rischio Elevato** si concentra soprattutto nelle province di Asti, Cuneo, Alessandria. Il gruppo Hinterland è maggiormente presente a Vercelli e Novara ed è assente nella provincia di Biella.

Come possono essere utilizzati e a cosa possono servire i risultati dello studio?

La predisposizione di un **DB integrato per la sicurezza stradale** è un pre-requisito indispensabile per progettare i futuri interventi di messa in sicurezza della rete regionale e governarne le ricadute. In questa direzione, come previsto a suo tempo nel progetto del CMRSS piemontese, del 2008, il DBSS sviluppato in questo lavoro, ne costituisce una realizzazione prototipale.

Per esemplificare le sue possibilità di utilizzo, è stata predisposta una nuova base dati organizzata per **segmenti di strada** che contiene, oltre ai dati incidentali, alcune informazioni sul profilo infrastrutturale e territoriale. Essa è costituita da 27.902 segmenti che rappresentano il 5,6% dei tratti dell'intera rete della viabilità regionale.

Un esame delle informazioni per arco stradale e per sezione censuaria mostra, ad esempio, come sui tratti con un numero elevato di incidenti, il traffico sia mediamente più intenso e più veloce. Esso evidenzia inoltre come i tratti a elevata incidentalità attraversino aree caratterizzate da un tessuto insediativo più denso.

La base dati per segmenti di strada è facilmente consultabile attraverso un semplice foglio elettronico (Excel) e nel testo si presentano alcuni esempi di utilizzo (casi d'uso). Pur con tutti i caveat del caso, si auspica che essa possa essere messa a disposizione (e testata) degli operatori della sicurezza stradale che operano in regione.

Studiare i casi d'uso tenendo conto del punto di vista di questi operatori, potrebbe fornire indicazioni utili adattare meglio le funzionalità del DBSS, rispetto alle attività stesse che gli operatori dovranno realizzare per mettere in opera i piani settoriali del nuovo Piano regionale della mobilità e dei trasporti.

La **dimensione territoriale** è un riferimento sempre più importante nei progetti di sicurezza stradale:

- è un background necessario per qualificare le **criticità incidentali** della rete infrastrutturale;
- è un supporto attraverso il quale **cogliere la varietà del fenomeno incidentale** nelle diverse parti del territorio regionale;
- è una componente di una piattaforma comunicativa, per **condividere le conoscenze sul fenomeno** incidentale da parte dei cittadini, degli stakeholders e dei decisori.

La costruzione di una **tipologia territoriale di profili incidentali** risponde a queste esigenze. Essa è un contributo analitico che serve a rafforzare la capacità di progettare interventi di policy meglio rispondenti ai problemi locali di incidentalità.

Naturalmente, la tipologia identificata in questo studio ha ampi margini di miglioramento, soprattutto per quanto riguarda le **relazioni tra determinanti di incidentalità e livelli di rischio incidentale**.

1. MOTIVAZIONI DELLO STUDIO

1.1 Studiare le determinanti dell'incidentalità per una mobilità sicura e sostenibile

Questo studio prende le mosse da un precedente lavoro dell'Ires Piemonte (Landini, Occelli e Scalzotto, 2018) sull'evoluzione del fenomeno incidentale in Piemonte nel periodo 2010-2015. In quel lavoro si ricordava come la sicurezza stradale fosse stata inclusa fra gli obiettivi per lo sviluppo sostenibile individuati dall'agenda 2030 delle Nazioni Unite, i cosiddetti Sustainable Development Goals (SDG).¹ L'incidentalità, infatti, ha una ricaduta diretta sulla sostenibilità dei sistemi socio-economici:

- ha un impatto diretto sulla salute pubblica e sui costi per la sua tutela (SDG 3);
- influenza la qualità degli insediamenti umani in termini di requisiti di inclusione, accessibilità e tutela degli utenti della strada più vulnerabili (i pedoni e i ciclisti) (SDG 11).

Disporre di descrizioni appropriate del fenomeno incidentale e formulare ipotesi teoriche e metodologiche per lo studio delle sue caratteristiche, pertanto, è indispensabile per definire interventi di policy, meglio capaci di:

- contrastare il fenomeno, riducendo il numero di incidenti e la gravità;
- migliorare la sostenibilità delle comunità locali, aumentando la consapevolezza del valore sociale della sicurezza stradale e promuovendo iniziative formative di prevenzione del fenomeno.

In questa direzione, il citato studio dell'Ires aveva condotto una prima investigazione del rischio incidentale (numero di incidenti in rapporto al livello di traffico) sulla rete di interesse regionale di I e di II livello. Questo lavoro arricchisce il percorso di analisi seguito in quello studio e ne approfondisce un certo numero di aspetti. Più precisamente:

- affina la lettura del fenomeno, ampliando il set delle variabili ritenute rilevanti nel determinare un certo livello di rischio incidentale;
- estende la base informativa, includendo variabili nuove per rappresentare le determinanti del fenomeno incidentale;
- introduce una certa flessibilità nel focus di osservazione: dà la possibilità, cioè, di osservare il fenomeno incidentale sia a livello regionale o sub-regionale, sia a livello comunale, sub-comunale e del singolo tratto stradale, consentendo di coglierne meglio la varietà delle manifestazioni a livello locale;
- irrobustisce il percorso analitico utilizzando metodi di geomatica (GIS) e di analisi statistica (Cluster Analysis) per il trattamento delle informazioni.

1.2 Perché la prospettiva territoriale è rilevante

L'incidente è un evento che, per definizione, è unico, imprevedibile, casuale, essendo determinato dall'interazione in uno stesso luogo e in uno stesso momento di un insieme di condizio-

¹ <https://sustainabledevelopment.un.org/?menu=1300>

ni che riguardano: i comportamenti degli utenti della strada, le caratteristiche dei veicoli, lo stato delle strade, il contesto insediativo, la situazione climatica .

I fattori che incidono su tali condizioni e, dunque, sulla probabilità del verificarsi di un incidente, il cosiddetto rischio incidentale, sono molteplici ed eterogenei (Hermans et al, 2008).²

Essi possono essere raggruppati secondo le seguenti categorie descrittive che servono anche a identificare le determinanti del fenomeno incidentale:

D1. Contesto territoriale

- geomorfologia
- livello funzionale e condizione delle infrastrutture
- struttura insediativa e uso del suolo
- flussi di traffico e servizi di trasporto disponibili
- variabili socio-demografiche e struttura economica.

D2. Tecnologia

- qualità e prestazioni del veicolo
- dispositivi per la sicurezza individuali e ambientali
- altri dispositivi potenziali cause di riduzione/aumento del rischio (per esempio, computer di bordo e cellulari).

D3. Componente umana

- età e caratteristiche cognitive
- comportamenti alla guida e stili di vita
- competenze di guida.

Tra queste, la categoria descrittiva relativa al territorio (D1) (e alle sue variabili strutturali), ha oggi un'importanza ancor più grande che in passato (Eksler e Lassare, 2008, Eksler, 2008, 2010, WHO, 2004, 2006). Essa, infatti, è una dimensione costitutiva essenziale per sviluppare un nuovo approccio all'analisi dell'incidentalità, che potremmo definire di meso-livello.

E' questo un approccio che si colloca in posizione intermedia tra:

- una lettura generale (macro) che privilegia la valutazione della consistenza/gravità del fenomeno incidentale in un certo luogo (regione, paese, città)³;
- e una lettura puntuale (micro) che guarda in profondità all'evento incidentale quale accade su un certo tratto di strada (*in-depth analysis* che studia la dinamica e i comportamenti dei veicoli/conducenti in un certo evento) (Larsen e Kines, 2002, Beanland et al, 2013).

L'opportunità di valorizzare la dimensione territoriale per sviluppare questo nuovo approccio si motiva alla luce di alcuni requisiti generali, che sono funzionali per realizzare gli interventi di sicurezza stradale. Più precisamente, questa dimensione:

- è un background geografico indispensabile per **qualificare le criticità incidentali** della rete infrastrutturale regionale;

² Il Piano Nazionale per la Sicurezza Stradale 2014, ad esempio, individua i seguenti fattori di rischio: velocità, condizioni ambientali (visibilità e stato dell'infrastruttura), vulnerabilità e uso di dispositivi di sicurezza, guida in stato di ebbrezza o sotto l'effetto di farmaci, stanchezza e stress, mancanza esperienza, massa e forma dei veicoli

³ I rapporti dell'European Transport Society Council sono un esempio di questa lettura macro (ad esempio, ETSC, 2018).

- è un riferimento analitico attraverso il quale **cogliere la varietà del fenomeno incidentale** nelle diverse parti del territorio regionale;
- è una componente di una piattaforma comunicativa per **condividere le conoscenze sul fenomeno** incidentale, da parte dei cittadini, degli stakeholders e dei decisori.

Esula dagli scopi di questo studio entrare nel merito delle caratteristiche e delle potenzialità analitiche dell'approccio di meso-livello. Alla luce dell'esperienza acquisita negli studi dell'Ires e delle caratteristiche attese delle nuove politiche per la sicurezza stradale⁴, possiamo tuttavia menzionarne alcuni requisiti, quali;

- il fatto di tener conto delle diverse determinanti dell'incidentalità e, al tempo stesso, di coglierne le interazioni possibili nei diversi contesti locali;
- Il fatto di adottare un approccio metodologico, scalabile, flessibile e articolato a diverse scale territoriali;
- Il fatto di non imporre aprioristicamente alcun metodo analitico, ma di riconoscere la possibilità di usare, in modo eventualmente integrato, una certa pluralità di metodi diversi per misurare, descrivere e governare un fenomeno.

In questo studio il territorio è, dunque, al cuore del lavoro. E' alla base di una categoria descrittiva indispensabile per investigare il fenomeno incidentale. Al tempo stesso, da un punto di vista sostantivo, è il riferimento necessario nei progetti per la messa in sicurezza della rete stradale, e per il governo dei determinanti stessi dell'incidentalità.

Porre il territorio al centro dell'analisi ha poi un indubbio vantaggio pratico, non secondario, di rendere disponibile un modo per integrare le informazioni.

La georeferenziazione dei dati, un attributo oggi sempre più diffuso, consente, tramite appropriate funzionalità tecnico informatiche⁵, di operare il collegamento tra domini e fonti informative diversi. Si tratta di una potenzialità analitica straordinaria: Essa permette di:

- mettere in relazione i diversi attributi di un fenomeno;
- registrarne i dati in archivi informativi convenientemente organizzati;
- abilitare l'uso di una certa molteplicità di possibili lenti di osservazione del fenomeno stesso.

⁴ Si veda in proposito lo studio dell'Ires che si interroga sui temi delle politiche per la sicurezza nei prossimi anni (Montaldo e Occelli, 2018).

⁵ http://www.treccani.it/enciclopedia/geomatica-per-la-gestione-del-territorio_%28Enciclopedia-della-Scienza-e-della-Tecnica%29/

1.3 Domande di ricerca e organizzazione del lavoro

Il riconoscimento della centralità del territorio si sostanzia in **due domande** che hanno guidato la realizzazione del lavoro:

- come e in che misura l'incidentalità stradale di un'area è legata al suo profilo insediativo (uso del suolo, rete viaria) e socioeconomico (caratteristiche della popolazione e delle attività economiche, modi di spostamento)?
- è possibile individuare dei profili territoriali dell'incidentalità e cioè ; in altre parole, è possibile qualificare le sub aree regionali in relazione alla presenza del fenomeno incidentale?

Per rispondere a queste domande lo studio costruisce una tipologia territoriale delle situazioni incidentali (Adanu et al 2017, Kaygisiz et al 2017, McAndrews et al. 2016). Il lavoro si concretizza in due principali prodotti:

- una base informativa, inedita, che integra alcune basi attinenti il fenomeno incidentale, esistenti per il Piemonte ma fino ad oggi utilizzate in modo separato;
- una lettura delle situazioni incidentali del territorio piemontese, realizzata attraverso la costruzione di una tipologia di *profili incidentali*. Si ritiene che questa tipologia (e i suoi affinamenti futuri) sia (no) un contributo analitico utile per progettare (e valutare⁶) interventi di policy a livello locale.

1.3.1 Costruzione di una base informativa integrata

I dati sull'incidentalità e quelli riguardanti le sue determinanti (contesto territoriale, tecnologia, comportamenti) sono contenuti in numerose basi dati, diverse per fonte istituzionale, struttura organizzativa e/o livello territoriale di riferimento.

In questo studio, le principali basi informative utilizzate sono:

- il Db Incidenti che raccoglie i dati sul singolo evento incidentale, per il periodo 2013-16 (Istat, Regione Piemonte);
- Db SezCensimento (al 2011) che contiene una selezione di variabili socioeconomiche e riguardanti l'edificato a livello di sezione censuaria (Istat);
- Db Strade che contiene i dati sulle caratteristiche geometriche, sui flussi di traffico e le velocità, per arco della rete di viabilità regionale (Regione Piemonte, TOC);
- alcuni Db comunali riguardanti la mobilità, il consumo di suolo, ecc., che provengono da varie fonti (Istat, CSI, Regione Piemonte, ARPA, Ministero dei Trasporti).

Per integrare queste basi dati (ma questo vale per qualsiasi altra base dati) è necessario avere un identificativo comune che consenta di associare le informazioni ivi contenute. Nel precedente lavoro dell'Ires (Landini, Occelli e Scalzotto, 2018), questo era stato fatto scegliendo un livello di osservazione sufficientemente alto in modo da poter predisporre una chiave di collegamento comune alle varie basi dati (che in quello studio era una stringa formata dal codice del comune e dal nome della strada).

In questo studio, il collegamento è realizzato considerando il più basso livello di osservazione, (diverso da quello comunale), adottato in ciascun Db. Si considerano pertanto il singolo inci-

⁶ Un uso pratico di questa tipologia, ad esempio, è quello di poter fare da riferimento nella selezione dei casi da includere nel gruppo di controllo in un approccio di valutazione.

dente, nel Db incidente, la sezione censuaria nel Db socioeconomico e dell'edificato, l'arco stradale nel Db strade.

