

The Harsh Truth: Segment-Level Analysis of Harsh Driving Events in Milan Using Large-Scale Telematics, Street Networks, and Google Street View

Andrea La Grotteria^a, Paolo Santi^{a,c}, Titus Venverloo^a, Umberto Fugiglando^a, Carlo Ratti^{a,b}

^a*Senseable City Lab, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, 02139, MA, USA*

^b*Politecnico di Milano, Department ABC, Via Giuseppe Ponzio 31, Milano, 20133, Italy*

^c*Consiglio Nazionale delle Ricerche, Via G. Moruzzi 1, Pisa, 56124, Italy*

Abstract

Police-reported crash statistics remain the standard input for urban road-safety assessment, but their incompleteness and reporting lag limit their usefulness for timely, fine-grained intervention design. Harsh acceleration and braking events are widely used as surrogate safety indicators but have so far been studied in comparatively small urban samples. This study analyses harsh events across the urban road network of Milan, Italy, combining high-resolution telematics from more than 4.2 million vehicles equipped with On-Board Units (collected via UnipolTech over six representative weeks of 2023), segment-level traffic metrics from TomTom, street-network and infrastructure attributes from OpenStreetMap, and visual streetscape features extracted from Google Street View via semantic segmentation with a OneFormer model. The harsh events are map-matched to 24,673 OSM segments. We employ an analytical framework that combines non-parametric Mann–Whitney U tests of segment-feature distributions between high- and low-harshness groups, supervised machine-learning regressors (Gradient Boosting, Random Forest) with SHAP-based feature attribution, and an exposure-controlled OLS specification for the cycling-infrastructure case study. We find that once exposure is controlled, wider carriageways, crossings and transit stops, and more open visual fields (higher sky- and road-pixel proportions) are associated with more harsh events, while denser built frontage is associated with fewer. Absolute harsh scores peak during commute hours, while per-vehicle risk peaks on weekend nights, highlighting the importance of exposure normalisation. Finally the cycling-infrastructure case study finds an observed gradient in harsh-event intensity across facility types: markings-only cycle lanes are associated with a 19.5% higher harsh score and shared-with-vehicles configurations with an 11.5% higher score relative to physically separated cycle paths, conditional on the included controls. The results support context-specific, rather than uniform, urban-safety interventions and illustrate how large-scale telematics combined with open geospatial and visual data can inform Vision Zero decisions at the metropolitan scale.

1. Introduction and Relevant Literature

Urban road safety has gained renewed policy attention, particularly with respect to vulnerable road users such as pedestrians and cyclists. The European Commission’s Vision Zero strategy (CINEA 2025) sets the long-term goal of eliminating traffic fatalities and serious injuries, and is typically implemented through infrastructure upgrades, public awareness activities, and policy reform coupled with enforcement. Several European cities, including Stockholm, Copenhagen, Amsterdam, Zürich, and Barcelona, have operationalised this framework through measures such as city-wide 30 km/h speed limits, restrictions on motor-vehicle access, and street redesign (Yannis and Michelaraki 2024). In Italy, Bologna has recently adopted a 30 km/h limit on most inner streets (Comune di Bologna 2023), while similar proposals are being debated in Milan and other cities. The issue is of particular relevance in the Italian context, where traffic fatality rates remain persistently above the European Union average (Eurostat 2025) and a substantial share of traffic injuries occurs in urban settings.

Email address: lagro@mit.edu (Andrea La Grotteria)

Reduced urban speed limits are an effective safety measure: Yannis and Michelaraki (2024) review European experiences with city-wide 30 km/h zones and report an associated 23% reduction in injury crashes, with no increase in congestion or travel time. Complementary levers such as automated enforcement, vertical calming, credible design (Sadeghi-Bazargani and Saadati 2016; Yao et al. 2019) and a range of contextual factors (temporal exposure, land use, network connectivity, intersection control (Cabrera-Arnau et al. 2020; Greibe 2003; Iranmanesh et al. 2024; Ng et al. 2001; Vorko-Jović et al. 2006)) further shape urban crash risk.

Conventional crash-risk evaluations rely almost exclusively on police-reported crashes. Police reports are, however, frequently incomplete, unavailable, or lag by months, which limits their usefulness for timely safety planning. This has motivated a growing interest in *surrogate safety indicators* – metrics capable of flagging crash risk before crashes occur. Harsh braking and acceleration events (harsh events) are widely used as proxies for aggressive or unsafe driving (Silva and Eugenio Naranjo 2020) and lend themselves to continuous monitoring through vehicle telematics.

Longitudinal accelerations and decelerations exceeding roughly $1.5\text{--}2\text{ m s}^{-2}$ are routinely treated as indicators of risky manoeuvres in the literature (Silva and Eugenio Naranjo 2020). Their spatial structure has been examined by Ziakopoulos (2021) using smartphone trajectory data for an urban suburb of Athens. The authors detected positive spatial autocorrelation of both harsh brakings (HBs) and harsh accelerations (HAs). Geographically Weighted Poisson Regression revealed that segment length and traffic exposure (pass count) are the strongest positive predictors of HBs, while gradient and neighbourhood complexity have negative effects. A follow-up study (Ziakopoulos et al. 2022) confirmed these trends and added that two-lane roads tend to produce more harsh brakings than single-lane segments, while low-volume residential streets produce fewer. Collectively, these analyses demonstrate that harsh events mirror features of the built environment and can therefore support the identification of potentially unsafe locations in advance of recorded crashes. Three data-side limitations nonetheless remain. First, the analysis drew on fewer than 1,500 drivers of a single smartphone application, which constrains its spatial coverage and statistical power. Second, the built environment was represented through OSM geometry alone, without visual streetscape information. Third, traffic exposure was controlled through pass counts rather than through an independent measure of traffic volume. The present study addresses these three data-side limitations.

Harsh events in the city of Milan are here analysed using high-resolution telematics records from more than 4.2 million vehicles equipped with On-Board Units are provided by UnipolTech, the telematics branch of one of Italy’s largest insurance groups. These data are integrated with segment-level traffic metrics from TomTom, network geometry and infrastructure attributes from OpenStreetMap (snapshot of 31 December 2023), and visual streetscape features extracted from Google Street View (GSV) imagery using a state-of-the-art semantic-segmentation model. The combined dataset covers 24,673 urban road segments and 3,407,876 map-matched harsh events recorded over six representative weeks of 2023.

The analysis combines non-parametric statistical tests, supervised machine-learning regressors with SHAP-based feature attribution, and an exposure-controlled OLS specification for the cycling-infrastructure case study. A dedicated case study on cycling infrastructure examines how different designs are associated with harsh-event intensity on adjacent road segments.

The analysis is organised around three research questions:

- RQ1.** Do large-scale telematics-based harsh-event measures reproduce the segment-level associations previously reported from smaller smartphone datasets, and does adding visual streetscape descriptors improve their explanatory power?
- RQ2.** After controlling for traffic exposure, how do road geometry, regulatory attributes, and visual openness relate to harsh-event intensity across the urban network?
- RQ3.** Can we use harsh events to inform on specific policies decisions, such as cycling infrastructure?

The remainder of the paper is organised as follows. Section 2.2 outlines the methodological background of the statistical and machine-learning models employed. Section 2.1 describes the data sources and Section ?? the processing and integration pipeline. Section ?? summarises the final spatial dataset. Section 3 presents the results of the statistical and machine-learning analyses, together with the cycling-infrastructure case study. Section 4 discusses the implications and limitations, and Section ?? concludes and outlines directions for future research.

2. Data and Methods

2.1. Data sources

The analysis integrates five complementary datasets: telematics-based harsh events from UnipolTech, segment-level traffic metrics from TomTom, street-network and infrastructure attributes from OpenStreetMap, the Milan Municipal Cycle Network, and Google Street View imagery. Each source and its relevant attributes are described below.

2.1.1. OpenStreetMap: street network and infrastructure

Street-network topology and infrastructure attributes were retrieved from *OpenStreetMap* (OSM) for the municipality of Milan, using the *osmnx* library (Boeing 2025). The snapshot corresponds to 31 December 2023 and comprises 18,037 nodes and 35,606 directed edges with a total length of approximately 3,640 km and an average segment length of 102 m. For each segment the dataset provides the geometry, number of lanes, street classification (e.g. residential, tertiary), posted speed limits (where available), traffic directionality, and the presence of sidewalks and cycleways.

Since the analysis targets urban driving, motorway segments (605 in total) were excluded, together with segments that did not contain any matched UnipolTech observation. The final OSM network comprises 24,673 segments.

Three feature groups were derived for every OSM segment. First, *geometric* features consist of segment length and sinuosity, defined as

$$\text{sinuosity} = \frac{\text{segment length}}{\|(x_d, y_d) - (x_o, y_o)\|}, \quad (1)$$

where (x_o, y_o) and (x_d, y_d) denote the coordinates of the origin and destination nodes. Second, *streetscape-context* features count the infrastructure elements within a 20 m buffer around each segment: traffic lights, pedestrian crossings, public-transport stops, cycleways, sidewalks, playgrounds, and green areas. Third, *road form and regulation* features comprise the grouped highway tag, the number of lanes, and an estimated street width computed via the *momepy.StreetProfile* method (Fleischmann 2019) following Araldi and Fusco (2019). The distribution of estimated street widths across Milan is shown in Figure 1.

2.1.2. UnipolTech: telematics harsh events

The primary source of driving-behaviour data is *UnipolTech*, the telematics branch of Unipol – one of the largest insurance groups in Italy. Data are collected from a nationwide fleet of more than 4.2 million vehicles equipped with On-Board Units (OBUs) that record sudden braking and rapid acceleration events. Each event record contains a timestamp, the event latitude and longitude, instantaneous speed, vehicle heading (degrees from North), the event type (braking or acceleration), a measure of severity, a GNSS signal-quality flag, the device type, and the vehicle category.

Two OBU models are used in the fleet:

- *Targa Telematics* devices measure acceleration directly from an on-board accelerometer and assign each event to one of four severity classes, with class boundaries (in cm s^{-2}) of 350–450, 450–550, 550–700, and above 700.
- *MetaSystem* devices estimate acceleration from GNSS-derived speed changes over a one-second window.

The two device families are harmonised onto the Targa class boundaries to yield a unified severity scale from 1 to 4.

Although the raw dataset covers the whole of Italy, the present study focuses on Milan for three reasons. First, the city offers consistent spatial coverage across all ancillary data sources (OSM, GSV, TomTom). Second, Milan is actively debating Vision Zero-oriented measures such as the extension of 30 km/h zones, which gives the analysis direct policy relevance. Third, restricting the analysis to a single metropolitan area keeps the computation tractable while still yielding several million driving events.

To balance seasonal representativeness against data volume, harsh events were retained for six representative weeks of 2023, corresponding to the first week of January, April, July, October, November, and December.