Nello specifico: ogni singolo incidente, è associato all'arco stradale in cui si è verificato e alla sezione di censimento in cui l'evento si è prodotto.

Questa operazione, semplice da un punto di vista concettuale, ha richiesto però una serie di elaborazioni piuttosto onerose dal punto di vista computazionale. Infatti, è stato necessario:

- valorizzare i tutti i segmenti del DB strade;
- collegare i database attraverso i dati di georeferenziazione presenti in tutte le basi dati.

A conclusione di queste elaborazioni è stato predisposto un Data-Base integrato, per la Sicurezza Stradale (DBSS) che collega il DB Incidenti con quello delle Strade e delle Sezioni censuarie. Esso può essere visto come il **nocciolo di un sistema scalabile** e rappresenta una realizzazione del database previsto a suo tempo nel progetto del CMRSS piemontese, del 2008.

Per esemplificare le possibilità di utilizzo del DBSS, è stata ricavata una nuova base dati organizzata per **segmenti di strada** che contiene oltre ai dati incidentali alcune informazioni sul profilo infrastrutturale e territoriale. Essa consiste di 27902 segmenti che rappresentano il 5,6% dei tratti dell'intera rete della viabilità regionale.

Un esame delle informazioni per arco stradale e per sezione censuaria mostra, ad esempio, come sui tratti con un numero elevato di incidenti, il traffico sia mediamente più intenso e più veloce. Esso evidenzia inoltre come i tratti a elevata incidentalità attraversino aree caratterizzate da un tessuto insediativo maggiormente denso (la quota di edifici multi-piano e di quelli non residenziali è più elevata).

La base dati per segmenti di strada è facilmente consultabile attraverso un semplice foglio elettronico (Excel) e nel testo si presentano alcuni esempi di utilizzo (casi d'uso).

Si auspica che, già in questa forma, essa possa essere usata dagli operatori regionali della sicurezza stradale, a fine di consultazione e/o di progettazione di interventi per la messa in sicurezza della rete stradale. A questo proposito, si segnala che **studiare i casi d'uso** tenendo conto del punto di vista di questi operatori, potrebbe essere una pista di lavoro da seguire: non solo per migliorare le funzionalità del DBSS, ma, soprattutto, per renderle più appropriate rispetto alle attività previste dai piani settoriali del Piano regionale della mobilità e dei trasporti.

1.3.2 Una lettura del fenomeno incidentale attraverso la costruzione di profili territoriali

In questo studio, costruire un profilo territoriale dell'incidentalità significa dare evidenza alle ipotesi seguenti:

- **H1: l'incidentalità non è un fenomeno omogeneo** nel territorio regionale. Non solo è diversa tra ambito urbano e ambito extra-urbano e questo è noto da tempo alla luce dei dati incidentali; ma è anche diversa tra comune e comune, un aspetto questo anch'esso messo in luce dai risultati delle analisi di quei dati, ma non ancora argomentato in modo sufficientemente soddisfacente;
- **H2: le caratteristiche insediative dei comuni influenzano il fenomeno incidentale**. Studiarne gli attributi e come questi influenzano il fenomeno è utile pertanto per governare meglio i determinanti territoriali dell'incidentalità; può aiutare, inoltre, a progettare

con maggiore attenzione gli interventi di messa in sicurezza della rete stradale, in modo da migliorarne l'efficacia.

Per identificare i profili territoriali dell'incidentalità, il lavoro utilizza la Cluster Analysis, (in particolare una tecnica ben consolidata, la k-means), che serve a raggruppare un insieme di unità di osservazione, sulla base della somiglianza degli attributi, preventivamente selezionati per descrivere le unità di osservazione.

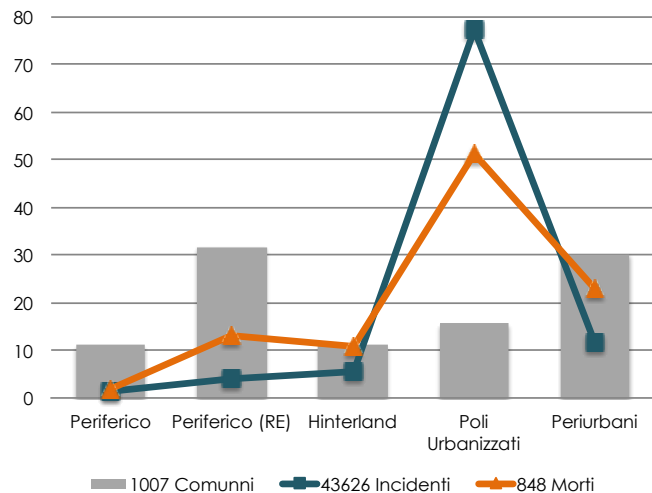
In questo studio, l'unità di osservazione è il comune, caratterizzato secondo un certo numero di variabili che riguardano:

- l'incidenza e le caratteristiche del fenomeno incidentale (profilo incidentale);
- la densità del tessuto urbanizzato (profilo insediativo);
- il livello e l'intensità del traffico sulla rete viaria e la quota di spostamenti con TPL (profilo infrastrutturale).

Tutte le variabili ad eccezione della quota di spostamenti sono state predisposte a partire dai dati individuali contenuti nel DBSS.

L'applicazione della Cluster Analysis ha identificato cinque gruppi (tipi) di comuni. La distribuzione geografica dei gruppi nel territorio regionale suggerisce che esiste un legame, tra profili incidentali e morfologia insediativa. I nomi dei cluster riflettono questo legame.

- **Periferico (11% dei comuni):** situato in aree periferiche e/o interstiziali del territorio regionale, ha valori bassi di infrastrutturazione e di criticità incidentale, in termini sia di rischio, numero di infortunati (feriti più morti) rispetto alla popolazione, sia di livello di infortunio (numero feriti e di morti) per 100 incidenti. Molto rilevante la quota di infortunati in ambito urbano. La quota di mobilità che usa il TPL è superiore alla media regionale.
- **Periferico Rischio Elevato (RE) (32% dei comuni):** situato in aree poco densamente abitate, ha una bassa infrastrutturazione ma un'alta criticità incidentale in termini sia di rischio sia di infortunio. Traffico e incidenti si concentrano soprattutto in ambito extra-urbano. Anche in questo gruppo, la quota di mobilità che usa il TPL è superiore alla media regionale. Pur concentrando poco meno del 5% degli incidenti totali, la quota di morti è quasi tre volte più elevata (14%).
- **Hinterland (11% dei comuni):** si colloca soprattutto lungo le direttrici più importanti. Ha una modesta infrastrutturazione, ma una presenza di strade a scorrimento veloce più alta che negli altri gruppi. Infortunati e volume di traffico si concentrano soprattutto in area extra-urbana.
- **Poli Urbanizzati (16% dei comuni):** Include i centri regionali più popolosi. Ha valori elevati di infrastrutturazione e di densità dell'edificato. In questo gruppo si concentrano il

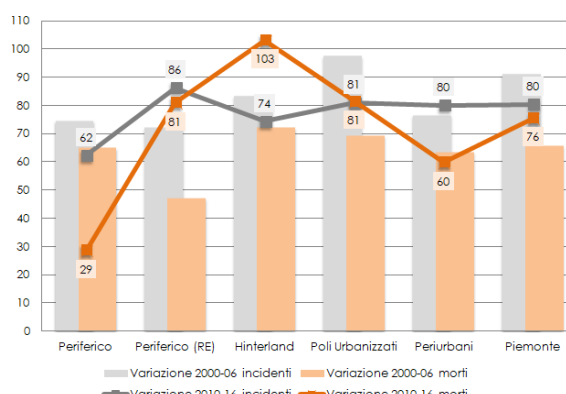


78% degli incidenti e il 51% dei morti. Il rischio di incidentalità e di infortunio hanno valori contenuti. Quasi tre infortuni su quattro sono causati da sinistri in ambito urbano.

- **Periurbani (30% dei comuni):** situato in prossimità dei comuni maggiori, ha un profilo di infrastrutturazione simile a quello medio regionale, associato però a livello di rischio (numero di infortunati rispetto alla popolazione) superiore. La quota di infortunati è quasi egualmente distribuita tra ambito urbano ed extra-urbano.

Nel periodo 2013-16, il 17% dei comuni piemontesi non ha avuto incidenti sulle proprie strade. A livello sub-regionale, molti comuni delle province di Torino, Novara, Biella e VCO appartengono ai gruppi **Poli Urbanizzati** e **Periurbani**; insieme questi tipi caratterizzano oltre il 50% dei loro comuni. Il gruppo **Periferico a Rischio Elevato** si concentra soprattutto nelle province di Asti, Cuneo, Alessandria. Il gruppo Hinterland è maggiormente presente a Vercelli e Novara ed è assente nella provincia di Biella.

La figura qui a lato confronta, per i diversi gruppi, la variazione del numero di incidenti e di morti⁷ nei primi sei anni, dei decenni 2000-10 e 2010-16. Se il confronto indica che le variazioni nel periodo più recente sono inferiori (o simili) a quelle del periodo precedente, allora è ragionevole supporre che il ridimensionamento del fenomeno incidentale stia continuando. E' questo il caso dei comuni che appartengono ai gruppi **Periferico** e **Periurbani**, dove il miglioramento della mortalità (e dell'incidentalità) in questo decennio è superiore o comunque non inferiore a quello osservato nel decennio scorso.



Diversamente, variazioni del periodo più recente, superiori a quelle osservate nel periodo precedente potrebbero indicare che esistono delle criticità nel percorso di miglioramento. E' questo il caso dei comuni che appartengono ai gruppi **Hinterland**, **Poli Urbanizzati** e in particolare al gruppo **Periferico RE**.

1.3.4 Indicazioni per approfondimenti futuri

Come molti studi che sperimentano percorsi innovativi di analisi, questo lavoro ha ampi margini di miglioramento.

Pur con tutti i caveat del caso, si auspica che i suoi risultati possano essere utilizzati, fin da ora, nella pratica della gestione del fenomeno incidentale, a fini di analisi e di progettazione degli interventi. Gli esiti di queste applicazioni, infatti, potranno costituire test molto utili per orientare le priorità degli approfondimenti analitici da realizzare nei prossimi mesi.

Oltre a quanto già menzionato più sopra circa lo **studio dei casi d'uso**, possiamo menzionare due piste possibili per migliorare lo **studio delle relazioni tra determinanti di incidentalità e livelli di rischio incidentale**:

⁷ I dati di questo confronto includono anche gli incidenti avvenuti sulle autostrade.

- esplorare altre metodologie di clustering che tengano conto della dinamicità dei processi; che sappiano cioè costruire la tipologia non solo sulla base dei valori delle variabili descrittive selezionate, ma anche della loro evoluzione nel tempo in relazione ai cambiamenti del conteso territoriale di appartenenza;
- utilizzare tecniche di regressione lineare (e spaziale) per studiare l'intensità delle relazioni tra determinanti e rischi incidentali e la loro variabilità in relazione alle diverse sub-aree regionale (ad esempio gli ambiti di trasporto).

2. IL PERCORSO METODOLOGICO

2.1 Costruire una base informativa integrata

Questa operazione si è articolata nei passi seguenti:

- ricognizione delle basi dati esistenti per il Piemonte;
- *integrazione delle basi dati* attraverso l'uso di gestori di database (MsAccess) e di funzioni di geoprocessing (Qgis).

2.1.1 Le basi dati

Si tratta di basi dati di fonte istituzionale. Nel seguito per comodità di esposizione, queste sono identificate con un'etichetta del tipo *DbNomeDimensione*. Per ognuna si indicano la fonte, la copertura temporale, l'unità di osservazione e la dimensione; si dà poi una sommaria descrizione del contenuto informativo e si precisa la frequenza di aggiornamento.

A. Db Incidenti

Fonte: Twist / CSI Piemonte

Anni disponibili: 2013 – 2016

Unità di osservazione: incidenti (44.719 records)⁸

Variabili: localizzazione (tipo e condizioni della strada, lat e long), circostanze e modalità, veicoli e persone coinvolte

Aggiornamento: annuale

B. Db SezCensimento

Fonte: ISTAT

Anni disponibili: 2011

Unità di osservazione: sezione censuaria (31.830 records)

Variabili: popolazione (genere, età), edificato (residenziale e non), unità locali, addetti

Aggiornamento: decennale

C. Db Strade

Fonte: 5T - CSI

Anni disponibili: 2015 - 2016

Unità di osservazione: arco stradale (484.494 records)

Variabili: (tutti i records) lunghezza, tipo di strada; (108.205 records) flusso veicoli leggeri e pesanti, capacità, velocità media

Aggiornamento: annuale

D. Db Comunali

Intesi come categoria di database che offrono informazioni aggregate a livello comunale, riguardanti, ad esempio, la mobilità, il consumo di suolo, i bilanci comunali, provenienti da varie fonti (Istat, Regione Piemonte, ARPA, Ministero dei Trasporti).

⁸ Di questi, solo 43.674 sono stati utilizzati in questa applicazione in quanto 1045 non avevano corretta georeferenziazione.

2.1.2 Integrazione delle basi dati

La principale criticità per integrare le basi dati è l'assenza di un identificativo comune che consenta di associare i diversi record individuali. In linguaggio da db relazionale ciò significa che mancano campi che siano possibili candidati (chiavi primarie e secondarie) utili per mettere in relazione (1-a-1 o 1-a-molti) i diversi data set.

In altre parole, è possibile incrociare le informazioni contenute nei diversi database solo scegliendo un livello di osservazione sufficientemente alto da consentire di predisporre una chiave di collegamento comune alle varie base dati, come già era stato fatto in un precedente lavoro (Landini, Occelli e Scalzotto, 2018).

A differenza di quel lavoro, nel quale la chiave di collegamento era costruita attraverso l'identificativo del nome della strada e di quello del comune, in questo studio il collegamento è realizzato a livello individuale: ogni singolo incidente è associato allo specifico arco stradale in cui si è verificato.

A tal fine è stato necessario:

- valorizzare i tutti i segmenti del DB strade;
- collegare i database attraverso i dati di georeferenziazione presenti in tutte le basi dati.

2.2 Valorizzazione dei segmenti del DB strade

Circa l'80% degli archi del DB strade sono privi dei dati di flusso (e delle informazioni sui volumi di traffico, la capacità e la velocità media del tratto stradale).

Per stimare questi dati e assegnarli ai rispettivi archi sono stati utilizzati software commerciali di gestione db (MsAccess e MsExcel). Si è inoltre adottata una logica di assegnazione che ha cercato valorizzare al massimo le informazioni contenuti nel DB. Nello specifico, si è tenuto conto di:

- tutti i dati di flusso disponibili associati ai 108mila segmenti;
- le informazioni riguardanti la localizzazione e tipologia dei segmenti;
- l'appartenenza dei segmenti a una stessa strada.

Di seguito, in sintesi, i principali passaggi

Fase 1. Informazioni di partenza

Il DbStrade contiene 484.494 segmenti di cui:

- 424.931 assegnati a 49.904 strade identificate da una denominazione e a un comune
- 59.563 non assegnati a strade né a un comune

108.205 segmenti sono valorizzati con i dati di flusso

8.986 strade hanno almeno 1 segmento valorizzato

Fase 2. Definizione della tipologia di segmenti

Per qualificare i segmenti del Dbstrade sono state prese in considerazione due caratteristiche:

- i segmenti sono valorizzati (campo flusso =1/0)
- i segmenti sono associati a una strada e a un comune (campo Istat =1/0)

In base a questo semplice schema si hanno 4 tipi di segmenti:

T1 -> Istat =1 ; Flusso =1 -> 103.133

T2 → Istat= 1 ; Flusso =0 -> 321.698

T3 → Istat = 0 ; Flusso =1 -> 5.072

T4 → Istat= 0 ; Flusso = 0 -> 54.591

Fase 3. Definizione delle regole di assegnazione, in base alle informazioni disponibili

L1 – A tutti segmenti di strada con almeno un arco valorizzato è attribuito lo stesso dato di flusso dell'arco

L2 – Assegnazione dei dati di flusso medi pesati sugli attributi della strada (tipologia e tronco) per comune di appartenenza

L3 – Assegnazione dei dati di flussi medi pesati sugli attributi della strada (come L2) e su quelli del comune (uso suolo, densità, altimetria) nel caso non esistano altre strade valorizzate a livello comunale per gli specifici attributi

Fase 4. Assegnazione dei dati di flusso ai segmenti non valorizzati

Prima di procedere all'assegnazione è stato necessario attribuire attraverso intersezione GIS (si veda il paragrafo seguente), i codici ISTAT ai segmenti di tipo T3 e T4.

V0. Valore iniziale: 108.205 segmenti valorizzati

V1. Adottando L1 si valorizzano 78.133 segmenti (per totali 186.138)

V2. Adottando L2 si valorizzano 167.420 (per totali 353.558)

V3. Adottando L3 si valorizzano i restanti 130.936 (per totali 484.494)

2.3 Collegamento dei database e funzioni di Geoprocessing

Oltre alla visualizzazione, la georeferenziazione delle basi informative consente, attraverso le coordinate geografiche, di unire informazioni contenuti in database diversi.

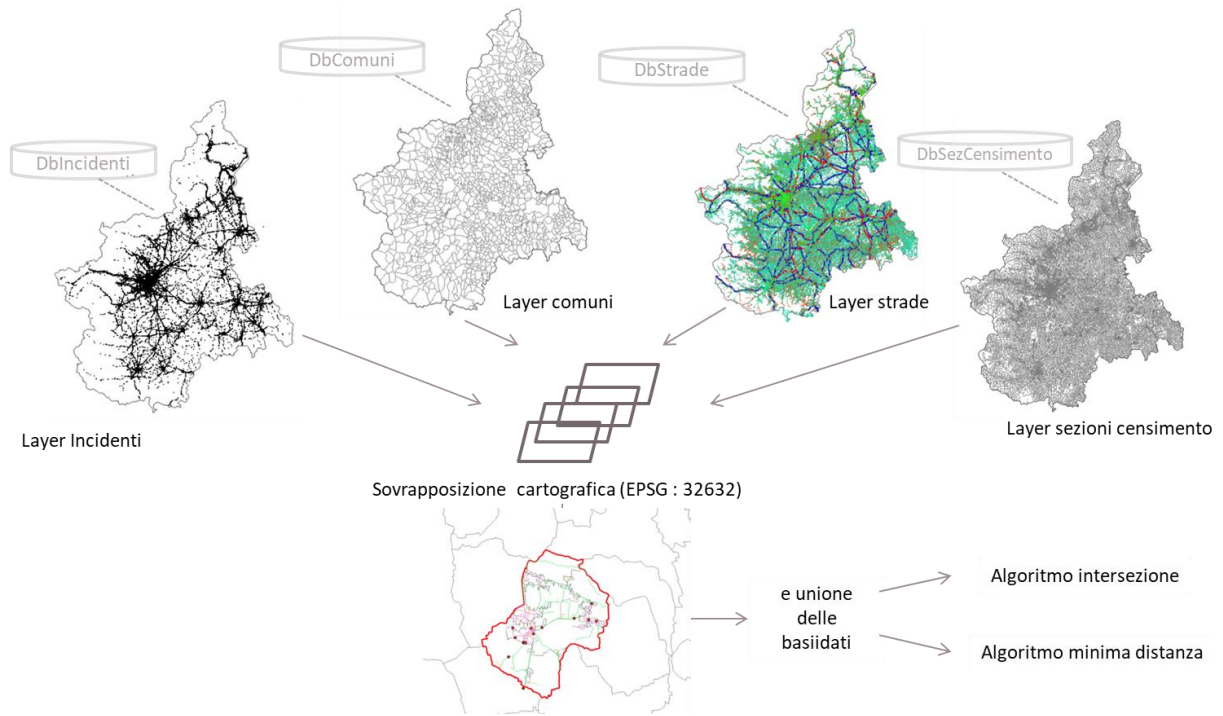
Perché questo sia possibile, è bene ricordare che i software GIS visualizzano file vettoriali o di coordinate che sono chiamati layers. A questi layers sono collegati (uno o più) database contenenti le informazioni che si vogliono rappresentare su mappa. La condizione essenziale perché si possa incrociare due database attraverso il riferimento geografico è quindi che a ogni database sia associato un layer georeferenziato. Nello specifico di questo studio:

- il DbIncidenti contiene le coordinate che consentono di creare un layer di punti corrispondenti alla localizzazione dell'incidente;
- i DbSezCensimento e il DbStrade sono associati rispettivamente a un layer di poligoni e a un layer di linee rilasciati congiuntamente ai database da ISTAT e da CSI Piemonte

Nel seguito, per semplicità, utilizzeremo il termine layer per fare riferimento sia ai file che contengono l'informazione cartografica, sia ai database ad essi collegati.

Le mappe in Fig.1 presentano alcune visualizzazioni (i layers) dei dati contenuti nei DB utilizzati.

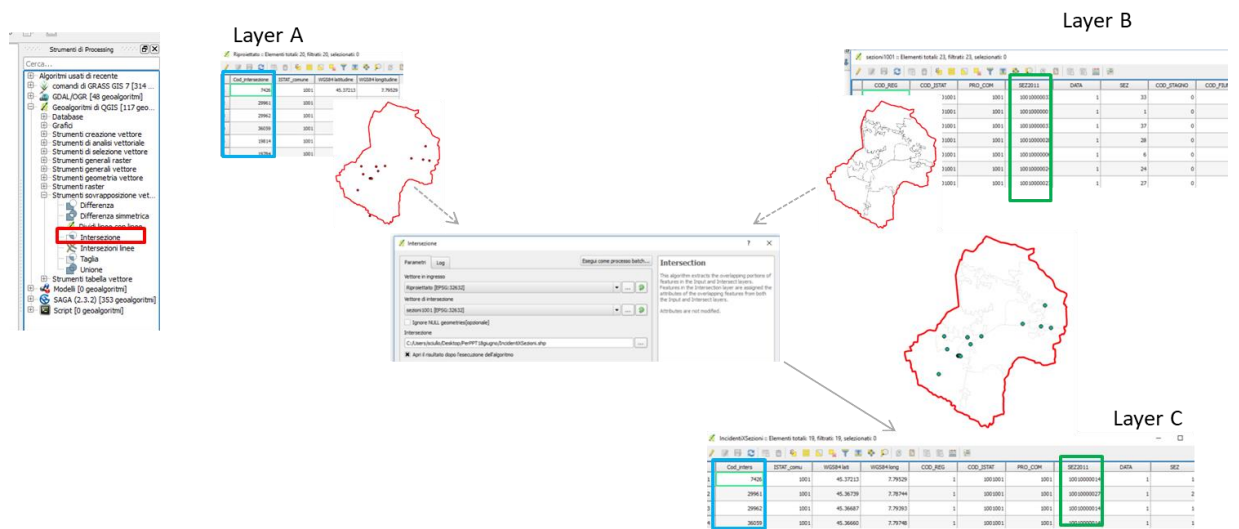
Fig.1 L'uso del GIS per integrare le basi dati



2.3.1 Unione dei DB incidenti con il DB sezioni censimento

E' stata realizzata attraverso la funzione di intersezione, una funzione che incrocia geometricamente, per sovrapposizione geografica, due layers: un layer A (vettore in ingresso) e un layer B (vettore di intersezione). Il risultato è un nuovo layer C con numerosità pari a A e l'assegnazione degli attributi corrispondenti di B, Fig. 2.

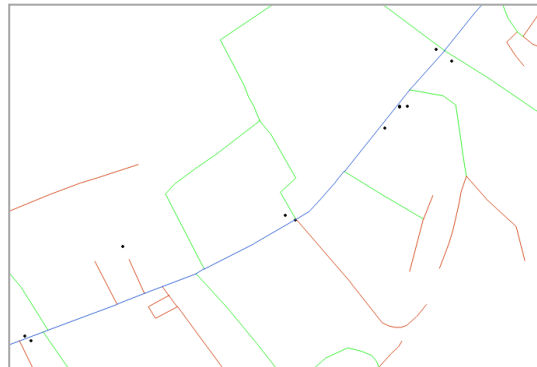
Fig. 2 Unire le basi dati (Layer C) attraverso l'intersezione tra layers (A e B)



2.3.2 Unione dei DB incidenti con il DB Strade

Ha richiesto una procedura ad hoc poiché la funzione di intersezione non è applicabile: i layer dei rispettivi DB non si sovrappongono, vedi Fig.3.

Fig.3 Esempio dei problemi di allineamento nella sovrapposizione tra il layer del DbIncidenti (punti) e quello del DbStrade (linee)



Nello specifico è stato necessario ricorrere a un algoritmo di distanza minima per assegnare il singolo evento incidentale all'arco.

L'algoritmo di minima distanza calcola la distanza tra i punti di un layer A (vettore di punti origine) e il punto più vicino di un layer B (vettore di linee o punti destinazione), Fig.4.

Il risultato è la creazione di un nuovo layer C associato a un database con numerosità pari a A e al quale sono aggiunti due campi: ID del punto/linea più vicino nel layer B e distanza tra i due punti.

Fig. 4 Unire le basi dati associate ai layer A e B attraverso l'assegnazione di un punto (layer A) al segmento più vicino (layer B) con l'algoritmo di minima distanza

Layer A (DbIncidenti)

ID	Anno	Data	Mese	WGS84 lat	WGS84 long
1	13	2022	4	45.08554	6.88863
2	14	2022	8	45.08123	6.88143
3	14	2022	7	45.07322	6.88877
4	13	2022	4	45.07903	6.76330
5	13	2022	9	45.08420	6.7
6	16	2022	8	45.08654	6.7

Layer B (DbStrade)

Linea	COD_B_2/1	COD_B_2/2	COD_B_2/3	COD_B_2/4	COD_B_2/5	COD_B_2/6	COD_B_2/7	ES_APPA_NOM
1	05	0,01	01	01	01	01	01	SP-41.8 Aggr
2	05	0,01	01	01	01	01	01	
3	05	0,01	01	01	01	01	01	
4	05	0,01	01	01	01	01	01	
5	05	0,01	01	01	01	01	01	
6	05	0,01	01	01	01	01	01	
7	05	0,01	01	01	01	01	01	
8	05	0,01	01	01	01	01	01	
9	05	0,01	01	01	01	01	01	
10	05	0,01	01	01	01	01	01	

Algoritmo di minima distanza

Layer C (DbIncidenti X DbStrade)

Cod_intersezione	WGS84 latitudine	WGS84 longitudine	HubName	HubDist	
5	16656	45.39915	7.7655	053C825E-89E0-2909-E054-000B8A0F36E5	89.54130
6	19796	45.39929	7.7655	053C825E-8A21-2909-E054-000B8A0F36E5	6.89287
7	20802	45.36218	7.7669	053C825F-A7A2-2909-E054-000B8A0F36E5	20.89720

L'attribuzione dell'ID del punto/linea (layer B) più vicino al punto che si vuole assegnare (layer A) comporta che nel database del layer C vengono replicati tutti i record e gli attributi del layer A e ad ogni record vengono associati i due nuovi valori, ID del punto/linea più vicino del layer B e la distanza. L'ID è la chiave di accesso a tutti gli attributi del database collegato al layer B e quindi consente l'integrazione tra le basi dati.