Each harsh event was assigned to an OSM segment through a two-step matching procedure. First, every OSM segment was sampled at 3 m intervals; for each sampled point p_i the coordinates (x_i, y_i) and the local tangent heading h_i were computed. For each event, the 20 geographically closest sampled points were retrieved, and the candidate with

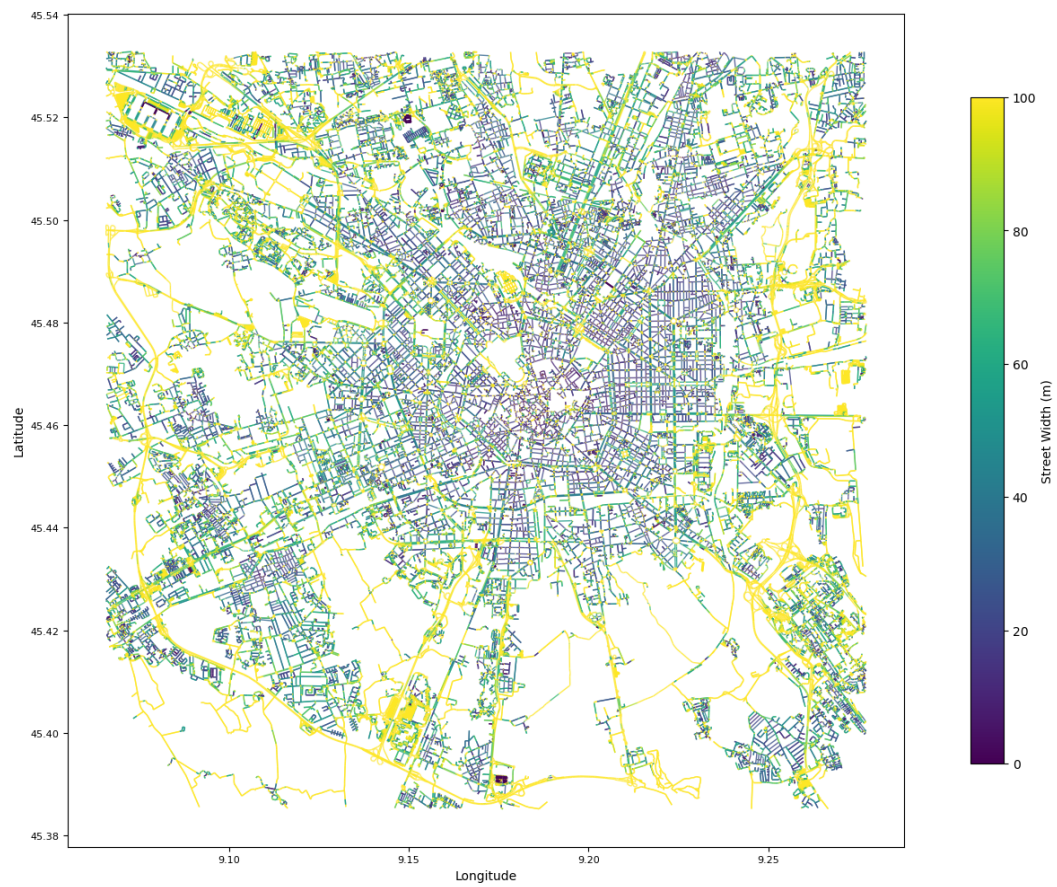


Figure 1: Distribution of estimated street widths across the city of Milan, obtained via the `momepy.StreetProfile` method.

the smallest absolute heading difference $|h_i - h_{\text{obs}}|$ to the observed vehicle heading was selected. A match was retained only if the spatial distance between the event and the selected point was smaller than 6 m and the heading difference smaller than 45° . These thresholds were calibrated on dense urban layouts to minimise mismatches between parallel carriageways and service roads.

Once events were mapped to OSM segments, counts were aggregated at the segment level. To control for exposure, a segment-level exposure metric was computed as

$$\text{Exposure}_s = \text{TrafficVolume}_s \times \text{Length}_s, \quad (2)$$

where TrafficVolume_s is the TomTom probe count for segment s (see Section 2.1.3) and Length_s is the OSM segment length. Because the raw exposure distribution is right-skewed, a log-transform was applied:

$$\text{LogExposure}_s = \log(1 + \text{Exposure}_s). \quad (3)$$

The weighted event count on segment s is defined as

$$\text{HarshScore}_s = n_1 + 2n_2 + 3n_3 + 4n_4, \quad (4)$$

where n_k is the number of events of severity level k . A normalised, log-transformed harshness rate is further defined as

$$\text{HarshRate}_s = \log\left(1 + \frac{\text{HarshScore}_s}{\text{Exposure}_s}\right). \quad (5)$$

The harsh score captures the absolute intensity of risky driving on a segment, while the harsh rate captures its per-unit-exposure intensity; both quantities are reported in the following.

2.1.3. TomTom: segment-level traffic metrics

To control for traffic exposure, segment-level traffic data were obtained from *TomTom*. TomTom aggregates probe data from in-vehicle navigation systems, smartphone applications, and third-party mapping platforms (including Apple Maps). For each of its proprietary segments the dataset provides average and median vehicle speed, posted speed limit, road class and name, speed percentiles (P5, P95), speed standard deviation, and the sample size of contributing probes. The probe sample size is used as a proxy for traffic volume: this is a deliberate simplification, but is supported by TomTom's coverage, which is reported by the data provider to capture signals from a substantial share of the vehicle fleet (TomTom 2025).

TomTom uses its own segmentation (mean length 56.1 m), which is finer than OSM (mean length 102.3 m). Matching between the two networks was carried out in a projected metric CRS (EPSG:32632) using an STR-tree spatial index. For each TomTom segment the 20 nearest OSM candidates were retrieved, the Hausdorff distance and the absolute heading difference were computed, and a match was accepted if the minimum Hausdorff distance $d_{\min}$ satisfied $d_{\min} \leq 10$ m and the heading difference $\Delta\theta$ satisfied $\Delta\theta \leq 45^\circ$. The directed Hausdorff distance between point sets A and B is

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|, \quad (6)$$

and captures the largest deviation between the two geometries. With these thresholds, all 71,867 TomTom segments were matched.

For each TomTom metric M (mean speed, percentiles, speed standard deviation, sample size), the aggregated value for OSM segment s was computed as a length-weighted average:

$$\bar{M}_s = \frac{\sum_{t \in \mathcal{T}(s)} L_t M_t}{\sum_{t \in \mathcal{T}(s)} L_t}, \quad (7)$$

where $\mathcal{T}(s)$ is the set of TomTom segments matched to s and L_t the length of TomTom segment t . Longer TomTom segments covering a greater portion of the OSM edge therefore contribute proportionally more to the aggregated estimate.

2.1.4. City of Milan: municipal cycle network

Cycling-infrastructure data were obtained from the *Milan Municipal Cycle Network* via the city's open data portal. Each record provides the geometry and type of the facility (physically separated cycle paths, on-street cycle lanes with markings only, shared-with-vehicles segments, low-traffic zones), the length, the year of construction, and a regulatory classification. The data was cleaned to remove road crossings and facilities outside the temporal window of the study (keeping only those completed by 2023). Each cycle-path geometry was projected to EPSG:32632 and split into contiguous 3 m fragments, discarding fragments shorter than 10 cm. Each fragment was then matched to its corresponding OSM segment following the procedure of Section 2.1.3.

Three facility types are compared (Figure 2):

1. *Separated cycle path* – physically segregated from motor-vehicle lanes by curbs, barriers, or grade separation (categories *ciclabile sede propria* and *promiscuo pedoni*).
2. *Cycle lane with markings only* – visually delimited lanes on the carriageway, without physical separation (category *ciclabile segnaletica*).
3. *Shared with vehicles* – mixed-traffic roads designated for both cyclists and motor vehicles (category *promiscuo veicoli*).

Of the 92,058 three-metre cycle-infrastructure fragments, 52% are markings-only, 25.8% are separated, and 21.8% are shared-with-vehicles. After removing fragments without matched exposure or harsh-event data, 84,118 fragments entered the analysis.

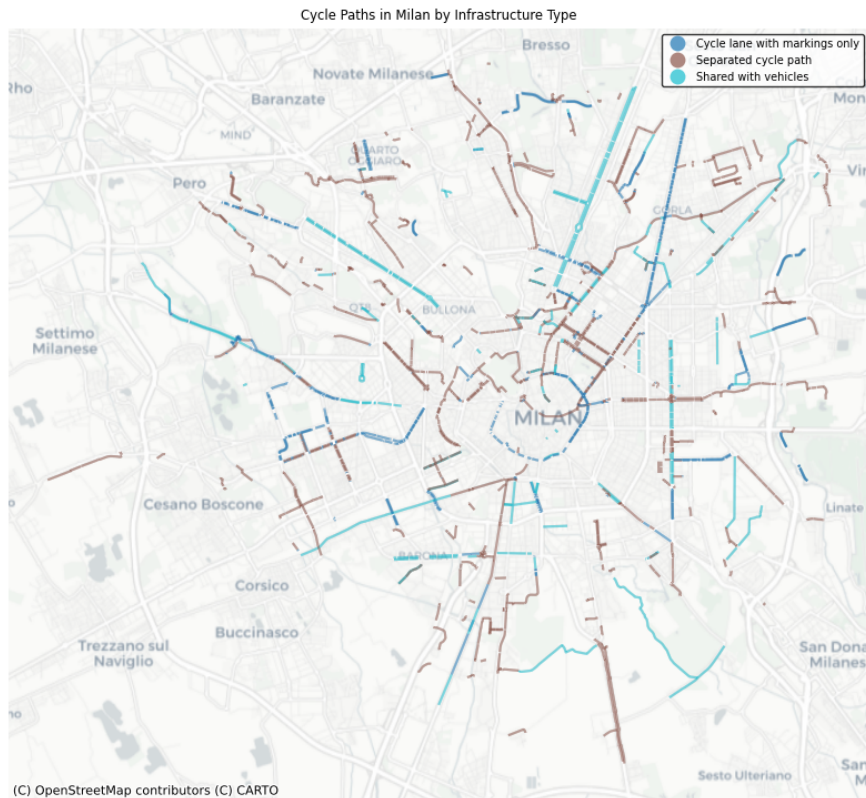


Figure 2: Milan cycling infrastructure by facility type: separated cycle paths, markings-only lanes, and shared-with-vehicles segments.

A log-linear regression was estimated with $\log(\text{HarshScore} + 1)$ as the dependent variable and with $\log(\text{Exposure} + 1)$, posted speed limit, number of lanes, road width, and two facility-type dummies (markings, shared) as predictors.

Separated cycle paths are the baseline:

$$\begin{aligned} \log(\text{HarshScore} + 1) = & \beta_0 + \beta_1 \log(\text{Exposure} + 1) + \beta_2 \text{SpeedLimit} + \beta_3 \text{Lanes} \\ & + \beta_4 \text{RoadWidth} + \beta_5 \text{Markings} + \beta_6 \text{Shared} + \varepsilon. \end{aligned} \quad (8)$$

These data is used to investigate the relationship between cycling-infrastructure design and harsh driving events on the adjacent carriageway.

2.1.5. Google Street View: visual streetscape context

Visual streetscape descriptors were extracted from *Google Street View* (GSV). For each of the 24,673 retained OSM segments, two 640×640 pixel images were retrieved via the GSV API, one in the direction of travel and one in the opposite direction. GSV was preferred over crowdsourced alternatives such as Mapillary because it provides a standardised, high-resolution imaging pipeline, which is essential for pixel-level comparisons across segments. The 360-degree coverage of each panoramic capture allows arbitrary view orientations to be sampled.

Figures 3a and 3b show two example GSV captures used as input to the semantic-segmentation pipeline..



(a)



(b)

Figure 3: Example Google Street View captures used as input to the semantic-segmentation pipeline.

Segmentation was performed with *OneFormer* (Jain et al. 2022), equipped with a DiNAT-L backbone pretrained on Mapillary Vistas v2.0 (47.8 PQ, 64.0 mIoU on the benchmark). The raw Mapillary categories were aggregated into eleven super-classes: road, sidewalk, building, vegetation, sky, street furniture, vehicles, pedestrians, signs, rails, and water. Two examples of segmented images are shown in Figures 4a and 4b.

For each segmented image a feature vector was built comprising (i) the proportion of pixels belonging to each super-class and (ii) the count of detected objects per class. The two images associated with a segment were subsequently averaged to produce a single segment-level visual descriptor vector.

2.2. Methods

Three complementary analytical approaches are applied to the segment-level dataset.

Non-parametric comparison. Segments are split at the median harsh score and distributions of OSM- and GSV-derived features are compared between the high- and low-harshness groups using two-sided Mann–Whitney U tests. For each feature the difference of means and the two-sided p -value are reported. The test is used as a directional



Figure 4: Two examples of GSV captures segmented by OneFormer into eleven super-classes.

screen rather than a formal effect-size estimator; at the per-group sample sizes available here even small distributional differences register as statistically significant, and multiple-comparison correction is not applied across the several dozen features compared (Section 4).

Machine-learning regression. Segment-level harsh scores are predicted from geometric, regulatory, traffic, and visual descriptors using Gradient Boosting and Random Forest regressors, trained on an 80% / 20% train-validation split with hyperparameters selected by five-fold cross-validated grid search. Predictive performance is reported through R^2 , RMSE, MAE, and MAPE on the held-out set. Feature importance is inspected through classical Gini-based rankings and SHapley Additive exPlanations (SHAP) values, the latter providing a consistent attribution of each prediction to its input features while accounting for interactions. The interaction between exposure as a predictor and exposure as a denominator in the harsh rate is discussed in Section 4.