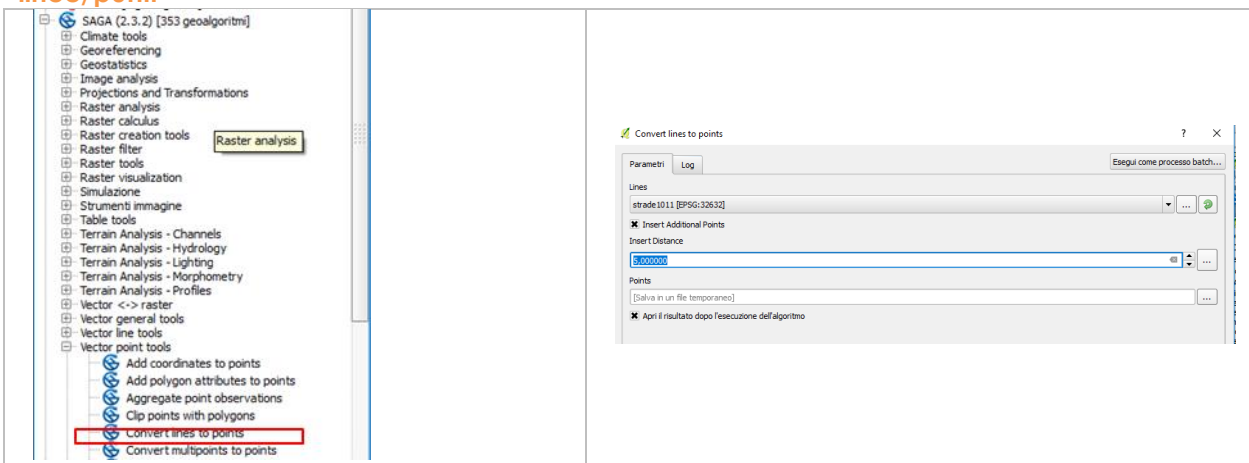
L'algoritmo di minima distanza ha però un limite di funzionamento per come è implementato in Qgis 2.18.18: Il problema consiste nel fatto che non sempre il punto è attribuito al segmento più vicino. Per quanto non sia chiara la causa del problema, dopo alcune prove, sembra ragionevole ipotizzare che il problema consista nel fatto che l'algoritmo considera la distanza tra il punto e gli estremi dei segmenti e non la perpendicolare che unisce il punto ai segmenti. Ne deriva che, seppur non sistematicamente, il punto è attribuito a un segmento scelto a caso tra quelli che terminano con l'estremo più vicino al punto: questo non garantisce però che il segmento scelto sia effettivamente quello più vicino.

il risultato migliora molto se la distanza è calcolata tra due layer di punti invece che tra un layer di punti e uno di linee, come è nel nostro caso.

La soluzione, molto dispendiosa in termini computazionali, è quella di trasformare il vettore di linee in vettore di punti in modo da non incorrere nella distorsione dei risultati dell'algoritmo e aumentare sensibilmente la probabilità che ogni punto del layer A sia assegnato effettivamente al punto (segmento) più vicino del layer B, Fig. 5.

In questa applicazione è sembrato ragionevole applicare una distanza di 5mt tra i punti che hanno sostituito il layer di linee. Questa trasformazione converte i 75000 km della rete stradale piemontese in un layer, e un database collegato, costituito da oltre 15 milioni di record.

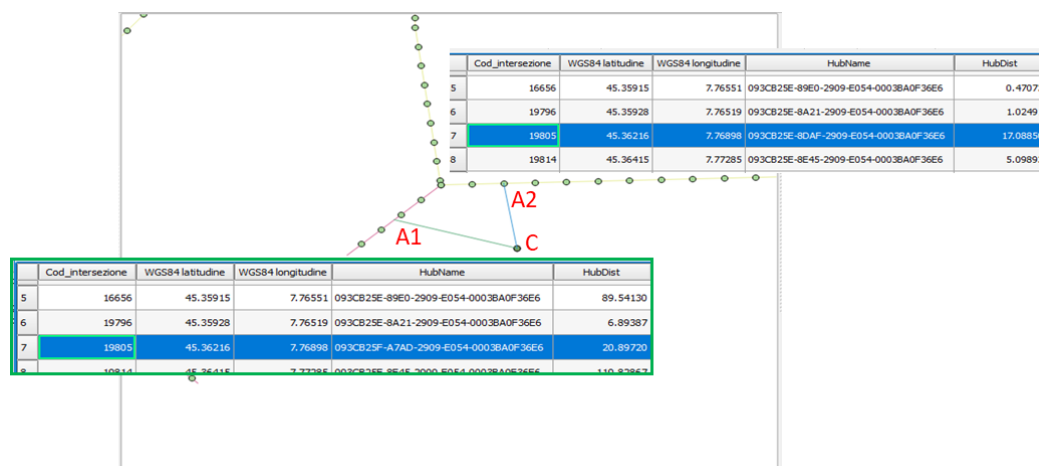
Fig.5 Trasformazione di un layer di linee in un layer di punti attraverso la funzione GIS converti linee/punti



Come mostrato in Fig. 6, l'attribuzione al vettore di linee (A1) associa il punto C a un segmento distante oltre 20 mt mentre l'attribuzione dello stesso punto C al vettore di punti (A2) lo as-

socia a un diverso punto (segmento) distante poco più di 17 mt⁹. Da questo punto di vista, la procedura utilizzata migliora la soluzione dell' algoritmo originale.

Fig.6 Assegnare un incidente alla rete stradale: la necessità di trasformare le linee in punti



2.4 La Cluster Analysis per identificare i profili territoriali

La cluster Analysis è una metodologia ampiamente utilizzata in ricerca sociale ed economica per costruire tipi di gruppi a partire da un raggruppamento delle unità elementari. Negli studi sull'incidentalità stradale, ad esempio, è stata utilizzata per individuare i punti critici di incidentalità su una rete infrastrutturale (i *blackspot*) (Anderson, 2009, Ng et al ,2002). In questo studio la utilizziamo per identificare una tipologia territoriale delle situazioni incidentali, dove queste ultime non derivano dalla sola considerazione delle caratteristiche degli incidenti.

2.4.1 Che cos'è la cluster analysis

Data una matrice di N unità di osservazione caratterizzate da M variabili, gli algoritmi di clustering sono metodi statistici che consentono di raggruppare le N osservazioni in base alla loro somiglianza per le M variabili a esse associate.

In pratica, applicando un algoritmo di clustering a una popolazione di individui otterremo un certo numero di gruppi, diversi fra loro, ma formati da individui simili, relativamente alle variabili descrittive originariamente selezionate per rappresentare gli individui.

Per comodità, chiameremo le m variabili utilizzate per alimentare l' algoritmo *variabili di clustering*, le altre le M-n variabili, utili qualificare meglio la tipologia individuata, *variabili descrittive* dei cluster.

⁹ L'efficienza è misurata in termini puramente geometrici nell'ipotesi che valga come approssimazione migliore della registrazione effettuata dalle forze dell'ordine.

BOX 1. L'algoritmo di clustering

Esistono due principali famiglie di tecniche di clustering.

Cluster gerarchico, in cui i gruppi sono individuati attraverso un processo iterativo che ad ogni step associa in successione gli elementi simili. Dagli individui singoli si formano gruppi sempre più ampi a ogni iterazione fino al raggiungimento di una certa soglia di aggregazione che soddisfa le esigenze del ricercatore.

Cluster partizionale, in cui si decide a priori il numero di cluster da individuare e attraverso un processo iterativo si assegnano i singoli individui ai cluster mediante la minimizzazione della distanza tra il singolo individuo e il punto rappresentativo del cluster.

L'algoritmo usato in questo studio è il k-means (MacQueen, 1967), il più diffuso tra gli algoritmi partizionali. In estrema sintesi, la procedura consiste nei passi seguenti:

1. si definiscono le variabili di clustering e il numero di clusters
2. attraverso un'assegnazione casuale gli individui sono assegnati ai cluster, ottenendo un raggruppamento iniziale, utilizzato per inizializzare l'algoritmo
3. si calcolano i punti rappresentativi dei cluster nello spazio vettoriale (centroidi)
4. I singoli individui vengono assegnati al cluster il cui centroide è più vicino in base alla minimizzazione della distanza
5. si ripetono i passi 3-4 fino a quando nessun individuo cambia di assegnazione a seguito del ricalcolo dei centroidi.

2.4.2 L'applicazione della cluster analysis all'analisi dell'incidentalità

In questo studio, si è deciso di utilizzare la cluster analysis per identificare i profili territoriali dell'incidentalità. Nello specifico, l'unità di osservazione è il comune, al quale è associato un certo insieme di variabili selezionate sulla base dei dati contenuti nei vari DB. Queste variabili riguardano i profili seguenti, Fig. 6:

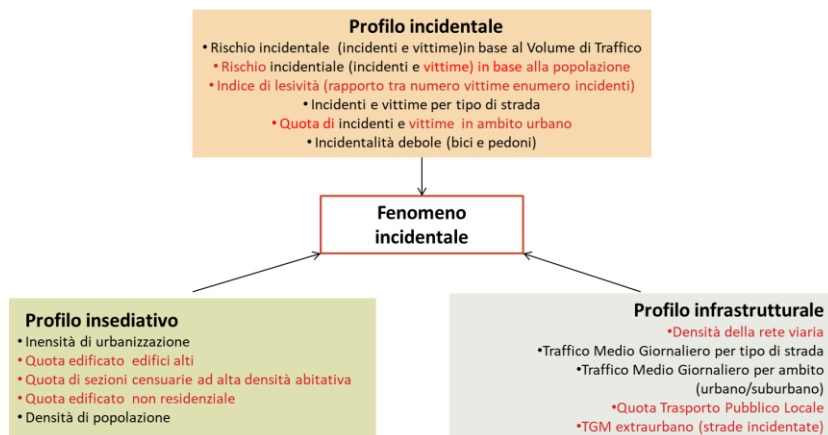
- incidentale, che coglie l'incidenza e le caratteristiche del fenomeno incidentale
- insediativo, che cerca render conto della densità del tessuto urbanizzato
- infrastrutturale, che considera il livello e l'intensità del traffico sulla rete viaria e i modi di spostamento.

Le *variabili di clustering* con il dettaglio della loro specificazione operativa sono:

- Rischio incidentale (numero di vittime / popolazione residente)
- Indice di lesività (numero vittime / numero incidenti)
- Quota di vittime in ambito urbano (numero vittime in ambito urbano / totale vittime)
- Densità della rete viaria (km strade / superficie comunale)
- Quota di Trasporto Pubblico (numero di spostamenti su TPL / totale spostamenti)
- Traffico Giornaliero Medio (per le sole strade con incidenti) (media pesata per la lunghezza dei segmenti del numero di veicoli leggeri e pesanti)
- Quota edifici alti (superiori ai 3 piani) (numero edifici residenziali superiori a i 3 piani / totale edifici residenziali)
- Quota di sezioni censuarie ad alta densità abitativa (sezioni con densità abitativa superiore ai 1500 ab per km² – soglia Espon/ totale delle sezioni)
- Quota edificato non residenziale (numero edifici non residenziali / totale edifici)

Tutte le variabili (ad eccezione di quella relativa al trasporto pubblico locale) sono ottenute dall'aggregazione su base comunale dei microdati contenuti nella base informativa integrata.

Fig. 6. Variabili di clustering (in rosso) e variabili descrittive dei cluster per profilo di analisi



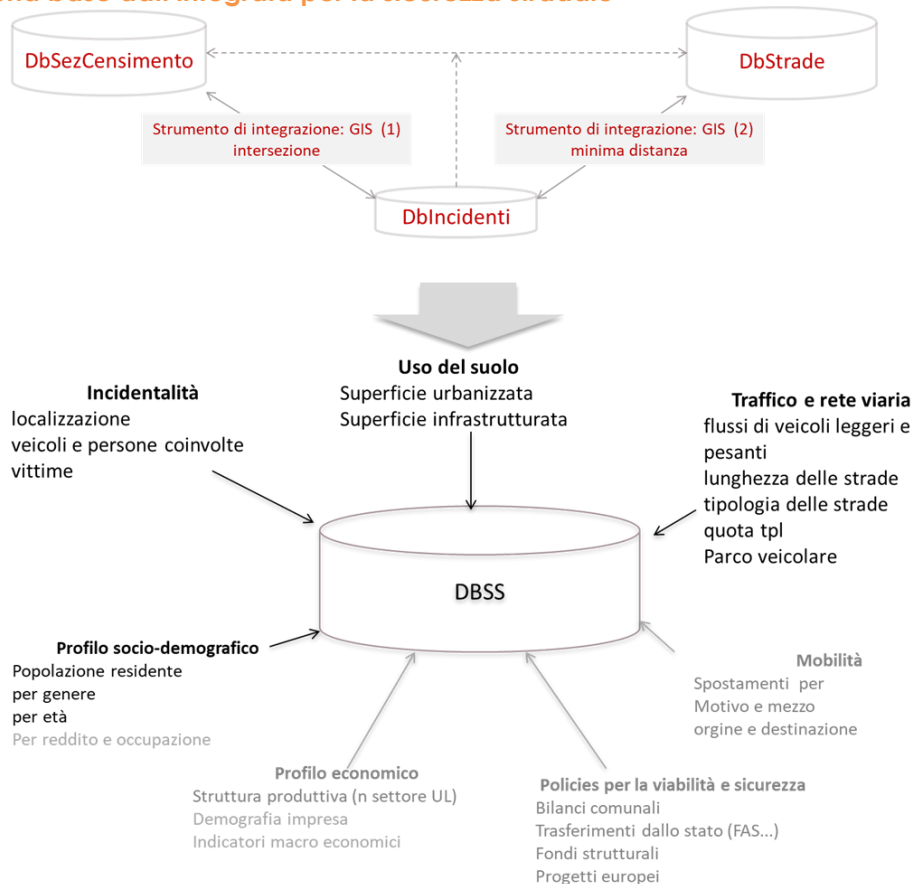
3. COSA HA PRODOTTO LO STUDIO

3.1 Una base dati integrata per la sicurezza stradale (DBSS)

A conclusione del percorso metodologico descritto nel capitolo precedente è stato predisposto un Data-Base integrato, per la Sicurezza Stradale DBSS¹⁰ (un DB relazionale in linguaggio tecnico), Fig. 7 che collega:

- DBIncidenti, che contiene informazioni relative a localizzazione dell'incidente, veicoli e persone coinvolte
- il DbStrade, che contiene informazioni relative al traffico veicolare ovvero flussi di veicoli leggeri e pesanti, lunghezza e tipologia delle strade, capacità e velocità media
- il DBSezCensimento che contiene informazioni relative a popolazione per gener ed età, intensità e tipologia dell'edificato, unità locali e addetti

Fig. 7 DBSS una base dati integrata per la sicurezza stradale



L'aggregazione a livello comunale dei dati individuali contenuti in questi data-set permette, attraverso il codice Istat del comune, di collegare questi dati con altri *DbComunali* contenenti

¹⁰ Tale database rappresenta una realizzazione di quello previsto a suo tempo nel progetto del CMRSS piemontese, sottoposto al Ministero dei Trasporti, nel 2008.

ti informazioni sulla mobilità e i trasporti. Nel caso specifico sono stati considerati i dati riguardanti la quota di Trasporto Pubblico Locale (Fonte: Istat).

Lo schema di Fig. 7 evidenzia come questo DBSS sia da considerarsi come il **nocciolo di un sistema informativo scalabile**.

In figura sono riportate, in nero, le variabili a oggi già integrate e, in grigio, quelle che potrebbe essere opportuno integrare per affinare l'analisi territoriale, quali, ad esempio, quelle riguardanti:

- le spese per viabilità e trasporti (contenute nei bilanci comunali) ;
- le spese sanitarie causate dall'incidentalità;
- il livello di servizio pubblico locale;
- le Unità locali, gli occupati, la popolazione;
- i progetti di intervento infrastrutturale (e ai relativi costi);
- le multe elevate dalle Forze dell'Ordine;
- i consumi delle famiglie in termini di mobilità e trasporto.

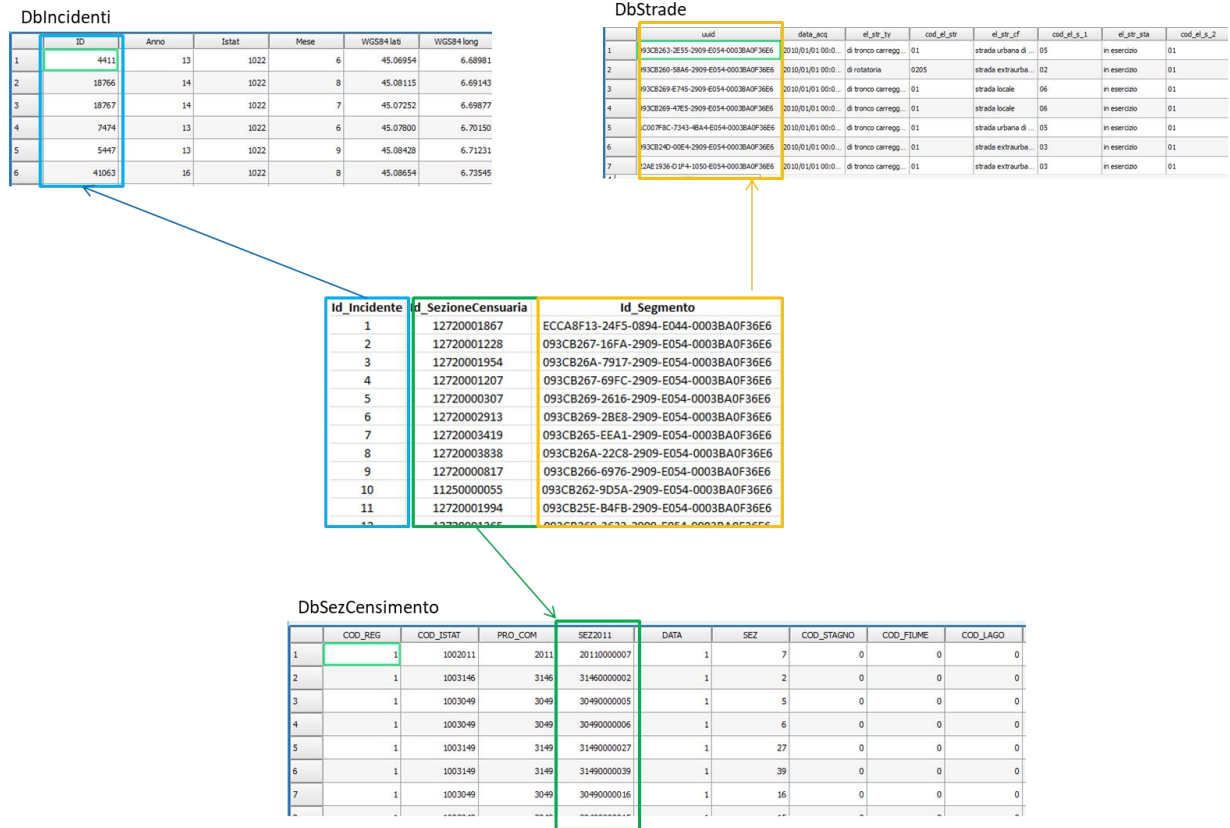
A beneficio degli utilizzatori del DBSS merita sottolineare che le relazioni tra i record individuali dei dataset sono così definite, Fig.8:

- relazione 1-a-molti tra DbStrade e DbIncidenti per cui a uno stesso segmento di strada possono corrispondere più incidenti
- relazione 1-a-molti tra DbSezCensimento e DbIncidenti per cui a una sezione censuaria possono corrispondere più incidenti
- relazione molti-a-molti tra DbStrade e DbSezCensimento per cui a uno stesso segmento di strada possono corrispondere diverse sezioni di censimento (quelle attraversate dal segmento) e viceversa.

La tabella di vettori di identificativi riportata in Fig. 8, è il cuore del DBSS. Essa mostra queste relazioni e illustra come è stata applicata la logica relazionale. Con questo termine si intende che:

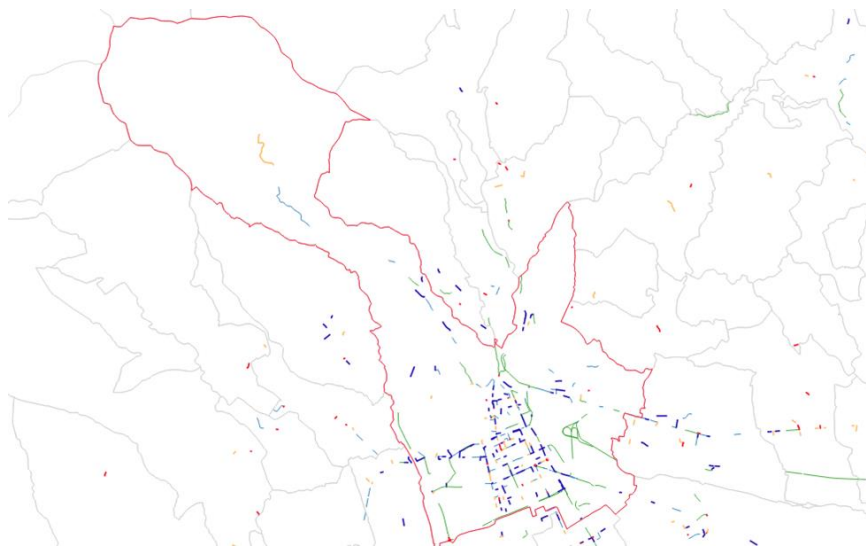
- all'interno della tabella in figura, riga per riga (*tuple* in gergo informatico), ciascun incidente è messo in relazione con una precisa sezione di censimento e uno specifico segmento;
- ogni singolo Id consente di accedere a tutte le informazioni contenute nei DB di cui l'oggetto rappresentato da quell'Id costituisce l'unità di osservazione (i.e. attraverso l'Id_segmento è possibile accedere a tutte le informazioni sul traffico veicolare contenute nel DbStrade e associate a quel segmento che saranno quindi associabili anche all'incidente con cui quel segmento è stato messo in relazione).

Fig. 8 Il cuore del DBSS – i vettori di identificativi



Un esempio di utilizzo delle informazioni contenute nel DBSS è mostrato in Fig. 9 dove i singoli archi della rete viaria del comune di Biella sono stati colorati per valori del rischio incidentale, qui misurato come rapporto tra incidenti e flusso di traffico per km.

Fig. 9 Un esempio di utilizzo del DBSS: rischio sugli archi stradali nel comune di Biella



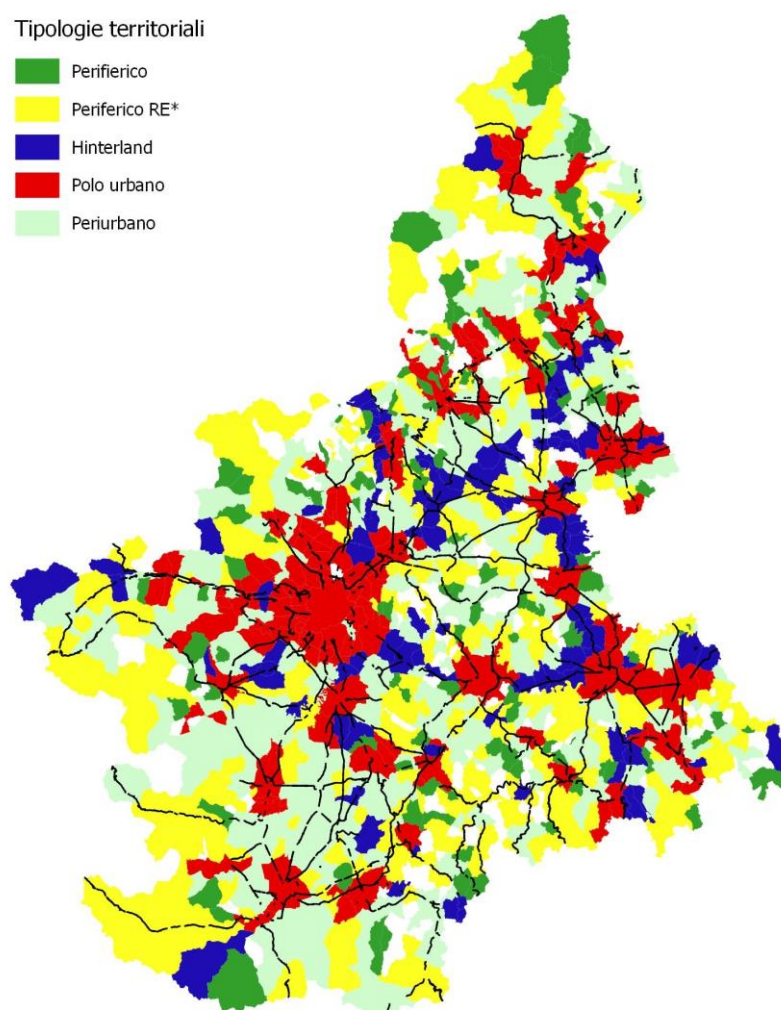
Nota : I segmenti sono colorati in base al rischio incidentale

3.2 I profili territoriali dell'incidentalità

3.2.1 La tipologia dei comuni

L'applicazione dell'algoritmo di cluster ha identificato 5 tipi di profili territoriali dell'incidentalità. La loro distribuzione nel territorio regionale è mostrata in Fig. 10.

Fig. 10 Distribuzione della tipologia di profili territoriali dell'incidentalità in Piemonte



Fonte: elaborazione Ires su dati, Istat, Regione Piemonte e TOC

Nota: in evidenza le strade identificate nel DbStrade come extraurbane principali (statali e provinciali di I e II livello)
(*) Periferico a Rischio Elevato, Tab. 1a e Tab. 2.

L'esame della distribuzione geografica dei gruppi suggerisce che esiste un legame, peraltro non inatteso, tra profili incidentali e morfologia insediativa. I nomi dei cluster riflettono questo legame. I comuni lasciati in bianco in Fig. 10 (il 17% dei comuni piemontesi) non hanno avuto incidenti sulle strade comunali, provinciali o statali, del proprio territorio, nel periodo 2013-16.

I gruppi più numerosi in termini di comuni, sono i gruppi denominati Periferico a Rischio Elevato (RE) e Periurbano, Tab. 1a.

Il gruppo che comprende i Poli Urbanizzati rappresenta il profilo maggiormente problematico dal punto di vista della numerosità dei sinistri e degli infortunati: è costituito dal 15% dei comuni nei quali avvengono oltre il 75% dei sinistri con infortunati.

Dal punto di vista degli incidenti fatali, la criticità maggiore si riscontra per i comuni Periferici a Rischio Elevato. Pur concentrando poco meno del 5% del totale degli incidenti, la quota di morti è quasi tre volte più elevata (14%). A livello provinciale, un numero relativamente elevato dei comuni di Torino, Novara, Biella e VCO appartiene ai gruppi Poli Urbanizzati e Periurbani: insieme caratterizzano oltre il 50% dei comuni di queste province, Tab.1b.

Il gruppo Periferico a Rischio Elevato è presente soprattutto nelle province di Asti, Cuneo, Alessandria. Il gruppo Hinterland è maggiormente presente a Vercelli e Novara ed è assente nella provincia di Biella.

Tab.1a Distribuzione dei comuni, incidenti, infortunati e morti nei gruppi

	Comuni		Incidenti		Infortunati (feriti + morti)		Morti	
	N	%	N	%	N	%	N	%
Periferico	112	11,1	583	1,4	851	1,4	16	2,1
Periferico (RE)	320	31,8	1792	4,4	2900	4,9	109	14,1
Hinterland	112	11,1	1342	3,3	1973	3,4	60	7,8
Polo Urbanizzato	158	15,7	31780	78,5	45481	77,3	397	51,3
Periurbano	305	30,3	4975	12,3	7613	12,9	192	24,8
Totale	1007	100,0	40472	100,0	58818	100,0	774	100,0

Fonte: elaborazione Ires su dati, Istat, Regione Piemonte e TOC

Nota: I comuni interessati dal fenomeno incidentale sono 1014. Per 7 comuni infatti gli incidenti avvengono solo sulle autostrade. Questi sono stati esclusi dall'analisi di cluster e non sono considerati in tabella. Considerando i 1014 comuni gli incidenti nel periodo 2013-16 sono 44.719.