Cycling-infrastructure OLS. The case study of Section 3.4 estimates a log-linear regression of $\log(\text{HarshScore} + 1)$ on $\log(\text{Exposure} + 1)$, posted speed limit, number of lanes, road width, and two facility-type dummies, with HC3 robust standard errors. The facility-type coefficients admit a direct percent-change interpretation relative to the separated-cycle-path baseline.

3. Results

3.1. Spatial and temporal distribution of harsh events

Posted speed limits in Milan cluster around three values – 30, 50, and 70 km/h – which together account for the vast majority of segments. Localised 20 km/h zones exist for pedestrian safety, while 40 km/h limits are reserved for highway ramps and limits above 90 km/h for motorways outside the study area. Figure 5 maps speed limits and TomTom probe counts across the city.

A direct comparison of harsh events across the three primary speed-limit categories (Figure 6) shows that the absolute number of events increases with the speed limit, with the highest counts observed on 70 km/h roads. This is consistent with higher traffic volumes on those arteries. However, the univariate relationship between the harsh score and either the speed limit or traffic volume alone is weak, with near-zero coefficients of determination ($R^2 < 0.001$ and $R^2 = 0.088$; Figure 7). Univariate proxies are therefore insufficient and multivariate analyses are required.

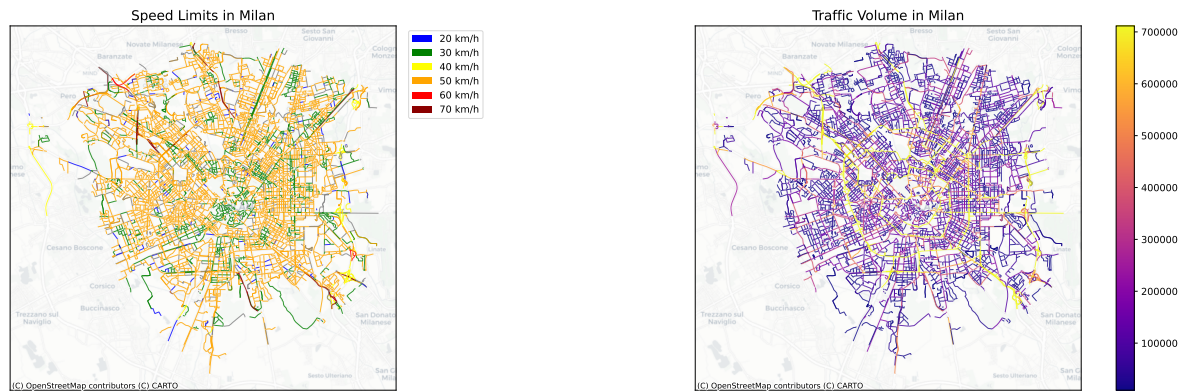


Figure 5: Posted speed limits (left) and TomTom-derived traffic volume proxy (right) across Milan.

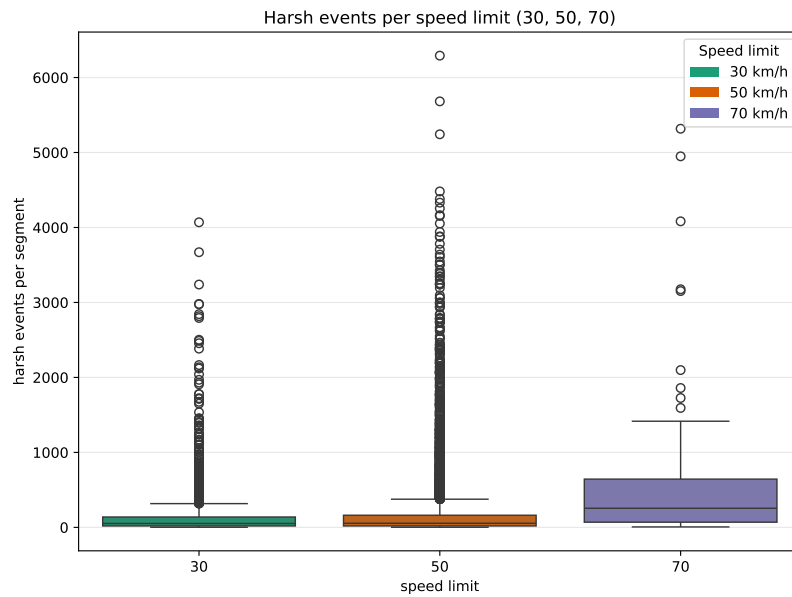


Figure 6: Distribution of harsh scores per segment across the three dominant speed-limit classes.



Figure 7: Univariate relationships between the harsh score and traffic volume (left) and speed limit (right).

Accounting for exposure reveals complementary spatial patterns (Figure 8). The harsh-score map highlights major arterials as the most affected corridors, reflecting high traffic volumes. The harsh-rate map, by contrast, displays more localised hotspots and suggests that certain street geometries and design features drive per-vehicle risk independently of volume.

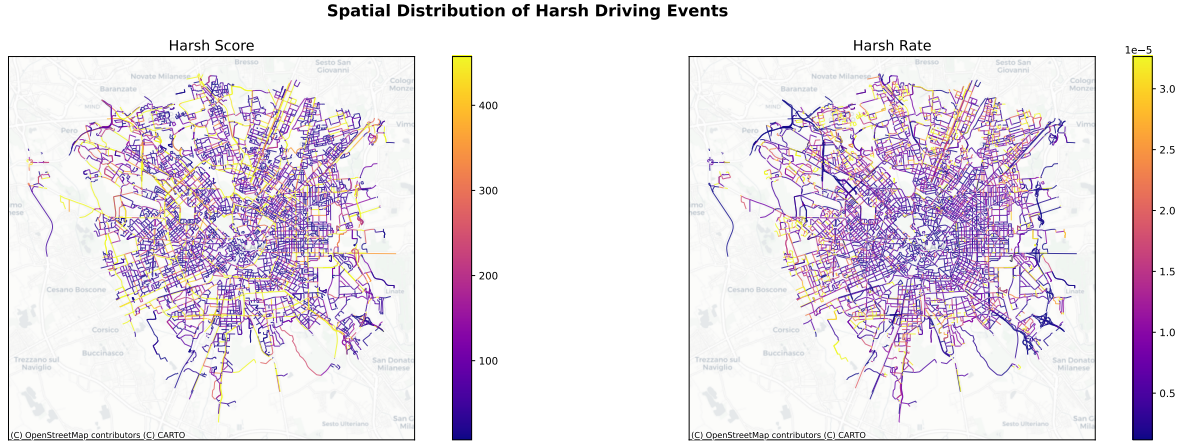


Figure 8: Spatial distribution of the absolute harsh score (left) and the exposure-normalised harsh rate (right) across Milan.

Temporal patterns echo this distinction. The harsh score peaks during morning and evening commute hours, with particularly pronounced spikes on Monday mornings and Friday afternoons (Figure 9). Normalising for exposure reverses the picture: the harsh rate peaks during night-time hours, especially on weekends (Figure 10).

Congestion metrics alone are similarly insufficient. Fitting a naive univariate model on the speed standard deviation or on the Travel Time Index (TTI), defined as

$$TTI = \frac{t_{\text{travel}}}{t_{\text{free-flow}}}, \quad (9)$$

yields $R^2 = 0.066$ and $R^2 = 0.062$ respectively (Figure 11), again motivating a multivariate treatment.

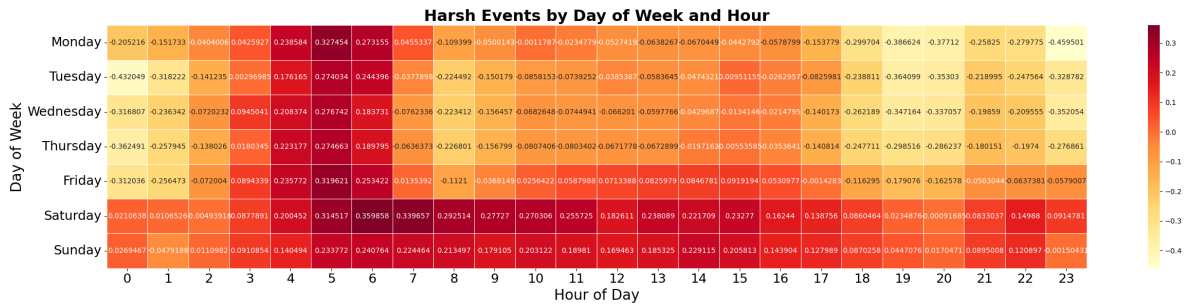


Figure 9: Hourly distribution of the absolute harsh score, by day of the week.

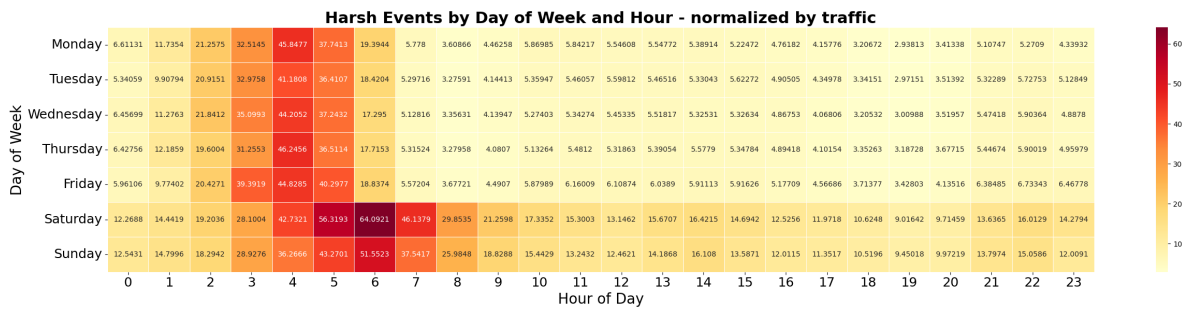


Figure 10: Hourly distribution of the exposure-normalised harsh rate, by day of the week.

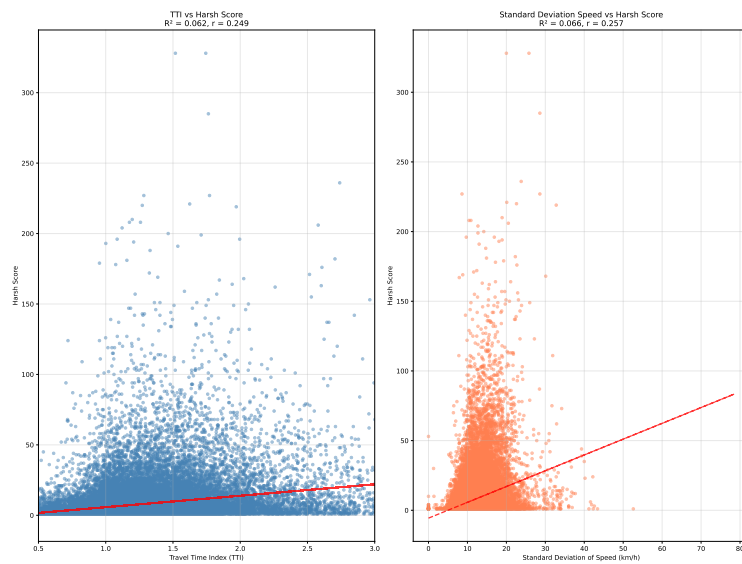


Figure 11: Univariate relationships between the harsh score and the Travel Time Index and speed standard deviation.

3.2. Mann–Whitney U comparisons of segment features

For the 9,267 segments for which GSV imagery was available, segments were split at the median harsh score (logarithm shown in Figure 12) and the distributions of OSM- and GSV-derived features were compared between the high- and low-harshness groups using two-sided Mann–Whitney U tests. For each feature f , the difference of means $d_f = \mu_f^{\text{high}} - \mu_f^{\text{low}}$ is reported together with the two-sided p -value. The analysis is restricted to segments with a 30 or 50 km/h speed limit.

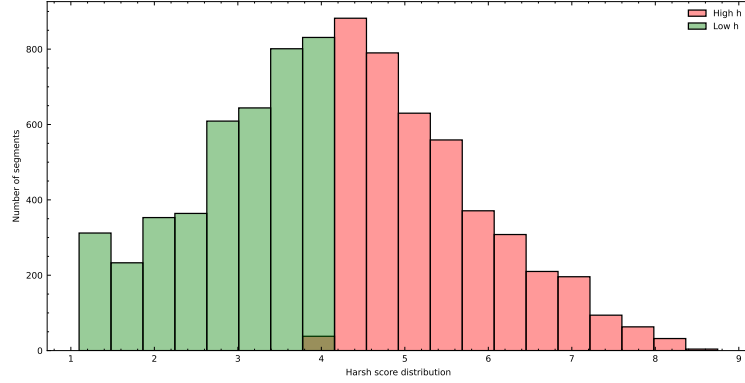


Figure 12: Distribution of the log harsh score across segments.