Tab.1b. Distribuzione percentuale dei comuni in ciascun gruppo per provincia

	Torino	Vercelli	Novara	Cuneo	Asti	Alessandria	Biella	VCO	Totale
Periferico	10,0	9,9	12,7	8,9	12,8	13,1	14,1	12,7	11,1
Periferico (RE)	23,8	33,8	11,4	42,9	46,5	36,3	26,6	28,6	31,8
Hinterland	12,1	26,8	19,0	5,4	10,5	13,1	0,0	4,8	11,1
Poli Urbanizzati	26,3	5,6	20,3	8,9	3,5	10,0	20,3	22,2	15,7
Periurbani	27,8	23,9	36,7	34,0	26,7	27,5	39,1	31,7	30,3
Totale	100,0								

Fonte: elaborazione Ires su dati, Istat, Regione Piemonte e TOC

La Tab. 2 riporta i valori delle variabili di clustering nei diversi gruppi. La loro lettura consente, in pratica, di evidenziare le principali differenze tra i gruppi e di qualificarne il profilo.

Tab. 2 Distribuzione delle variabili di clustering nei gruppi (*)

	Periferico	Periferico (RE)	Hinterland	Poli Urbanizzati	Periurbani	Totale
Infortunati rispetto alla popolazione (x 1000 ab)	1,52	2,90	2,16	2,55	2,63	2,53
Infortunati per 100 incidenti	138	157	154	143	151	150
Infrastrutturazione (km strade per kmq)	3,2	2,9	3,3	5,0	3,5	3,5
Rapporto tra edifici alti (3 e più piani) e edifici bassi (<= 2 piani)	0,66	0,51	0,38	0,76	0,54	0,56
Quota % edificato non residenziale	8,1	7,1	8,5	11,3	8,7	8,5
Velocità media sulla rete (escluse le autostrade)	45,2	49,4	51,4	48,3	50,5	49,3
Quota TPL mobilità sistemica 2011	15,0	15,1	13,6	12,5	13,7	14,1
Quota % infortunati in ambito urbano	93,1	6,7	37,1	73,5	45,6	42,0
% sez.cens ad elevata densità popolazione	16,6	16,0	24,2	55,1	22,3	25,0
Quota % volume traffico in extra-urbano	9,3	96,6	70,0	44,2	71,1	68,0

(*) In giallo e in grigio il valore rispettivamente più e meno elevato.

Fonte: elaborazione Ires su dati, Istat, Regione Piemonte e TOC

- **Periferico (11% dei comuni)**: situato in aree periferiche e/o interstiziali del territorio regionale, ha valori bassi di infrastrutturazione e di criticità incidentale, in termini sia di rischio, numero di infortunati (feriti più morti) rispetto alla popolazione, sia di livello di infortunio (numero feriti e di morti) per 100 incidenti. Molto rilevante la quota di infortunati in ambito urbano. La quota TPL è relativamente più elevata.
- **Periferico Rischio Elevato (RE) (32% dei comuni)**: situato in aree poco densamente abitate, ha una bassa infrastrutturazione ma un'alta criticità incidentale in termini sia di rischio sia di infortunio. Traffico e incidenti si concentrano soprattutto in ambito extra-urbano. La quota TPL è relativamente più elevata.
- **Hinterland (11% dei comuni)**: si colloca soprattutto lungo le direttrici più importanti. Ha una modesta infrastrutturazione, con strade a scorrimento veloce e velocità media più alta che negli altri gruppi. Infortunati e volume di traffico si concentrano soprattutto in area extra-urbana.
- **Poli Urbanizzati (16% dei comuni)**: Include i centri regionali più popolosi. Ha valori elevati di infrastrutturazione e di densità dell'edificato. Concentra il 78% degli incidenti e il 51% dei morti. Il rischio di incidentalità e di infortunio hanno valori contenuti. Quasi 3 infortuni su quattro sono causati da sinistri in ambito urbano.
- **Periurbani (30% dei comuni)**: situato in prossimità dei comuni maggiori, ha un profilo d'infrastrutturazione simile a quello medio regionale, associato però ad un livello di rischio (numero di infortunati rispetto alla popolazione) superiore. La quota di infortunati è quasi egualmente distribuita tra ambito urbano e extra-urbano.

Le descrizioni dei diversi profili può poi essere ulteriormente affinata alla luce di alcune delle variabili elencate in Fig.6, e riportate nelle Tabb. 3 e 4.

Profilo incidentale, Tab.3

- *Periferico*: associa il più basso livello di rischio rispetto alla popolazione con il più alto livello di rischio rispetto alle strade incidentate. Incidenti e infortunati si concentrano soprattutto sulle strade in ambito urbano con il più elevato coinvolgimento degli utenti vulnerabili.
- *Periferico (RE)*: si caratterizza per i più alti valori del rischio di incidenti e infortuni rispetto alla popolazione, soprattutto sulle strade in ambito extra-urbano. Ha il più basso livello di coinvolgimento degli utenti vulnerabili.
- *Hinterland*: presenta valori più bassi del rischio incidentale in relazione al volume di traffico sulle strade comunali e statali. Ha un elevato rischio di infortuni sulle strade provinciali.
- *Poli Urbanizzati*: hanno i valori più elevati di incidentalità e di infortunio rispetto alla lunghezza totale delle strade e al volume di traffico totale. Presentano, però i più bassi valori rispetto al traffico sulle strade incidentate sia in ambito urbano che extraurbano.
- *Periurbano*: hanno un profilo relativamente simile a quello regionale. Si caratterizzano per il più alto valore di rischio di incidenti e di infortunio sulle strade comunali.

Tab. 3a Profilo incidentale dei gruppi: aspetti generali

	Periferico	Periferico (RE)	Hinterland	Poli Urbanizzati	Periurbani	Totale
Incidenti per Km di strade totali	0,12	0,14	0,19	0,75	0,25	0,27
Incidenti per Km di strade incidentate	8,29	3,76	4,57	5,42	4,40	4,77
Infortunati per Km di strade totali	0,18	0,23	0,29	1,07	0,38	0,40
Infortunati per Km di strade incidentate	11,20	5,99	7,07	7,70	6,43	7,04
Quota % di incidenti che coinvolgono pedoni	10,7	1,5	3,6	9,7	4,6	5,0
Quota % di incidenti che coinvolgono pedoni o bici	19,1	6,6	9,8	16,9	10,7	11,1

Fonte: elaborazione Ires su dati, Istat, Regione Piemonte e TOC

Tab. 3b Profilo incidentale dei gruppi: il rischio incidentale totale e per tipo di strada

	Periferico	Periferico (RE)	Hinterland	Poli Urbanizzati	Periurbani	Totale
Incidenti rispetto alla popolazione (x 1000 ab)	1,03	1,82	1,43	1,79	1,74	1,65
Incidenti rispetto al volume di traffico (tot strade)*	0,11	0,11	0,09	0,20	0,13	0,13
Incidenti rispetto al volume di traffico (tot strade incidentate)	11,11	4,50	3,14	1,70	2,35	3,76
Comunali: Incidenti rispetto al volume di traffico	4,15	3,74	2,74	3,85	8,58	5,16
Provinciali: Incidenti rispetto al volume di traffico	9,42	4,11	6,15	1,72	2,79	4,15
Statali: Incidenti rispetto al volume di traffico	0,39	0,10	0,07	0,20	0,18	0,17
Urbano: Incidenti rispetto al volume di traffico	11,32	4,43	3,96	2,55	5,33	5,12
Extra-urbano: Incidenti rispetto al volume di traffico	3,25	4,42	2,16	1,23	2,50	2,96

(*) I valori rispetto al volume di traffico sono dati dal rapporto: numerosità incidenti/volume x 1000000

Fonte: elaborazione Ires su dati, Istat, Regione Piemonte e TOC

Tab. 3c Profilo incidentale dei gruppi: il rischio di infortunio totale e per tipo di strada

	Periferico	Periferico (RE)	Hinterland	Poli Urbanizzati	Periurbani	Totale
Infortunati rispetto alla popolazione (x 1000 ab)	1,50	2,88	2,15	2,56	2,62	2,52
Infortunati rispetto al volume di traffico totale	0,159	0,175	0,134	0,283	0,199	0,188
Infortunati rispetto al volume di traffico incidentato	14,40	7,44	4,68	2,34	3,28	5,51
Comunali: Infortunati rispetto al volume di traffico	5,19	4,76	3,70	5,09	10,56	6,50
Provinciali: Infortunati rispetto al volume di traffico	12,26	6,89	19,11	2,37	3,92	7,24
Statali: Infortunati rispetto al volume di traffico	1,25	0,16	0,12	0,31	0,25	0,34
Urbano: Infortunati rispetto al volume di traffico	14,91	5,22	4,73	3,55	8,26	6,85
Extra-urbano: Infortunati rispetto al volume di traffico	3,92	7,37	4,76	1,67	3,30	4,54

(*) I valori rispetto al volume di traffico sono dati dal rapporto: numerosità incidenti/volume x 1000000
 Fonte: elaborazione Ires su dati, Istat, Regione Piemonte e TOC

Profilo infrastrutturale, Tab.4

- Il TGM è nettamente più elevato per il gruppo Poli Urbanizzati; i gruppi Periferici a Rischio Elevato e non, mostrano i valori più contenuti, Tab. 4a;
- per il gruppo dei Periferici (RE) si osserva il valore più elevato del volume di traffico sulle strade provinciali e nell'ambito extra-urbano; nei Poli Urbanizzati la quota maggiore di traffico insiste sulle strade comunali. I Periferici concentrano in ambito urbano oltre il 90% del loro volume di traffico, Tab. 4b;
- per quanto riguarda la velocità di scorrimento, i valori più alti si osservano nel gruppo Hinterland; i comuni che appartengono al gruppo Poli Urbanizzati sono quelli più connessi dalle principali reti della viabilità regionale. Tab.4c.

Tab. 4a Profilo infrastrutturale dei gruppi: Traffico Giornaliero Medio totale e per tipo di strada

	Periferico	Periferico (RE)	Hinterland	Poli Urbanizzati	Periurbani	Totale
Strade Comunali	658	401	1616	4005	1529	1471
Strade Provinciali	1739	2519	4331	6414	3639	3583
Strade Statali	191	385	517	1812	825	737
Strade in ambito urbano						
Strade in ambito urbano	2011	947	3100	5487	3010	2641
Strade in ambito extra-urbano						
Strade in ambito extra-urbano	958	2794	4266	6496	4283	3785
Totale Strade incidentate						
Totale Strade incidentate	1931	2719	4124	5924	3814	5312
Totale Strade						
Totale Strade	1303	1474	2574	3757	2206	2501

Tab. 4b Profilo infrastrutturale dei gruppi: ripartizione del volume di traffico per tipo di strada

	Periferico	Periferico (RE)	Hinterland	Poli Urbanizzati	Periurbani	Totale
Strade Comunali	15,5	5,9	9,5	23,5	8,6	10,9
Strade Provinciali	79,2	85,5	84,5	65,1	78,9	79,5
Strade Statali	5,2	8,2	5,3	11,0	12,1	9,1
Totale	100*					
Strade in ambito urbano	90,7	2,9	29,2	55,5	28,5	31,6
Strade in ambito extra-urbano	9,3	96,6	70,0	44,2	71,1	68,0
Totale	100*					

Nota: Per alcuni cluster il totale non somma a 100 poiché il tipo 'Altra strada' non attribuibile a nessuno dei due ambiti
 Fonte: elaborazione Ires su dati, Istat, Regione Piemonte e TOC

Tab. 4c Profilo infrastrutturale dei gruppi: velocità sui tratti di strada incidentati e attraversamento strade principali

	Periferico	Periferico (RE)	Hinterland	Poli Urbanizzati	Periurbani	Totale
Velocità Media sulle strade incidentate	46,7	58,8	59,0	53,9	58,0	56,1
% Comuni attraversati da Strade principali – STATALI	9,8	13,1	25,0	38,0	24,3	21,4
% Comuni attraversati da Strade principali - PROVINCIALI 1 LIV	17,9	28,4	32,1	52,5	36,4	33,9
% Comuni attraversati da Strade principali - PROVINCIALI 2 LIV	23,2	29,4	24,1	43,7	35,1	32,1

Fonte: elaborazione Ires su dati, Istat, Regione Piemonte e TOC

3.2.2 Il database degli archi stradali incidentati: un esempio applicativo del dbss

Per esemplificare le possibilità di utilizzo del DBSS, da questo è stata predisposta una nuova base dati che contiene tutti i **segmenti di strada** sui quali è avvenuto almeno un incidente nel periodo 2013-2016. Oltre all'identificativo e alla tipologia di comuni di appartenenza (cluster), le variabili riportate in questo database riguardano aspetti riguardanti, Fig.11:

- il profilo infrastrutturale (TGM, velocità media, lunghezza, tipo di strada);
- il profilo incidentale: numero di incidenti, infortunati e vittime; incidenti e infortunati per natura dell'incidente (scontro tra veicoli, investimento di pedoni, ostacoli ed errori di guida) e luogo dell'incidente (presenza o meno di intersezione, rettilineo o curva e segnaletica);
- il profilo territoriale (densità e tipologia di edificato nella zona attraversata dall'arco stradale).

Fig. 11 Screen shot del database dei segmenti di strada incidentati

Descrizione segmento da DB CSI															CLUSTER				
Cluster	incidenti			Luogo dell'incidente VITTIME						Luogo dell'incidente VITTIME						Cluster			
	N_incidenti	Morti	Vittime	incrocio/rotatoria (1,2)	int_segnaleta (3,4)	int_nosegnalata (5)	NoInt_Rett ilineo (7)	NoInt_Curva (8)	Altro (6,9,10,11,12)	frontale	frontale-laterale	laterale	tamponamento	pedoni	veicolo fermo o in arresto		ostacolo	fuoruscita	frenata/caduta
4	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
4	2	0	3	0	2	0	0	1	0	0	2	0	0	0	0	0	0	1	0

Luogo dell'incidente INCIDENTI															Profilo territoriale da SezCensi		
natura incidente INCIDENTI															Popolazione	Tot_edifici	Quota_EdificiBassi(1,2,3 piani)
8929	9562	773	19488	4419	494	2017	14791	3890	7766	5848	1809	2285	4334	925			
incrocio/rotatoria (1,2)	int_segnaleta (3,4)	int_nosegnalata (5)	NoInt_Rett ilineo (7)	NoInt_Curva (8)	Altro (6,9,10,11,12)	frontale	frontale-laterale	laterale	tamponamento	pedoni	veicolo fermo o in arresto	ostacolo	fuoruscita	frenata/caduta			
0	0	0	1	0	0	0	0	0	0	0	0	1	0	0			
0	1	0	0	1	0	0	1	0	0	0	0	0	1	0			
															153	33	0,76
															9	4	0,75

Questo database contiene 27902 segmenti che rappresentano il 5.6% dei tratti della rete della viabilità regionale. La loro distribuzione per tipologia dei comuni e tipo di strade è mostrata in Tab.5.

Tab.5 Tratti di strada con almeno 1 incidente nel periodo 2013-16 per tipologia dei comuni e tipo di strada

Gruppo	Autostrada	Comunali		Provinciali		Statali		Altro	Totale	
		Urbane	Extra-urbane	Urbane	Extra-urbane	Urbane	Extra-urbane			
Periferico		4	89	4	277	56	45	10	2	487
Periferico RE		18	78	41	142	1027	9	136	7	1458
Hinterland		462	249	36	236	521	33	45	3	1585
Poli urbanizzati		653	15249	295	1456	1974	259	302	74	20262
Periurbano		108	774	104	986	1642	171	299	12	4096
non assegnato		14								14
Totale		1259	16439	480	3097	5220	517	792	98	27902

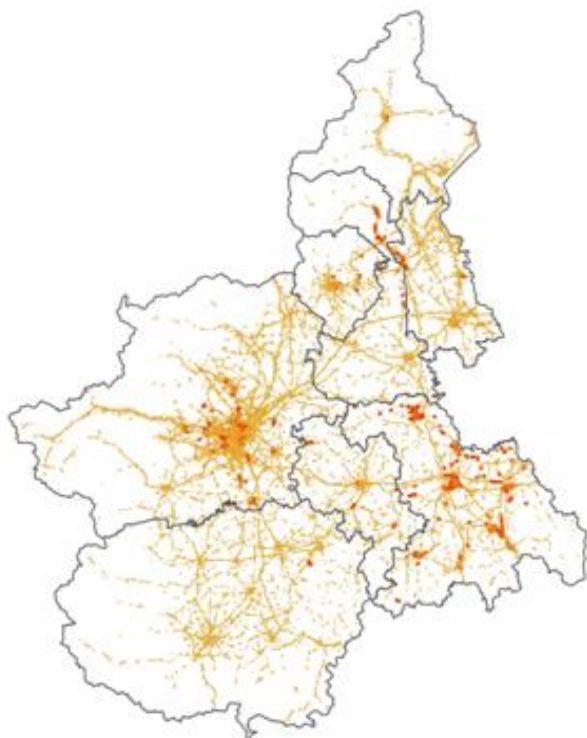
Fonte: elaborazione Ires su dati, Istat e Regione Piemonte

La mappa di Fig. 12 visualizza questi tratti di strada, evidenziando in rosso quelli sui quali si sono verificati più di 4 incidenti nel periodo preso in esame.

Un'analisi congiunta delle informazioni per arco stradale e per sezione censuaria mostra come sui tratti con un numero elevato d'incidenti, il traffico sia mediamente più intenso e più veloce, Tab. 6.

I risultati riportati in tabella mostrano poi come i tratti a elevata incidentalità attraversino aree nelle quali l'urbanizzazione è caratterizzata da una presenza più elevata di edifici multi-piano e di edifici non residenziali.

Fig. 12 Segmenti delle strade piemontesi sui quali si sono verificati incidenti nel periodo 2013-16



Nota: sono visualizzati solo i segmenti con almeno 1 incidente nel periodo 2013-16. I tratti in rosso hanno più di 4 incidenti.

Tab. 6 Confronto del traffico e di alcune caratteristiche insediative per i tratti di strada a bassa e alta incidentalità

		Media regionale	Livelli di incidentalità	
			Bassa: tratti con al massimo 4 incidenti	Alta: tratti con più di 4 incidenti
Traffico	Traffico Giornaliero Medio	2750	6110	12294
	Velocità Media	49,6	54,9	56,3
Caratteristiche insediative	Rapporto tra edifici alti e edifici bassi	8,8	11,3	18,9
	% edifici non residenziali sul totale	15,5	23,4	36,2
	% edifici fino a 3 piani sul totale	50	37,8	22,7

Nota: in questo caso si utilizzano dati provenienti da Dbincidenti, DbStrade e DbSezCensimento

Nel seguito, si forniscono alcuni esempi pratici di utilizzo del database dei tratti di strada, mostrato in Fig.12. Essi rispondono ad alcune domande circa la quantificazione e qualificazione degli eventi incidentali dal *punto di vista* (molto puntuale) dei segmenti di strada dove si sono verificati.

3.2.3 Casi d'uso

Caso 1 – Quanti sono i tratti di strada ad alta incidentalità (con più di 4 incidenti ? Come si distribuiscono tra i profili comunali individuati dalla Cluster Analysis?

I segmenti con la maggior concentrazione di incidenti sono 739, pari al 2,8% del totale.

Nonostante questa ridotta numerosità su di essi accadono il 12% degli incidenti (4846) che causano il 12,6% (7414) degli infortuni e il 7,9% dei morti (61).

Questi tratti di strada si trovano per la quasi totalità (94%) nei comuni che appartengono al gruppo Poli Urbanizzati. Per questo gruppo, i tratti ad alta incidentalità rappresentano circa il 3% dei tratti totali, sui quali avvengono oltre il 14% degli incidenti e il 12% delle morti.

Tab.7 Tratti di strada ad alta incidentalità nei gruppi: incidenza percentuale dei tratti, degli incidenti, degli infortunati e dei morti sui rispettivi totali

Tipologia comune	N tratti	Incidenza % sul totale			
		Tratti di strada	Incidenti	Infortunati	Morti
Periferico	5	1,0	4,8	5,7	0,0
Periferico (RE)	13	0,9	4,7	4,7	4,4
Hinterland	3	0,3	1,2	1,3	0,0
Polo Urbanizzato	695	3,5	14,4	15,3	12,3
Periurbano	23	0,6	2,8	2,9	3,7
Totale	739	2,8	12,0	12,6	7,9

Caso 2 – Come si caratterizzano gli eventi incidentali sui segmenti di strada ad alta incidentalità, dal punto di vista della natura dell'evento e del luogo in cui avvengono?

I tratti con elevata incidentalità si distinguono per una presenza ancor più decisa degli scontri frontali-laterali (40% degli incidenti) e dei tamponamenti (20%).

Per quanto riguarda il luogo dell'incidente, gli archi con alta incidentalità sono maggiormente critici nelle intersezioni segnalate con semaforo: qui avvengono il 35% degli eventi che causano il 37% degli infortuni.

Incroci e rotonde sono i luoghi relativamente più critici per gli archi altri a minore frequenza incidentale.

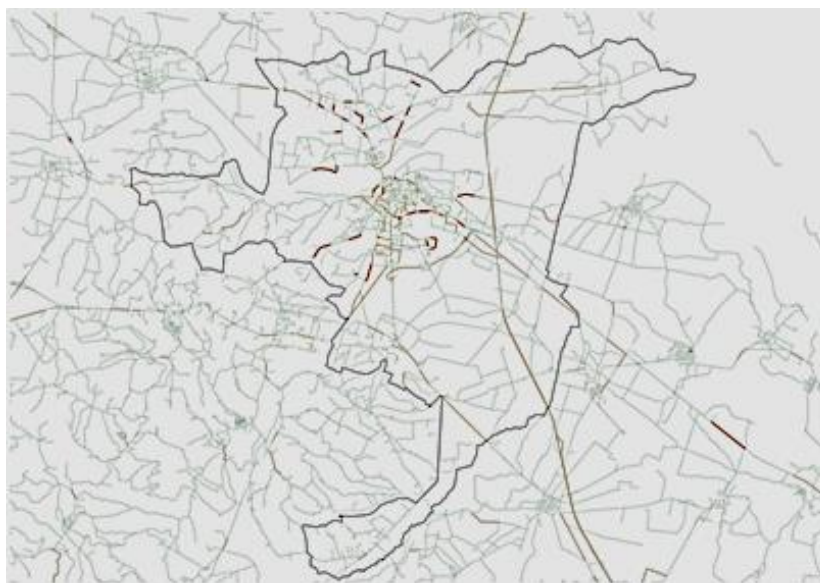
Tab. 8 Confronto tra tratti di strada a bassa e alta incidentalità: natura e luogo dell'incidente

			Incidenti		Infortunati	
			Tratti di strada a bassa incidentalità	Tratti di strada ad alta incidentalità	Tratti di strada a bassa incidentalità	Tratti di strada ad alta incidentalità
Natura dell'incidente	Scontro	Frontale	5,0	3,9	6,6	5,1
		Frontale-laterale	35,7	39,7	38,2	43,8
		Laterale	9,2	7,5	8,5	6,4
		Tamponamento	15,8	19,6	18,3	21,5
	Investimento Pedoni		14,5	13,7	11,1	10,1
	Veicolo fermo o ostacolo		9,0	7,0	8,1	6,0
	Errori di guida (fuoriuscita/frenata/caduta)		10,8	8,5	9,2	7,1
	Totale		100,0			
Luogo dell'incidente	Intersezione	Incrocio/rotatoria	23,1	14,0	21,9	13,9
		Segnalata o semaforo	22,1	34,9	23,8	37,0
		non segnalata	2,0	1,1	2,0	1,0
	Non intersezione	Rettilineo	42,5	41,4	41,7	39,0
		Curva	9,1	7,8	9,4	8,2
	Altro (dosso, galleria, pendenza...)		1,2	0,8	1,2	0,9
	Totale		100,0			

Nota: in questo caso si utilizzano dati provenienti da Dbincidenti e DbStrade

Caso 3 – Come può essere utilizzato il DB per esaminare le criticità incidentali a livello sub-comunale?

Consideriamo come esempio, il comune di Casale Monferrato che appartiene al gruppo Poli Urbanizzati; nella provincia di Alessandria questo gruppo comprende il 10% dei comuni che hanno avuto incidenti nel periodo 2013-16 (Tab. 1b). La mappa dei segmenti di strada incidentati del comune (un ingrandimento della mappa di Fig.12) è mostrata in Fig. 13.

Fig. 13 Segmenti delle strade di Casale Monferrato sui quali si sono verificati incidenti nel periodo 2013-16


La Tab. 9 presenta un bilancio riassuntivo della situazione incidentale per tipo di strada. Essa mostra, ad esempio, come le strade comunali urbane concentrino il numero di gran lunga più elevato di incidenti. Escludendo l'autostrada, il maggior numero di sinistri sono avvenuti sulla SP31 del Monferrato e sulla SP457 di Moncalvo, entrambe strade provinciali di interesse regionale di I livello.

Tab. 9 Numero di incidenti e di infortunati per tipo di strada del comune di Casale Monferrato, nel periodo 2013-16

incidenti	Altro	Autostrada	com_EU	com_U	prov_EU	prov_U	Totale	Indice di infortunio
Comunali e altre provinciali	1		3	293	10	2	309	128
A26 Genova Voltri - Gravellona Toce		22					22	195
SP 26 Balzola - Casal Popolo					2		2	200
SP 31 Bis del Monferrato di Chivasso						5	5	120
SP 31 del Monferrato					9	8	17	182
SP 43 San Germano - Pozzo Sant'Evasio					1	1	2	150
SP 457 di Moncalvo					8		8	163
SP 457 di Moncalvo - VAR di Casale Monferrato					6		6	217
SP 50 Alessandria - Casale					5	3	8	175
SP 54 Casale - Ticineto					4		4	100
SP 55 Casale - Valenza				1	6	1	8	200
SP 596 dei Cairoli (Diramazione Casale - Mortara)					3	2	5	160
SP 7 Casale - Trino						1	1	100
Totale complessivo	1	22	3	294	54	23	397	139
Infortunati (feriti + morti)	Altro	Autostrada	com_EU	com_U	prov_EU	prov_U	Totale	Somma di morti
Comunali e altre provinciali	3		3	368	18	3	395	3
A26 Genova Voltri - Gravellona Toce		43					43	3
SP 26 Balzola - Casal Popolo					4		4	0
SP 31 Bis del Monferrato di Chivasso						6	6	1
SP 31 del Monferrato					16	15	31	0
SP 43 San Germano - Pozzo Sant'Evasio					2	1	3	0
SP 457 di Moncalvo					13		13	0
SP 457 di Moncalvo - VAR di Casale Monferrato					13		13	1
SP 50 Alessandria - Casale					9	5	14	1
SP 54 Casale - Ticineto					4		4	0
SP 55 Casale - Valenza				2	13	1	16	2
SP 596 dei Cairoli (Diramazione Casale - Mortara)					6	2	8	1
SP 7 Casale - Trino						1	1	0
Totale complessivo	3	43	3	370	98	34	551	12

Concentrando l'attenzione sulla SP31, la Tab.10 ne dettaglia alcuni aspetti profilo incidentale.