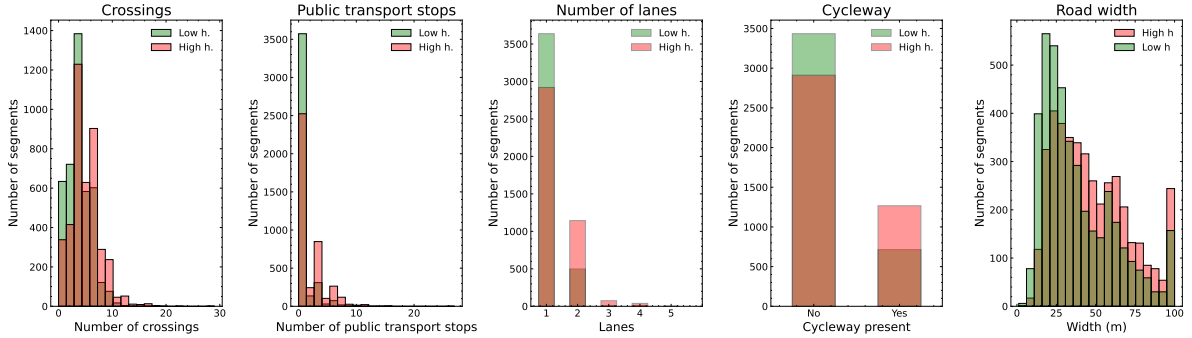


Figure 13: Mann–Whitney U comparisons between high- and low-harshness groups for OSM-derived features.

Road-design and infrastructure attributes differ systematically between the two groups (Figure 13). Segments with pedestrian crossings ($d = +1.20$), public-transport stops ($d = +1.20$), additional lanes ($d = +0.21$), binary cycleway presence ($d = +0.13$), and greater road width ($d = +9.13$ m) all tend to appear more often in the high-harshness group. GSV-derived visual features show the same direction (Figure 14): higher proportions of sky ($d = +0.02$) and road surface ($d = +0.02$) pixels co-occur with higher harsh scores, whereas a higher proportion of building pixels ($d = -0.05$) co-occurs with lower harsh scores.

3.3. Machine-learning models

Two feature configurations were compared: a design-only baseline (Model A) using geometric, regulatory, traffic, and visual descriptors, and a temporal-augmented configuration (Model B) adding the `time_of_day` and `is_week` predictors. Two tree-ensemble algorithms – Gradient Boosting and Random Forest – were trained for each configuration, tuned by grid search with five-fold cross-validation. Gradient Boosting was the best-performing algorithm in the design-only configuration ($R^2 = 0.565$ on the held-out validation set), while Random Forest was the best performer in the temporal-augmented configuration ($R^2 = 0.714$). Table 1 therefore reports the best algorithm per configuration.

Adding temporal predictors reduces RMSE and MAE and raises R^2 from 0.565 to 0.714.

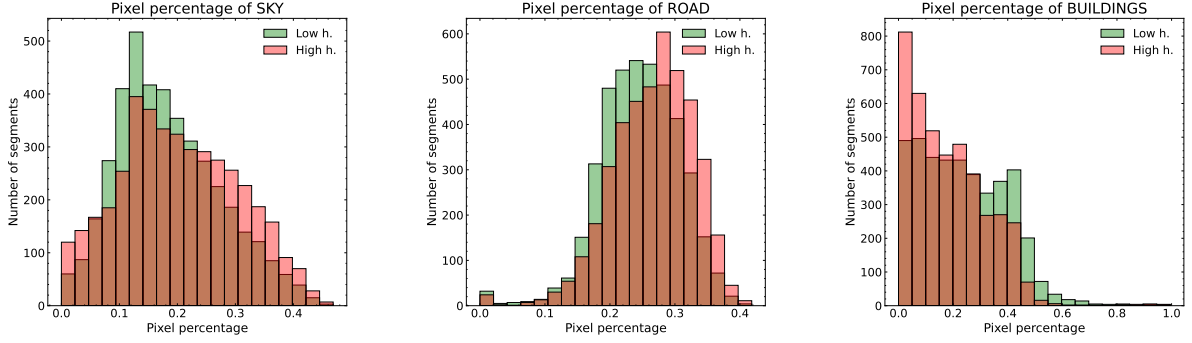


Figure 14: Mann–Whitney U comparisons between high- and low-harshness groups for GSV-derived features.

Table 1: Predictive performance on the held-out validation set. Each row reports the best-performing tree-ensemble algorithm for the corresponding feature configuration.

Model	RMSE	MAE	MAPE (%)	R^2
Design-only (Gradient Boosting)	0.9867	0.7691	24.60	0.565
With temporal features (Random Forest)	0.4763	0.3689	29.80	0.714

SHAP-based feature attribution is shown for both models in Figure 15. In the design-only model, segment-level exposure ranks as the single most influential predictor, followed by the number of traffic signals, estimated street width, and the presence of bus stops and pedestrian crossings. In the temporal-augmented model, temporal indicators join the top predictors without displacing exposure.

3.4. Case study: cycling infrastructure

As a targeted application of the framework, we examined whether different designs of cycling infrastructure are associated with differences in harsh driving behaviour on the adjacent carriageway.

The facility-type dummies admit a direct percent-change interpretation: $\log(Y_{\text{treatment}} / Y_{\text{baseline}}) = \beta$, so the proportional change in the expected harsh score relative to the baseline is $(e^\beta - 1) \times 100\%$.

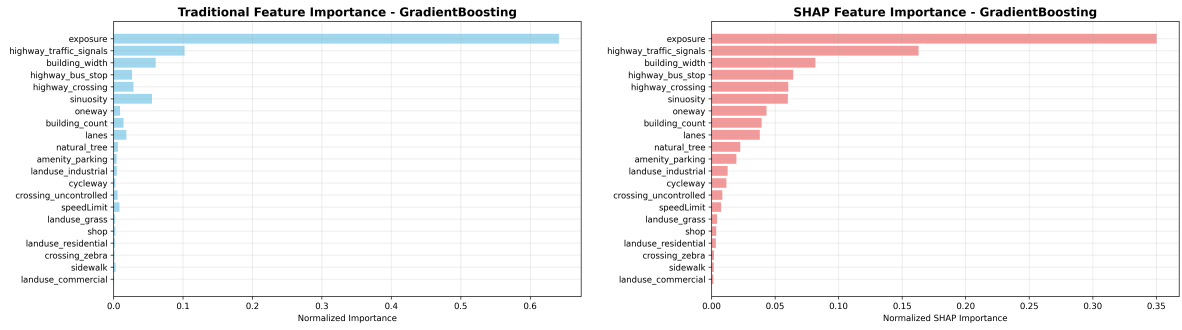
Table 2 reports the fitted coefficients with HC3 robust 95% confidence intervals. The model achieves $R^2 = 0.394$.

Table 2: OLS on $\log(\text{HarshScore} + 1)$ by cycle-infrastructure type, with robust (HC3) 95% confidence intervals.

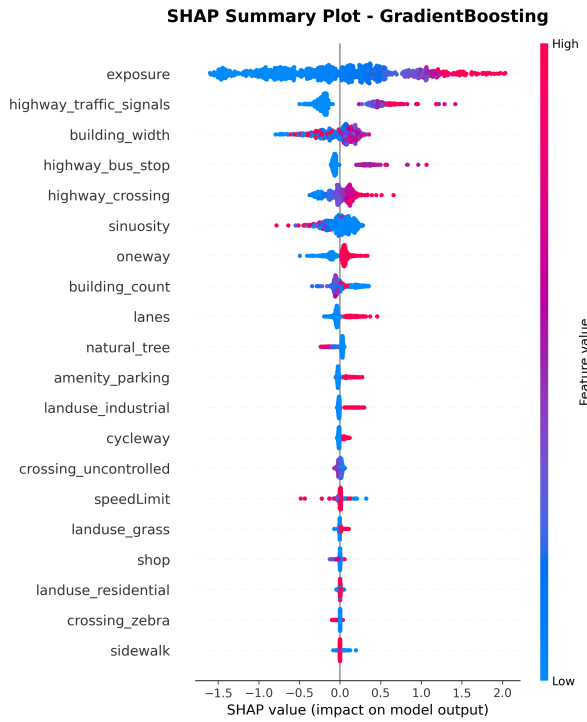
Variable	Coef. (β)	95% CI low	95% CI high
$\log(\text{Exposure} + 1)$	0.584***	0.576	0.592
Markings only (vs. separated)	0.178***	0.156	0.201
Shared with vehicles (vs. separated)	0.109***	0.089	0.130
<i>Percent change in ($y + 1$)</i>			
Markings only	19.5%	16.9%	22.3%
Shared with vehicles	11.5%	9.3%	13.9%
Observations		84,118	
Adjusted R^2		0.394	

*** $p < 0.001$. Facility-type effects are reported as $100(e^\beta - 1)$. The coefficient on $\log(\text{Exposure} + 1)$ is an elasticity of ($y + 1$) with respect to ($\text{Exposure} + 1$).

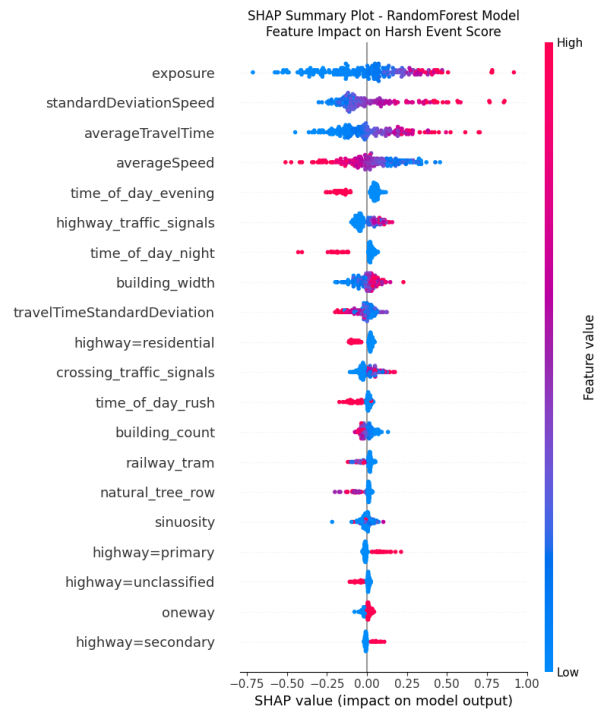
Relative to separated cycle paths, markings-only lanes are associated with a 19.5% higher harsh score ($p < 0.001$) and shared-with-vehicles configurations with an 11.5% higher harsh score ($p < 0.001$). The direction of the gradient is consistent across exposure quintiles and is summarised graphically in Figure 16.



(a) Gini vs. SHAP feature-importance rankings, design-only Gradient Boosting model.



(b) SHAP summary, design-only (Gradient Boosting).



(c) SHAP summary, temporal-augmented (Random Forest).

Figure 15: Feature attribution for the two machine-learning regressors. (a) Gini vs. SHAP rankings for the design-only Gradient Boosting model. (b) SHAP summary for the design-only model. (c) SHAP summary for the temporal-augmented Random Forest model.

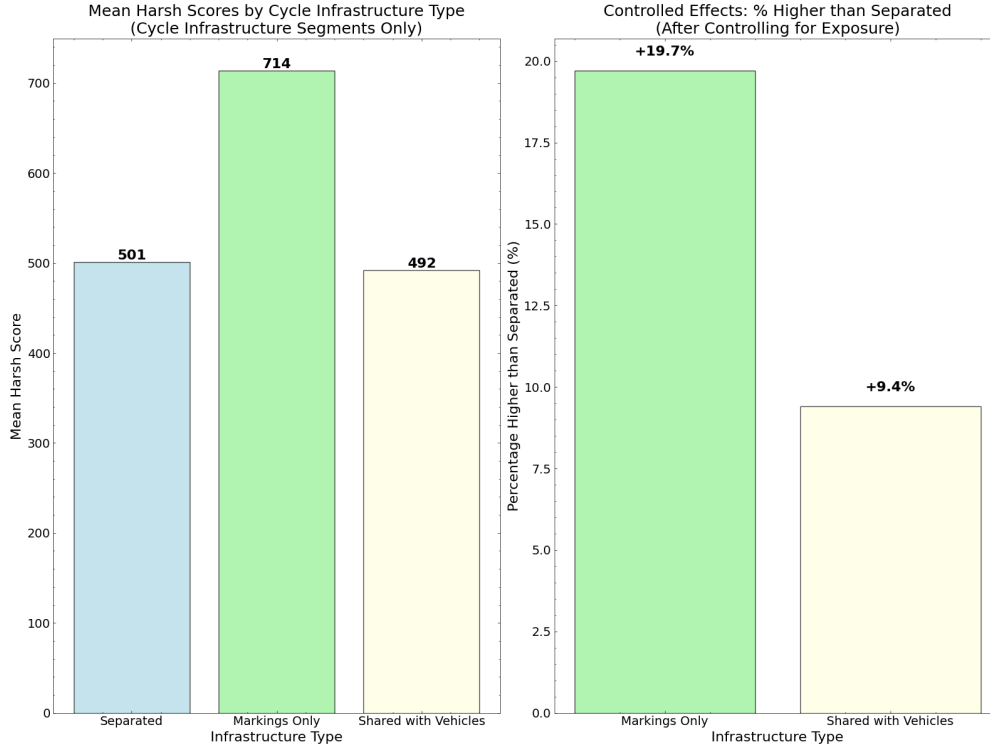


Figure 16: Fitted percent-change in the harsh score by cycling-infrastructure type, relative to separated cycle paths.

4. Discussion and Conclusions

This study presents a large-scale spatial analysis of harsh driving events in the urban road network of Milan, using high-resolution telematics data from more than 4.2 million vehicles together with open geospatial sources (OpenStreetMap, TomTom) and visual streetscape features extracted from Google Street View via semantic segmentation. A multi-scale analytical framework (Mann–Whitney U tests and supervised tree-ensemble regressors with SHAP-based feature attribution) was applied to 24,673 OSM segments in the metropolitan area. The framework extends the spatial harsh-event analysis of Ziakopoulos (2021) to an order of magnitude larger sample, adds visual streetscape descriptors, and introduces an independent exposure control.

The results converge on two main findings. First, exposure sets the absolute count baseline for harsh-event frequency; once exposure is controlled, segment form and visual openness also exert significant effects. Second, temporal normalisation matters: absolute harsh scores peak during commute hours, while per-vehicle risk peaks on weekend nights. Finally the cycling-infrastructure case study illustrates the framework’s practical applicability.

Beyond the quantitative findings, the study contributes a reproducible methodology for rapid construction and analysis of urban road-network spatial datasets that integrate large-scale telematics with open geospatial and visual sources.

4.1. Interpretation of the spatial, temporal, and design patterns

The univariate analyses confirm that no single attribute, speed limit, traffic volume, congestion index, explains the spatial distribution of harsh events on its own. The near-zero coefficients of determination reported in Section 3.1 motivate the multivariate Mann–Whitney and machine-learning analyses, which together indicate that exposure, street form, and visual openness jointly shape the observed spatial pattern. Exposure sets the absolute count baseline for harsh-event frequency; once exposure is controlled, the design and temporal features drive residual variation. This interpretation is consistent with the SHAP rankings, in which exposure is the single most influential predictor but

is accompanied at the top of the list by infrastructure attributes (traffic signals, street width, bus stops, pedestrian crossings) and, in the temporal model, by the time-of-day and weekend indicators.

The temporal pattern is itself informative. Absolute harsh scores peak during morning and evening commute hours on weekdays, consistent with commuting stress and traffic density during these periods. Exposure-normalised rates peak at night, especially on weekends, a pattern consistent with reduced visibility, driver fatigue, and alcohol-related impairment reported in the wider crash literature (Vorko-Jović et al. 2006). The divergence between the absolute and normalised temporal profiles underscores the importance of reporting both quantities: policies driven by absolute counts will target commute hours, whereas policies driven by per-vehicle risk will target weekend nights.

The Mann–Whitney comparisons and SHAP attributions point in the same direction on road design. Wider carriageways, longer sightlines, more open visual fields (higher sky and road-pixel proportions), and public-transport and pedestrian infrastructure co-occur with higher harsh scores, while denser built frontage and narrower cross-sections co-occur with lower harsh scores. These are associations rather than causal effects, but they are consistent with the broader streetscape-perception literature (Fan et al. 2023; Li et al. 2022) and suggest that the openness of the driver’s visual field is a plausible contributor to aggressive acceleration and braking.

4.2. Cycling-infrastructure case study

The cycling-infrastructure comparison recovers an observed gradient in harsh-event intensity across facility types, conditional on the included controls (exposure, posted speed limit, number of lanes, road width). Markings-only lanes are associated with a 19.5% higher harsh score and shared-with-vehicles configurations with an 11.5% higher harsh score, both relative to separated cycle paths. These associations are consistent with the direction suggested by the wider cycling-safety literature but do not on their own establish a causal ranking of facility types. Cycle-path type is highly correlated with road hierarchy, neighbourhood context, and demand patterns: separated paths in Milan tend to be on wider higher-class roads or peripheral greenway corridors, while markings-only and shared facilities appear more frequently on smaller streets in dense districts. The 6-predictor OLS specification reaches $R^2 = 0.394$, which leaves substantial heterogeneity unmodelled.

4.3. Implications for urban-safety interventions

Taken together, the patterns reported above support context-specific rather than uniform safety interventions. Streets combining wide cross-sections, open sightlines, and unobstructed visual fields emerge as candidate locations for lane narrowing, visual or physical traffic-calming elements, and added roadside built frontage. The temporal patterns suggest that time-of-day enforcement or signal retiming may be appropriate on streets that concentrate per-vehicle risk in the late-night or weekend hours. The cycling-infrastructure gradient is consistent with prioritising physical separation where feasible and, where physical separation is impractical, preferring explicit shared arrangements over paint-only delimitation, subject to further causal validation.

4.4. Limitations

The analysis is observational and cannot establish causality. Unmeasured confounders – enforcement intensity, signal coordination, socio-economic context, cyclist volumes, land-use mix – may drive part of the observed patterns. The OSM snapshot is fixed to 31 December 2023, but Google Street View capture dates vary across locations; transient factors such as weather, roadworks, or special events are not modelled. The probe-count proxy for traffic volume, although derived from a large and representative sample, remains a simplified measure of exposure.

The telematics fleet combines two device families with different noise characteristics: Targa devices record accelerometer-derived events, whereas MetaSystem devices derive events from GNSS-measured speed changes over a one-second window. The two distributions were harmonised onto the Targa severity boundaries, but a sensitivity analysis restricted to Targa-only events, or with a device fixed effect in the regressions, was not performed. How much of the observed spatial pattern reflects the spatial distribution of device types across the fleet therefore cannot be quantified with the current analysis.

Finally, for the cycling infrastructure case study, neighbourhood fixed effects and road-class stratification were not applied here; the reported coefficients should therefore be read as associations within this specific sample rather than as a causal ranking of design types, and any infrastructure-policy implication would require further causal investigation.

4.5. Future directions

Four directions for future research follow from the limitations discussed in Section 4.4. First, additional predictors – weather, daylight, detected traffic signs, pedestrian volumes – would improve model interpretability and capture transient effects. Second, a more advanced exposure estimate, for instance the CatBoost-based reconstruction proposed by TomTom (2025) that integrates probe counts with free-flow speeds, lane counts, functional road class, and socio-demographic attributes, would strengthen the denominator of the harsh-rate metric. Third, linking the harsh-event dataset to police-reported crashes – not possible in this study because of data-access constraints – would allow the surrogate to be validated against injury outcomes. Fourth, replicating the analysis in additional European cities would support a direct assessment of transferability. In addition, a formal validation of the map-matching pipeline against a held-out manually-labelled set, a sensitivity analysis to the choice of severity weights and to the device family used to record the events, and the use of multi-scale or hierarchical modelling frameworks that permit non-stationary spatial relationships are all natural follow-on steps that would sharpen several of the associations reported here into more defensible causal claims.

CRedit authorship contribution statement

Andrea La Grotteria: Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review and editing. **Paolo Santi:** Conceptualization, Supervision, Methodology, Writing – review and editing. **Titus Venverloo:** Conceptualization, Methodology, Writing – review and editing. **Umberto Fugiglando:** Conceptualization, Funding acquisition, Writing – review and editing. **Carlo Ratti:** Conceptualization, Funding acquisition, Writing – review and editing.

References

- Araldi, Alessandro and Giovanni Fusco (2019). “From the street to the metropolitan region: Pedestrian perspective in urban fabric analysis”. In: *Environment and Planning B: Urban Analytics and City Science* 46.7, pp. 1243–1263.
- Boeing, Geoff (2025). “Modeling and analyzing urban networks and amenities with OSMnx”. In: *Geographical Analysis*.
- Cabrera-Arnau, Carmen, Rafael Prieto Curiel, and Steven Richard Bishop (2020). “Uncovering the behaviour of road accidents in urban areas”. In: *Royal Society open science* 7.4, p. 191739.
- CINEA (2025). *EU Road Safety: Towards Vision Zero*. URL: https://cinea.ec.europa.eu/news-events/news/eu-road-safety-towards-vision-zero-2025-03-18_en (visited on 08/27/2025).
- Comune di Bologna (2023). *Bologna diventa Città 30*. URL: <https://www.comune.bologna.it/novita/notizie/citta-30> (visited on 08/27/2025).
- Eurostat (2025). *EU road fatalities decreased by 1.3% in 2023*. URL: <https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20250429-1> (visited on 08/27/2025).
- Fan, Zhuangyuan et al. (2023). “Urban visual intelligence: Uncovering hidden city profiles with street view images”. In: *Proceedings of the National Academy of Sciences* 120.27, e2220417120.
- Fleischmann, Martin (2019). “momepy: Urban Morphology Measuring Toolkit”. In: *Journal of Open Source Software* 4.43, p. 1807. DOI: 10.21105/joss.01807. URL: <https://doi.org/10.21105/joss.01807>.
- Greibe, Poul (2003). “Accident prediction models for urban roads”. In: *Accident Analysis & Prevention* 35.2, pp. 273–285.
- Iranmanesh, Aminreza, Can Kara, and Tuğşad Tülbentçi (2024). “Mapping the relationship between traffic accidents, road network configuration, and urban land use”. In: *International Journal of Injury Control and Safety Promotion* 31.4, pp. 672–685.
- Jain, Jitesh et al. (2022). “OneFormer: One Transformer to Rule Universal Image Segmentation. arXiv 2022”. In: *arXiv preprint arXiv:2211.06220*.
- Li, Yongchang et al. (2022). “Street View Imagery (SVI) in the built environment: A theoretical and systematic review”. In: *Buildings* 12.8, p. 1167.
- Ng, KS, WT Hung, and WG Wong (2001). “Effects of urban design on road accidents in Hong Kong”. In: *WIT Transactions on The Built Environment* 52.

- Sadeghi-Bazargani, Homayoun and Mohammad Saadati (2016). "Speed management strategies; a systematic review". In: *Bulletin of Emergency & Trauma* 4.3, p. 126.
- Silva, Iván and Jose Eugenio Naranjo (2020). "A systematic methodology to evaluate prediction models for driving style classification". In: *Sensors* 20.6, p. 1692.
- TomTom (2025). *Historical Traffic Volumes*. Combines probe data and machine learning to estimate accurate traffic volumes network-wide.
- Vorko-Jović, Ariana, Josipa Kern, and Zrinka Biloglav (2006). "Risk factors in urban road traffic accidents". In: *Journal of safety research* 37.1, pp. 93–98.
- Yannis, George and Eva Michelaraki (2024). "Review of city-wide 30 km/h speed limit benefits in Europe". In: *Sustainability* 16.11, p. 4382.
- Yao, Yao et al. (2019). "Exploring the relationship between risk perception, speed limit credibility and speed limit compliance". In: *Transportation research part F: traffic psychology and behaviour* 62, pp. 575–586.
- Ziakopoulos, Apostolos (2021). "Spatial analysis of harsh driving behavior events in urban networks using high-resolution smartphone and geometric data". In: *Accident Analysis & Prevention* 157, p. 106189.
- Ziakopoulos, Apostolos et al. (2022). "Spatial predictions of harsh driving events using statistical and machine learning methods". In: *Safety science* 150, p. 105722.