Tab. 10 Numero di infortunati e percentuali per alcune sui tratti incidentati della SP31

Arco	Infortunati	% Incrocio/rotatoria (1,2)	% Int segnalata (3,4)	% Nolnt_Rettilineo (7)	% scontro (frontale/laterale/front-lat)	% frontale-laterale	% tamponamento
093CB26A-1508-2909-E054-0003BA0F36E6	9	33	33	33	44	44	56
093CB26B-5EA4-2909-E054-0003BA0F36E6	3	0	0	33	100	0	0
03F7D1B1-F630-6DDE-E054-0003BA0F36E6	2	0	100	0	0	0	100
03F7D1B1-F676-6DDE-E054-0003BA0F36E6	2	0	0	100	0	0	100
03F7D1B1-F694-6DDE-E054-0003BA0F36E6	1	0	0	100	0	0	100
03F7D1B1-F70C-6DDE-E054-0003BA0F36E6	3	0	0	100	0	0	100
03F7D1B1-F71B-6DDE-E054-0003BA0F36E6	1	0	0	100	0	0	100
093CB264-9D67-2909-E054-0003BA0F36E6	1	0	0	100	0	0	100
093CB265-2822-2909-E054-0003BA0F36E6	2	0	0	100	100	100	0
093CB265-CA34-2909-E054-0003BA0F36E6	1	0	100	0	100	100	0
093CB265-CA48-2909-E054-0003BA0F36E6	1	0	100	0	0	0	0
093CB26A-16B1-2909-E054-0003BA0F36E6	2	100	0	0	100	100	0
093CB26B-43F1-2909-E054-0003BA0F36E6	2	0	0	100	0	0	100
093CB26B-5E77-2909-E054-0003BA0F36E6	1	0	0	100	0	0	100
211E7019-BDA5-605D-E054-0003BA0F36E6	3	0	0	100	0	0	100
F840954A-2961-49D1-E044-0003BA0F36E6	1	0	0	100	0	0	100
F840954A-29CF-49D1-E044-0003BA0F36E6	1	0	0	100	0	0	100
F840954A-2C5E-49D1-E044-0003BA0F36E6	1	0	100	0	100	100	0
Totale %	100	14	22	59	35	27	62
Totale va	37	5	8	22	13	10	23

4. PROSPETTIVE DI APPROFONDIMENTO

L'analisi condotta e i risultati ottenuti suggeriscono due piste possibili di approfondimento futuro:

- ottimizzare il metodo di integrazione delle basi informative
- affinare gli strumenti analitici

4.1 Ottimizzare il metodo di integrazione basi dati

Il percorso metodologico utilizzato per costruire il DBSS ha permesso di investigare a fondo i problemi operativi che esistono per costruire e mantenere nel tempo una base informativa a supporto dell'analisi incidentale in Piemonte.

Essa ha, però carattere sperimentale e in quanto tale ha ampi margini di miglioramento.

In un'ottica *user centered* e in considerazione degli stringenti vincoli di bilancio delle PA italiane, sarebbe bene che il percorso migliorativo prestasse attenzione a:

- abbassare le barriere all'entrata in termini di competenza degli utenti;
- utilizzare dotazioni hardware e software già disponibili o comunque liberamente accessibili.

Ciò premesso, le direzioni più urgenti per ottimizzare la performance computazionale e l'usabilità della metodologia sono:

1. Migliorare l'algoritmo di assegnazione degli incidenti ai segmenti stradali.

La funzione GIS per il calcolo della minima distanza tra un punto (l'incidente) e una linea (il segmento di strada) necessaria per integrare DbStrade e DbIncidenti ha richiesto la trasformazione, molto pesante dal punto di vista computazionale, dei segmenti in punti per migliorare il risultato dell'algoritmo.

Una possibile soluzione a questo problema è di implementare una funzione, richiamabile dalla console python di Qgis, che esegua in modo efficiente l'algoritmo di calcolo della minima distanza da un punto a una retta basato sulla minimizzazione delle distanze dal punto alle linee passanti per gli estremi di un numero ristretto di segmenti prossimi al punto.

2. Sviluppare una funzionalità appropriata per stimare e assegnare i dati di flusso ai segmenti di strada non valorizzati.

Si tratta di una funzionalità indispensabile per utilizzare e valorizzare i dati di incidentalità per arco stradale; essa consentirebbe inoltre di calcolare indicatori plausibili dei valori potenziali di esposizione al rischio di incidentalità di un'area.

3. Costruire un'interfaccia appropriata per collegare i diversi dataset.

Questa applicazione avrà il doppio vantaggio di conferire solidità e replicabilità agli algoritmi di assegnazione e facilitare le operazioni di data-linkare. Consentirà, attraverso lo sviluppo di un'interfaccia di facile utilizzo, di alimentare e interrogare la base dati da parte di utenti non specialisti.

Quanto al collegamento con altri data-set tre paiono prioritari:

- il collegamento con i dati sanitari per identificare la gravità dell'infortunio incidentale. Si tratta di dar seguito al progetto pilota della regione Piemonte, previsto nel Piano Statistico Nazionale¹¹. Questa operazione è necessaria per determinare i costi sociali del fenomeno e affinare gli indicatori di criticità incidentale della rete;
- il collegamento con i dati economico-finanziari riguardanti i progetti di miglioramento infrastrutturale e di manutenzione (straordinaria e ordinaria) della rete. Si tratta di un prerequisito indispensabile per realizzare appropriati bilanci costi-benefici degli interventi di messa in sicurezza della rete stradale¹² ;
- il collegamento con le multe elevate dalle Forze dell'Ordine e/o dai dispositivi di controllo sui vari tratti di strada. E' questa un'operazione che può essere utile non solo per monitorare l'impatto di alcuni interventi riguardanti la segnaletica sulla velocità, ma anche per identificare, preventivamente, tratti di strada potenzialmente più a rischio.

4.2 Approfondire l'analisi

Pur con tutti i caveat del caso, si auspica che i risultati di questo lavoro (questo report e il database per arco stradale) possano essere utilizzati, già da ora, in casi pratici di analisi e di intervento.

Queste applicazioni potranno costituire test molto utili per identificare le priorità degli approfondimenti da realizzare nei prossimi mesi.

A questo proposito, una pista interessante da seguire riguarda lo **studio dei casi d'uso** che, dal punto di vista degli operatori della sicurezza stradale del Piemonte, sarebbe opportuno privilegiare, nel quadro inoltre della realizzazione del nuovo Piano regionale della mobilità e dei trasporti.

Quanto agli affinamenti possibili, essi sono numerosi. Qui ci limitiamo a citarne due entrambi i quali sono orientati a migliorare lo **studio delle relazioni tra determinanti di incidentalità e livelli di rischio incidentale**. In questa direzione sarebbe opportuno:

- esplorare altre metodologie di clustering che tengano conto della dinamicità dei processi; ovvero che sappiano costruire la tipologia non solo sulla base dei valori delle variabili descrittive selezionate, ma anche della loro evoluzione nel tempo in relazione ai cambiamenti del conteso territoriale di appartenenza;
- utilizzare tecniche di regressione lineare (e spaziale) per studiare l'intensità delle relazioni tra determinanti e rischi incidentali e la loro variabilità in relazione alle diverse sub-aree regionale (ad esempio gli ambiti di trasporto).

PIE-00008 Data linkage dei record degli infortunati negli incidenti stradali regionali con i dati del sistema. In ISTAT - Programma Statistico Nazionale 2017-19. sanitario regionale http://www.sistan.it/fileadmin/Repository/Home/PSN/Programma_statistico_nazionale/Psn_2017-2019/Vol._2_Psn_2017-2019.pdf (p. 129)

¹² in proposito si veda il recente lavoro dell'Ires (Grieco e Occelli, 2018)

BIBLIOGRAFIA

- Adanu E.K., Smith R., Powell L., Jones S. (2017) Multilevel analysis of the role of human factors in regional disparities in crash outcomes, *Accident Analysis and Prevention*, 109, 10-17.
- Anderson T.L. (2009) Kernel density estimation and K-means clustering to profile road accident hotspots, *Accident Analysis and Prevention*, 41, 359-364.
- Beanland V., Fitzharris M., Young K.L., Lenné M.G.(2013) Driver inattention and driver distraction in serious casualty crashes: Data from the Australian National Crash In-depth Study, *Accident Analysis and Prevention*, 54, 99-107.
- Chen F., Wu J., Chen X., Wang J., Wang D (2016) Benchmarking road safety performance: Identifying a meaningful reference (best-in-class). *Accident Analysis and Prevention*, 86, 76-89.
- Eksler V., Lassarre S.(2008) Evolution of road risk disparities at small-scale level: Example of Belgium. *Journal of Safety Research*, 39, 417-427.
- Eksler V. (2008) Regional analysis of road mortality in Europe. *Public Health*, 122, 826-837.
- Eksler V. (2010) Measuring and understanding road safety performance at local territorial level, *Safety Science*, 48, 1197-1202.
- ETSC (2018) RANKING EU PROGRESS ON ROAD SAFETY. 12th Road Safety Performance Index Report. https://etsc.eu/wp-content/uploads/PIN_ANNUAL_REPORT_2018_final.pdf
- Hakkert A.S., Braimaister L. (2002) The use of exposure and risk in road safety studies. SWOV Institute for Road Safety Research, R-2002-12.
- Hermans E., Van den Bosscheb F., Wets G.(2008) Combining road safety information in a performance index. *Accident Analysis and Prevention*, 40, 1337-1344.
- Kaygisiz O., Senbil M., Yildiz A. (2017) Influence of urban built environment on traffic accidents: The case of Eskisehir (Turkey), *Case Studies on Transport Policy*, 5, 306-313.
- Landini, Occelli, Scalzotto (2017) Quanto è sicuro il IPiemonte? Uno sguardo alla situazione dell'incidentalità stradale a metà del decennio. <http://archive.digibess.eu/islandora/object/librib:911652/datastream/PDF/content/librib%20911652.pdf>
- Grieco A., Occelli S. (2018) SICUREZZA STRADALE E VALUTAZIONE: PERCORSI POSSIBILI (in pubblicazione).
- Larsen L, Kines P.(2002) Multidisciplinary in-depth investigations of head-on and left-turn road collisions. *Accident Analysis and Prevention*, 34, 367-380.
- Mannering F.L. , Bhat C.R. (2014)Methodological frontiers and future directions. *Analytic methods in accident research*, 1,1-22
- McAndrews C., Beyer K., Guse C.E., Layde P. (2016) How do the definitions of urban and rural matter for transportation safety? Re-interpreting transportation fatalities as an outcome of regional development processes. *Accident Analysis and Prevention*, 97, 231-241.
- MacQueen J.B.(1967) Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1:281-297.
- Ministero delle Infrastrutture e dei Trasporti – Direzione generale per la sicurezza stradale (2014) PNSS Orizzonte 2020. <http://www.mit.gov.it/documentazione/piano-nazionale-sicurezza-stradale-orizzonte-2020->
- Mondaldo C., Occelli S. (2018) ELEMENTI DI ATTENZIONE PER LA FORMULAZIONE DELLE NUOVE POLITICHE SULLA SICUREZZA (in pubblicazione).
- Ng K., Hung W., Wong W. (2002) An algorithm for assessing the risk of traffic accident. *Journal of Safety Research*, 33, 387- 410
- Regione Piemonte (2018) Piano regionale della mobilità e dei trasporti. http://www.regione.piemonte.it/trasporti/dwd/PRMT/20180116_DCR_All_A_PRMT.pdf
- World Health Organization (2004) World report on road traffic injury prevention. http://www.who.int/violence_injury_prevention/road_traffic/activities/training_manuals/en/
- World Health Organization (2006) Risk factors for road traffic injuries. In Road Safety Training Manual. http://www.who.int/violence_injury_prevention/publications/road_traffic/world_report/en

NOTE EDITORIALI

Editing

IRES Piemonte

Ufficio Comunicazione

Maria Teresa Avato

© IRES

Giugno 2019

Istituto di Ricerche Economico Sociali del Piemonte

Via Nizza 18 -10125 Torino

www.ires.piemonte.it

si autorizzano la riproduzione, la diffusione e l'utilizzazione del contenuto con la citazione della fonte.

Ambiente e Territorio

Cultura

Finanza locale

Immigrazione

Industria e Servizi

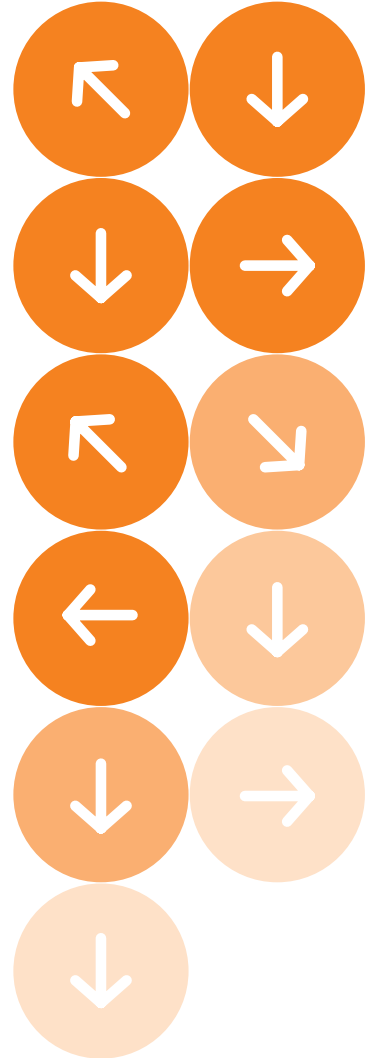
Istruzione e Lavoro

Popolazione

Salute

Sviluppo rurale

Trasporti



IRES Piemonte

Via Nizza, 18

10125 TORINO

+39 0116666-461

www.ires.piemonte.it