La dura verità

Analisi a livello di segmento degli eventi di guida brusca a Milano mediante telematica su larga scala, reti stradali e Google Street View

Traduzione italiana non ufficiale del paper MIT / Politecnico di Milano / CNR

Andrea La Grotteria¹, Paolo Santi^{1,3}, Titus Venverloo¹, Umberto Fugiglando¹, Carlo Ratti^{1,2}

¹ Senseable City Lab, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

² Politecnico di Milano, Dipartimento ABC, Via Giuseppe Ponzio 31, 20133 Milano, Italia

³ Consiglio Nazionale delle Ricerche, Via G. Moruzzi 1, 56124 Pisa, Italia

Abstract

Le statistiche sugli incidenti riportati dalla polizia restano la fonte standard per valutare la sicurezza stradale urbana, ma la loro incompletezza e il ritardo con cui vengono rese disponibili ne limitano l'utilità per progettare interventi tempestivi e di dettaglio. Le accelerazioni e frenate brusche sono ampiamente utilizzate come indicatori surrogati di sicurezza, ma finora sono state studiate su campioni urbani relativamente piccoli. Questo studio analizza gli eventi di guida brusca sull'intera rete stradale urbana di Milano, combinando dati telematici ad alta risoluzione provenienti da oltre 4,2 milioni di veicoli dotati di unità di bordo (raccolti tramite UnipolTech in sei settimane rappresentative del 2023), metriche di traffico a livello di segmento da TomTom, attributi della rete stradale e delle infrastrutture da OpenStreetMap e caratteristiche visive dell'ambiente stradale estratte da Google Street View mediante segmentazione semantica con un modello OneFormer. Gli eventi di guida brusca sono stati associati a 24.673 segmenti OSM. Il quadro analitico combina test non parametrici Mann-Whitney U sulle distribuzioni delle caratteristiche dei segmenti tra gruppi ad alta e bassa intensità di eventi, regressori supervisionati di machine learning (Gradient Boosting e Random Forest) con attribuzione delle caratteristiche basata su SHAP e una specificazione OLS controllata per l'esposizione nel caso di studio sulle infrastrutture ciclabili.

I risultati mostrano che, una volta controllata l'esposizione, carreggiate più larghe, attraversamenti, fermate del trasporto pubblico e campi visivi più aperti (percentuali maggiori di pixel cielo e strada) sono associati a più eventi di guida brusca, mentre fronti edificati più densi sono associati a meno eventi. I punteggi assoluti di guida brusca raggiungono il massimo nelle ore di pendolarismo, mentre il rischio per veicolo raggiunge il massimo nelle notti del fine settimana, evidenziando l'importanza della normalizzazione per l'esposizione. Infine, il caso di studio sulle infrastrutture ciclabili rileva un gradiente osservato nell'intensità degli eventi tra le diverse tipologie: le corsie ciclabili delimitate solo da segnaletica sono associate a un punteggio superiore del 19,5% e le configurazioni promiscue con i veicoli a un punteggio superiore dell'11,5% rispetto alle piste ciclabili fisicamente separate, a parità dei controlli inclusi. I risultati sostengono interventi di sicurezza urbana specifici per il contesto, anziché uniformi, e mostrano come telematica su larga scala, dati geospaziali aperti e informazioni visive possano supportare le decisioni di Vision Zero alla scala metropolitana.

1. Introduzione e letteratura rilevante

La sicurezza stradale urbana ha ricevuto una rinnovata attenzione nelle politiche pubbliche, in particolare per quanto riguarda gli utenti vulnerabili come pedoni e ciclisti. La strategia Vision Zero della Commissione europea (CINEA 2025) fissa l'obiettivo di lungo periodo di eliminare morti e feriti gravi dovuti al traffico ed è generalmente attuata attraverso miglioramenti infrastrutturali, attività di sensibilizzazione e riforme delle politiche,

accompagnate da controlli. Diverse città europee, tra cui Stoccolma, Copenaghen, Amsterdam, Zurigo e Barcellona, hanno reso operativo questo quadro con misure quali limiti generalizzati a 30 km/h, restrizioni all'accesso dei veicoli a motore e riprogettazione delle strade (Yannis e Michelaraki 2024). In Italia, Bologna ha recentemente adottato il limite di 30 km/h sulla maggior parte delle strade interne (Comune di Bologna 2023), mentre proposte simili sono oggetto di dibattito a Milano e in altre città. La questione è particolarmente rilevante nel contesto italiano, dove i tassi di mortalità stradale rimangono persistentemente superiori alla media dell'Unione europea (Eurostat 2025) e una quota significativa dei feriti si registra in ambito urbano.

La riduzione dei limiti di velocità urbani è una misura di sicurezza efficace: Yannis e Michelaraki (2024), passando in rassegna le esperienze europee con zone 30 estese a tutta la città, riportano una riduzione associata del 23% degli incidenti con feriti, senza aumento della congestione né dei tempi di viaggio. Ulteriori leve, come il controllo automatizzato, gli elementi verticali di moderazione del traffico, una progettazione che renda credibile il limite di velocità (Sadeghi-Bazargani e Saadati 2016; Yao et al. 2019) e vari fattori contestuali – esposizione temporale, uso del suolo, connettività della rete e controllo delle intersezioni – contribuiscono ulteriormente a determinare il rischio di incidente.

Le valutazioni convenzionali del rischio di incidente si basano quasi esclusivamente sugli incidenti registrati dalla polizia. Tali registrazioni, tuttavia, sono spesso incomplete, non disponibili o pubblicate con mesi di ritardo, cosa che ne limita l'utilità per una pianificazione tempestiva della sicurezza. Da qui nasce un crescente interesse per gli indicatori surrogati di sicurezza: metriche in grado di segnalare il rischio di incidente prima che gli incidenti si verifichino. Le frenate e accelerazioni brusche ("harsh events") sono ampiamente usate come proxy di una guida aggressiva o insicura e possono essere monitorate in continuo attraverso la telematica veicolare. In letteratura, accelerazioni e decelerazioni longitudinali superiori a circa 1,5–2 m/s² vengono normalmente considerate indicatori di manovre rischiose.

La struttura spaziale di questi eventi è stata studiata da Ziakopoulos (2021) usando dati di traiettoria da smartphone in un sobborgo urbano di Atene. Gli autori hanno rilevato autocorrelazione spaziale positiva sia per le frenate brusche (HB) sia per le accelerazioni brusche (HA). Una regressione di Poisson geograficamente pesata ha mostrato che la lunghezza del segmento e l'esposizione al traffico (conteggio dei passaggi) sono i più forti predittori positivi delle frenate brusche, mentre la pendenza e la complessità del quartiere hanno effetti negativi. Uno studio successivo (Ziakopoulos et al. 2022) ha confermato queste tendenze, aggiungendo che le strade a due corsie tendono a produrre più frenate brusche rispetto ai segmenti a corsia singola, mentre le strade residenziali a basso volume ne producono meno. Nel complesso, tali analisi dimostrano che gli eventi di guida brusca riflettono caratteristiche dell'ambiente costruito e possono quindi aiutare a individuare luoghi potenzialmente pericolosi prima che emergano dagli incidenti registrati.

Restano tuttavia tre limiti relativi ai dati. Primo, le analisi precedenti si basavano su meno di 1.500 conducenti di una singola applicazione per smartphone, limitando copertura spaziale e potenza statistica. Secondo, l'ambiente costruito era rappresentato soltanto attraverso la geometria OSM, senza informazioni visive sul paesaggio stradale. Terzo, l'esposizione al traffico era controllata mediante conteggi di passaggi, non attraverso una misura indipendente del volume di traffico. Il presente studio affronta questi tre limiti.

Gli eventi di guida brusca nella città di Milano sono qui analizzati mediante registrazioni telematiche ad alta risoluzione provenienti da oltre 4,2 milioni di veicoli dotati di On-Board Unit, fornite da UnipolTech, la divisione telematica di uno dei maggiori gruppi assicurativi italiani. I dati sono integrati con metriche di traffico a livello di segmento da TomTom, geometria della rete e attributi infrastrutturali da OpenStreetMap (snapshot al 31 dicembre 2023) e caratteristiche visive dell'ambiente stradale estratte da immagini Google Street View mediante un modello di segmentazione semantica allo stato dell'arte. Il dataset combinato copre 24.673 segmenti stradali urbani e 3.407.876 eventi di guida brusca associati alla rete, registrati in sei settimane rappresentative del 2023.

L'analisi combina test statistici non parametrici, regressori supervisionati di machine learning con attribuzione delle caratteristiche basata su SHAP e una specificazione OLS controllata per l'esposizione per il caso di studio sulle infrastrutture ciclabili. Un caso di studio specifico esamina in che modo diverse configurazioni delle infrastrutture ciclabili siano associate all'intensità degli eventi di guida brusca sui segmenti stradali adiacenti.

L'analisi è organizzata attorno a tre domande di ricerca:

1. RQ1. Le misure di eventi di guida brusca basate su telematica su larga scala riproducono le associazioni a livello di segmento già osservate in dataset più piccoli basati su smartphone? E l'aggiunta di descrittori visivi dell'ambiente stradale ne migliora la capacità esplicativa?
2. RQ2. Dopo aver controllato l'esposizione al traffico, come si relazionano all'intensità degli eventi di guida brusca la geometria stradale, gli attributi regolamentari e l'apertura visiva?
3. RQ3. Possiamo utilizzare gli eventi di guida brusca per orientare decisioni di policy specifiche, ad esempio sulle infrastrutture ciclabili?

Il resto dell'articolo è organizzato come segue. La sezione 2.2 illustra il quadro metodologico dei modelli statistici e di machine learning impiegati. La sezione 2.1 descrive le fonti dati, mentre nell'originale alcune successive sezioni sono indicate con “??” per la pipeline di elaborazione e integrazione e per il riepilogo del dataset spaziale finale. La sezione 3 presenta i risultati delle analisi statistiche e di machine learning, insieme al caso di studio sulle infrastrutture ciclabili. La sezione 4 discute implicazioni e limiti e la conclusione indica le direzioni per ricerche future.

2. Dati e metodi

2.1. Fonti dei dati

L'analisi integra cinque dataset complementari: eventi di guida brusca basati sulla telematica UnipolTech, metriche di traffico a livello di segmento TomTom, attributi della rete stradale e delle infrastrutture da OpenStreetMap, rete ciclabile comunale di Milano e immagini Google Street View. Di seguito sono descritte ciascuna fonte e le relative variabili.

2.1.1. OpenStreetMap: rete stradale e infrastrutture

La topologia della rete stradale e gli attributi infrastrutturali sono stati ricavati da OpenStreetMap (OSM) per il Comune di Milano mediante la libreria osmnx (Boeing 2025). Lo snapshot è quello del 31 dicembre 2023 e comprende 18.037 nodi e 35.606 archi direzionati, per una lunghezza totale di circa 3.640 km e una lunghezza media dei segmenti di 102 m. Per ogni segmento il dataset fornisce geometria, numero di corsie, classificazione della strada (per esempio residenziale o terziaria), limiti di velocità indicati quando disponibili, direzione di marcia e presenza di marciapiedi e infrastrutture ciclabili.

Poiché l'analisi riguarda la guida urbana, sono stati esclusi 605 segmenti autostradali e i segmenti privi di qualunque osservazione UnipolTech associata. La rete OSM finale comprende 24.673 segmenti.

Per ciascun segmento OSM sono stati derivati tre gruppi di caratteristiche. Primo, caratteristiche geometriche: lunghezza del segmento e sinuosità, definita come rapporto tra lunghezza effettiva del segmento e distanza euclidea fra nodo di origine e nodo di destinazione:

$$\text{sinuosità} = \text{lunghezza del segmento} / \|(x_d, y_d) - (x_o, y_o)\|$$

Secondo, le caratteristiche del contesto stradale conteggiano gli elementi infrastrutturali entro un buffer di 20 m attorno a ciascun segmento: semafori, attraversamenti pedonali, fermate del trasporto pubblico, infrastrutture ciclabili, marciapiedi, aree gioco e aree verdi. Terzo, le caratteristiche di forma e regolazione della strada

comprendono il tag highway raggruppato, il numero di corsie e una larghezza stradale stimata mediante il metodo momepy.StreetProfile (Fleischmann 2019), seguendo Araldi e Fusco (2019).

Figura 1. Distribuzione della larghezza stradale stimata nella città di Milano mediante il metodo momepy.StreetProfile.

2.1.2. UnipolTech: eventi telematici di guida brusca

La fonte primaria dei dati sul comportamento di guida è UnipolTech, divisione telematica di Unipol, uno dei maggiori gruppi assicurativi italiani. I dati sono raccolti da una flotta nazionale di oltre 4,2 milioni di veicoli dotati di On-Board Unit (OBU), che registrano frenate improvvise e accelerazioni rapide. Ogni evento contiene timestamp, latitudine e longitudine, velocità istantanea, direzione del veicolo (in gradi rispetto al Nord), tipo di evento (frenata o accelerazione), misura della severità, indicatore di qualità del segnale GNSS, tipo di dispositivo e categoria del veicolo.

Nella flotta sono utilizzati due modelli di OBU:

- Dispositivi Targa Telematics: misurano direttamente l'accelerazione tramite accelerometro di bordo e assegnano ogni evento a una di quattro classi di severità, con soglie (in cm/s^2) di 350–450, 450–550, 550–700 e oltre 700.
- Dispositivi MetaSystem: stimano l'accelerazione dalle variazioni di velocità ricavate dal GNSS su una finestra di un secondo.

Le due famiglie di dispositivi sono armonizzate sulle soglie di classe Targa, ottenendo una scala unificata di severità da 1 a 4.

Benché il dataset grezzo copra tutta Italia, lo studio si concentra su Milano per tre ragioni. Primo, la città offre una copertura spaziale coerente per tutte le fonti ausiliarie (OSM, GSV, TomTom). Secondo, Milano sta discutendo attivamente misure orientate a Vision Zero, come l'estensione delle zone 30, rendendo l'analisi direttamente rilevante per le politiche. Terzo, limitare l'analisi a una singola area metropolitana mantiene gestibile il carico computazionale pur producendo diversi milioni di eventi di guida.

Per bilanciare rappresentatività stagionale e volume di dati, sono stati mantenuti gli eventi relativi a sei settimane rappresentative del 2023: la prima settimana di gennaio, aprile, luglio, ottobre, novembre e dicembre.

Ogni evento di guida brusca è stato associato a un segmento OSM con una procedura in due fasi. Prima, ogni segmento OSM è stato campionato a intervalli di 3 m; per ciascun punto campionato sono state calcolate le coordinate e la direzione tangente locale. Per ogni evento sono stati individuati i 20 punti campionati geograficamente più vicini e si è selezionato il candidato con la minima differenza assoluta tra la direzione locale e la direzione osservata del veicolo. L'associazione è stata mantenuta soltanto se la distanza spaziale tra evento e punto selezionato era inferiore a 6 m e la differenza di direzione inferiore a 45° . Le soglie sono state calibrate su tessuti urbani densi per ridurre al minimo errori tra carreggiate parallele e strade di servizio.

Una volta associati gli eventi ai segmenti OSM, i conteggi sono stati aggregati a livello di segmento. Per controllare l'esposizione, è stata calcolata la seguente metrica:

$$\text{Esposizione}_s = \text{VolumeTraffico}_s \times \text{Lunghezza}_s$$

dove VolumeTraffico_s è il conteggio dei probe TomTom del segmento s e Lunghezza_s è la lunghezza del segmento OSM. Poiché la distribuzione grezza dell'esposizione è fortemente asimmetrica a destra, è stata applicata una trasformazione logaritmica:

$$\text{LogEsposizione}_s = \log(1 + \text{Esposizione}_s)$$

Il conteggio pesato degli eventi sul segmento s è definito come:

$$\text{PunteggioBrusco}_s = n_1 + 2n_2 + 3n_3 + 4n_4$$

dove n_k è il numero di eventi di severità k . Viene inoltre definito un tasso normalizzato e log-trasformato:

$$\text{TassoBrusco}_s = \log(1 + \text{PunteggioBrusco}_s / \text{Esposizione}_s)$$

Il punteggio di guida brusca cattura l'intensità assoluta della guida rischiosa su un segmento, mentre il tasso di guida brusca ne cattura l'intensità per unità di esposizione. Entrambe le grandezze sono riportate nei risultati.

2.1.3. TomTom: metriche di traffico a livello di segmento

Per controllare l'esposizione al traffico sono stati utilizzati dati TomTom a livello di segmento. TomTom aggrega dati probe provenienti da sistemi di navigazione a bordo veicolo, applicazioni smartphone e piattaforme cartografiche di terze parti, fra cui Apple Maps. Per ogni segmento proprietario il dataset fornisce velocità media e mediana, limite di velocità, classe e nome della strada, percentili di velocità P5 e P95, deviazione standard della velocità e numerosità del campione di probe. Quest'ultima è utilizzata come proxy del volume di traffico: è una semplificazione deliberata, ma sostenuta dalla copertura TomTom, che secondo il fornitore intercetta segnali provenienti da una quota sostanziale della flotta veicolare.

TomTom usa una segmentazione propria, con lunghezza media di 56,1 m, più fine di quella OSM, pari a 102,3 m. L'associazione tra le due reti è stata eseguita in un sistema di riferimento metrico proiettato (EPSG:32632) usando un indice spaziale STR-tree. Per ogni segmento TomTom sono stati recuperati i 20 candidati OSM più vicini; sono state calcolate la distanza di Hausdorff e la differenza assoluta di direzione e l'associazione è stata accettata se la distanza minima era ≤ 10 m e la differenza di direzione $\leq 45^\circ$. Con queste soglie, tutti i 71.867 segmenti TomTom sono stati associati.

Per ciascuna metrica TomTom (velocità media, percentili, deviazione standard della velocità, numerosità del campione), il valore aggregato per il segmento OSM è stato calcolato come media pesata per la lunghezza. I segmenti TomTom più lunghi, che coprono una quota maggiore dell'arco OSM, contribuiscono quindi proporzionalmente di più alla stima aggregata.

2.1.4. Comune di Milano: rete ciclabile comunale

I dati sulle infrastrutture ciclabili sono stati ricavati dalla rete ciclabile comunale di Milano tramite il portale open data della città. Ogni record fornisce geometria e tipologia dell'infrastruttura – piste ciclabili fisicamente separate, corsie ciclabili su strada delimitate solo da segnaletica, segmenti promiscui con veicoli, zone a basso traffico – oltre a lunghezza, anno di realizzazione e classificazione regolamentare. I dati sono stati ripuliti eliminando gli attraversamenti stradali e le infrastrutture fuori dalla finestra temporale dello studio, mantenendo soltanto quelle completate entro il 2023. Ogni geometria ciclabile è stata proiettata in EPSG:32632 e suddivisa in frammenti contigui di 3 m, eliminando quelli più corti di 10 cm. Ciascun frammento è stato poi associato al corrispondente segmento OSM.

Sono confrontate tre tipologie di infrastruttura:

4. Pista ciclabile separata – fisicamente segregata dalle corsie dei veicoli a motore mediante cordoli, barriere o separazione altimetrica (categorie “ciclabile sede propria” e “promiscuo pedoni”).
5. Corsia ciclabile delimitata solo da segnaletica – corsia visivamente delimitata sulla carreggiata, senza separazione fisica (categoria “ciclabile segnaletica”).
6. Promiscua con i veicoli – strade a traffico misto designate per ciclisti e veicoli a motore (categoria “promiscuo veicoli”).

Dei 92.058 frammenti di infrastruttura ciclabile da tre metri, il 52% appartiene alla categoria “solo segnaletica”, il 25,8% è separato e il 21,8% è promiscuo con i veicoli. Dopo avere eliminato i frammenti privi di dati associati sull’esposizione o sugli eventi di guida brusca, nell’analisi sono entrati 84.118 frammenti.

Figura 2. Infrastrutture ciclabili di Milano per tipologia: piste separate, corsie delimitate solo da segnaletica e segmenti promiscui con i veicoli.

È stata stimata una regressione log-lineare con $\log(\text{PunteggioBrusco} + 1)$ come variabile dipendente e con $\log(\text{Esposizione} + 1)$, limite di velocità, numero di corsie, larghezza stradale e due dummy per la tipologia di infrastruttura (solo segnaletica, promiscua) come predittori. Le piste separate costituiscono la categoria di riferimento:

$$\log(\text{PunteggioBrusco} + 1) = \beta_0 + \beta_1 \log(\text{Esposizione} + 1) + \beta_2 \text{LimiteVelocità} + \beta_3 \text{Corsie} + \beta_4 \text{LarghezzaStrada} + \beta_5 \text{SoloSegnaletica} + \beta_6 \text{Promiscua} + \varepsilon$$

Questi dati sono utilizzati per studiare la relazione tra progettazione delle infrastrutture ciclabili ed eventi di guida brusca sulla carreggiata adiacente.

2.1.5. Google Street View: contesto visivo della strada

I descrittori visivi dell’ambiente stradale sono stati estratti da Google Street View (GSV). Per ciascuno dei 24.673 segmenti OSM mantenuti sono state scaricate tramite API GSV due immagini da 640×640 pixel, una nella direzione di marcia e una in direzione opposta. GSV è stato preferito ad alternative crowdsourced come Mapillary perché fornisce una pipeline di acquisizione standardizzata e ad alta risoluzione, essenziale per confronti pixel per pixel tra segmenti. La copertura a 360° di ciascuna panoramica consente di campionare orientamenti arbitrari della vista.

Figura 3. Esempi di immagini Google Street View usate come input della pipeline di segmentazione semantica.

La segmentazione è stata eseguita con OneFormer (Jain et al. 2022), dotato di backbone DiNAT-L preaddestrato su Mapillary Vistas v2.0. Le categorie grezze di Mapillary sono state aggregate in undici super-classi: strada, marciapiede, edificio, vegetazione, cielo, arredo stradale, veicoli, pedoni, segnali, rotaie e acqua.

Figura 4. Due esempi di immagini GSV segmentate da OneFormer nelle undici super-classi.

Per ciascuna immagine segmentata è stato costruito un vettore di caratteristiche comprendente: (i) la proporzione di pixel appartenenti a ogni super-classe e (ii) il numero di oggetti rilevati per classe. Le due immagini associate a ciascun segmento sono state quindi mediate per ottenere un unico vettore di descrittori visivi a livello di segmento.

2.2. Metodi

Al dataset a livello di segmento vengono applicati tre approcci analitici complementari.

Confronto non parametrico. I segmenti sono divisi alla mediana del punteggio di guida brusca e le distribuzioni delle caratteristiche derivate da OSM e GSV vengono confrontate tra gruppi ad alta e bassa intensità tramite test Mann–Whitney U a due code. Per ogni caratteristica vengono riportate la differenza delle medie e il p-value a due code. Il test è utilizzato come screening direzionale e non come stima formale della dimensione dell’effetto; con campioni così grandi anche piccole differenze risultano statisticamente significative e non viene applicata una correzione per confronti multipli alle diverse decine di caratteristiche considerate.

Regressione di machine learning. I punteggi a livello di segmento vengono predetti a partire da descrittori geometrici, regolamentari, di traffico e visivi usando regressori Gradient Boosting e Random Forest, addestrati con suddivisione 80%/20% fra training e validazione e con iperparametri selezionati mediante grid search con cross-validation a cinque fold. Le prestazioni vengono riportate tramite R^2 , RMSE, MAE e MAPE sul set di validazione.

L'importanza delle caratteristiche viene ispezionata sia con ranking classici basati su Gini sia con valori SHAP (SHapley Additive exPlanations), che attribuiscono in modo coerente a ciascuna variabile il contributo alla previsione, tenendo conto delle interazioni.

OLS sulle infrastrutture ciclabili. Il caso di studio stima una regressione log-lineare di $\log(\text{PunteggioBrusco} + 1)$ su $\log(\text{Esposizione} + 1)$, limite di velocità, numero di corsie, larghezza stradale e due dummy di tipologia, con errori standard robusti HC3. I coefficienti delle tipologie possono essere interpretati direttamente come variazioni percentuali rispetto alla categoria di riferimento delle piste ciclabili separate.

3. Risultati

3.1. Distribuzione spaziale e temporale degli eventi di guida brusca

I limiti di velocità indicati a Milano si concentrano attorno a tre valori – 30, 50 e 70 km/h – che rappresentano la grande maggioranza dei segmenti. Esistono zone localizzate a 20 km/h per la sicurezza pedonale, mentre i 40 km/h sono riservati alle rampe autostradali e i limiti superiori a 90 km/h alle autostrade, esterne all'area di studio.

Figura 5. Limiti di velocità indicati (sinistra) e proxy del volume di traffico derivato da TomTom (destra) a Milano.

Un confronto diretto degli eventi bruschi fra le tre principali categorie di limite mostra che il numero assoluto di eventi aumenta con il limite di velocità, con i conteggi maggiori sulle strade a 70 km/h. Ciò è coerente con i maggiori volumi di traffico di queste arterie. Tuttavia, la relazione univariata tra il punteggio di guida brusca e il solo limite di velocità o il solo volume di traffico è debole, con coefficienti di determinazione prossimi a zero ($R^2 < 0,001$ e $R^2 = 0,088$). Le proxy univariate sono quindi insufficienti e servono analisi multivariate.

Figura 6. Distribuzione dei punteggi di guida brusca per segmento nelle tre classi dominanti di limite di velocità.

Figura 7. Relazioni univariate tra punteggio di guida brusca e volume di traffico (sinistra) e limite di velocità (destra).

Considerare l'esposizione rivela pattern spaziali complementari. La mappa del punteggio di guida brusca evidenzia le principali arterie come corridoi più colpiti, riflettendo gli elevati volumi di traffico. La mappa del tasso normalizzato, invece, mostra hotspot più localizzati e suggerisce che alcune geometrie stradali e caratteristiche progettuali determinino un rischio per veicolo indipendente dal volume.

Figura 8. Distribuzione spaziale del punteggio assoluto di guida brusca (sinistra) e del tasso normalizzato per l'esposizione (destra) a Milano.

I pattern temporali riflettono la stessa distinzione. Il punteggio assoluto raggiunge il massimo nelle ore di punta del mattino e della sera, con picchi particolarmente marcati il lunedì mattina e il venerdì pomeriggio. Normalizzando per l'esposizione, il quadro si rovescia: il tasso di guida brusca raggiunge il massimo nelle ore notturne, soprattutto nei fine settimana.

Figura 9. Distribuzione oraria del punteggio assoluto di guida brusca per giorno della settimana.

Figura 10. Distribuzione oraria del tasso normalizzato per l'esposizione per giorno della settimana.

Anche le metriche di congestione, da sole, sono insufficienti. Un modello univariato ingenuo basato sulla deviazione standard della velocità o sul Travel Time Index (TTI = tempo di viaggio / tempo a flusso libero) produce rispettivamente $R^2 = 0,066$ e $R^2 = 0,062$, rafforzando la necessità di un approccio multivariato.

Figura 11. Relazioni univariate tra punteggio di guida brusca e Travel Time Index e deviazione standard della velocità.

3.2. Confronti Mann–Whitney U delle caratteristiche dei segmenti

Per i 9.267 segmenti per cui erano disponibili immagini GSV, i segmenti sono stati divisi alla mediana del punteggio di guida brusca e le distribuzioni delle caratteristiche derivate da OSM e GSV sono state confrontate tra i gruppi ad alta e bassa intensità mediante test Mann–Whitney U a due code. L'analisi è limitata ai segmenti con limite di 30 o 50 km/h.

Figura 12. Distribuzione del logaritmo del punteggio di guida brusca fra i segmenti.

Gli attributi di progettazione stradale e infrastruttura differiscono sistematicamente tra i due gruppi. I segmenti con attraversamenti pedonali ($d = +1,20$), fermate del trasporto pubblico ($d = +1,20$), corsie aggiuntive ($d = +0,21$), presenza binaria di infrastruttura ciclabile ($d = +0,13$) e maggiore larghezza stradale ($d = +9,13$ m) compaiono più spesso nel gruppo ad alta intensità. Le caratteristiche visive derivate da GSV mostrano la stessa direzione: percentuali più elevate di pixel cielo ($d = +0,02$) e superficie stradale ($d = +0,02$) co-occorrono con punteggi più elevati, mentre una maggiore percentuale di pixel edificio ($d = -0,05$) co-occorre con punteggi più bassi.

Figura 13. Confronti Mann–Whitney U tra gruppi ad alta e bassa intensità per caratteristiche derivate da OSM.

Figura 14. Confronti Mann–Whitney U tra gruppi ad alta e bassa intensità per caratteristiche visive derivate da GSV.

3.3. Modelli di machine learning

Sono state confrontate due configurazioni di caratteristiche: un modello di base “solo design” (Modello A), con descrittori geometrici, regolamentari, di traffico e visivi, e una configurazione “con variabili temporali” (Modello B), che aggiunge i predittori `time_of_day` e `is_week`. Per ciascuna configurazione sono stati addestrati due algoritmi ensemble ad alberi – Gradient Boosting e Random Forest – ottimizzati tramite grid search con cross-validation a cinque fold. Gradient Boosting è risultato il migliore nella configurazione solo design ($R^2 = 0,565$ sul set di validazione), mentre Random Forest è risultato il migliore nella configurazione con variabili temporali ($R^2 = 0,714$). L'aggiunta delle variabili temporali riduce RMSE e MAE e aumenta R^2 da 0,565 a 0,714.

Tabella 1. Prestazioni predittive sul set di validazione.

Modello	RMSE	MAE	MAPE (%)	R^2
Solo design (Gradient Boosting)	0,9867	0,7691	24,60	0,565
Con variabili temporali (Random Forest)	0,4763	0,3689	29,80	0,714

L'attribuzione delle caratteristiche basata su SHAP mostra, nel modello solo design, che l'esposizione a livello di segmento è il singolo predittore più influente, seguita dal numero di semafori, dalla larghezza stradale stimata e dalla presenza di fermate bus e attraversamenti pedonali. Nel modello con variabili temporali, gli indicatori temporali entrano tra i principali predittori senza scalzare l'esposizione.

Figura 15. Attribuzione dell'importanza delle caratteristiche per i due regressori di machine learning: ranking Gini e SHAP per il modello Gradient Boosting solo design e sintesi SHAP per i modelli solo design e con variabili temporali.

3.4. Caso di studio: infrastrutture ciclabili

Come applicazione mirata del quadro analitico, è stato esaminato se differenti configurazioni delle infrastrutture ciclabili siano associate a differenze nel comportamento di guida brusca sulla carreggiata adiacente.

Le dummy di tipologia ammettono un'interpretazione diretta in termini percentuali: se $\log(Y_{\text{trattamento}} / Y_{\text{riferimento}}) = \beta$, la variazione proporzionale del punteggio atteso rispetto alla categoria di riferimento è $(e^\beta - 1) \times 100\%$. Il modello raggiunge $R^2 = 0,394$.

Tabella 2. OLS su $\log(\text{PunteggioBrusco} + 1)$ per tipologia di infrastruttura ciclabile, con intervalli di confidenza robusti (HC3) al 95%.

Variabile	Coeff. (β)	IC 95% basso	IC 95% alto
$\log(\text{Esposizione} + 1)$	0,584***	0,576	0,592
Solo segnaletica (vs separata)	0,178***	0,156	0,201
Promiscua con veicoli (vs separata)	0,109***	0,089	0,130
Variazione % di (y + 1): solo segnaletica	19,5%	16,9%	22,3%
Variazione % di (y + 1): promiscua con veicoli	11,5%	9,3%	13,9%
Osservazioni	84.118		
R ² aggiustato	0,394		

*** $p < 0,001$. Gli effetti delle tipologie sono riportati come $100(e^{\beta} - 1)$. Il coefficiente di $\log(\text{Esposizione} + 1)$ è un'elasticità di (y + 1) rispetto a (Esposizione + 1).

Rispetto alle piste ciclabili separate, le corsie delimitate solo da segnaletica sono associate a un punteggio di guida brusca superiore del 19,5% ($p < 0,001$) e le configurazioni promiscue con veicoli a un punteggio superiore dell'11,5% ($p < 0,001$). La direzione del gradiente è coerente in tutti i quintili di esposizione.

Figura 16. Variazione percentuale stimata del punteggio di guida brusca per tipologia di infrastruttura ciclabile, rispetto alle piste separate.

4. Discussione e conclusioni

Questo studio presenta un'analisi spaziale su larga scala degli eventi di guida brusca nella rete stradale urbana di Milano, usando dati telematici ad alta risoluzione provenienti da oltre 4,2 milioni di veicoli insieme a fonti geospaziali aperte (OpenStreetMap, TomTom) e caratteristiche visive dell'ambiente stradale estratte da Google Street View mediante segmentazione semantica. Un quadro analitico multiscala – test Mann–Whitney U e regressori supervisionati ensemble ad alberi con attribuzione SHAP – è stato applicato a 24.673 segmenti OSM dell'area metropolitana. Il metodo estende l'analisi spaziale degli eventi di Ziakopoulos (2021) a un campione di un ordine di grandezza maggiore, aggiunge descrittori visivi del paesaggio stradale e introduce un controllo indipendente dell'esposizione.

I risultati convergono su due conclusioni principali. Primo, l'esposizione determina la base del conteggio assoluto degli eventi; una volta controllata, anche la forma del segmento e l'apertura visiva mostrano effetti significativi. Secondo, la normalizzazione temporale conta: i punteggi assoluti raggiungono il massimo nelle ore di pendolarismo, mentre il rischio per veicolo raggiunge il massimo nelle notti del fine settimana. Infine, il caso di studio sulle infrastrutture ciclabili illustra l'applicabilità pratica del quadro analitico.

Oltre ai risultati quantitativi, lo studio contribuisce con una metodologia riproducibile per costruire e analizzare rapidamente dataset spaziali di reti stradali urbane che integrano telematica su larga scala, dati geospaziali aperti e fonti visive.

4.1. Interpretazione dei pattern spaziali, temporali e progettuali

Le analisi univariate confermano che nessun singolo attributo – limite di velocità, volume di traffico, indice di congestione – spiega da solo la distribuzione spaziale degli eventi di guida brusca. I coefficienti di determinazione prossimi a zero motivano le analisi multivariate Mann–Whitney e di machine learning, che indicano congiuntamente come esposizione, forma stradale e apertura visiva determinino il pattern osservato. L'esposizione stabilisce la base del conteggio assoluto; una volta controllata, caratteristiche progettuali e temporali guidano la variazione residua. Questa interpretazione è coerente con i ranking SHAP, in cui l'esposizione è il predittore

singolo più influente ma ai primi posti compaiono anche attributi infrastrutturali quali semafori, larghezza stradale, fermate bus e attraversamenti pedonali e, nel modello temporale, ora del giorno e indicatore di weekend.

Il pattern temporale è di per sé informativo. I punteggi assoluti raggiungono il massimo nelle ore di punta mattutine e serali dei giorni feriali, coerentemente con stress da pendolarismo e densità di traffico. I tassi normalizzati per l'esposizione raggiungono invece il massimo di notte, soprattutto nei fine settimana, un pattern coerente con la ridotta visibilità, la stanchezza del conducente e l'alterazione correlata all'alcol riportate nella più ampia letteratura sugli incidenti. La divergenza tra profili assoluti e normalizzati sottolinea l'importanza di riportare entrambe le grandezze: politiche guidate dai conteggi assoluti punterebbero sulle ore di punta, mentre politiche guidate dal rischio per veicolo punterebbero sulle notti del weekend.

I confronti Mann–Whitney e le attribuzioni SHAP indicano la stessa direzione per quanto riguarda il design stradale. Carreggiate più larghe, visuali più lunghe, campi visivi più aperti (percentuali maggiori di cielo e strada) e infrastrutture pedonali e di trasporto pubblico co-occorrono con punteggi più elevati, mentre fronti edificati più densi e sezioni trasversali più strette co-occorrono con punteggi inferiori. Si tratta di associazioni, non di effetti causali, ma sono coerenti con la più ampia letteratura sulla percezione dell'ambiente stradale e suggeriscono che l'apertura del campo visivo del conducente possa contribuire ad accelerazioni e frenate aggressive.

4.2. Caso di studio sulle infrastrutture ciclabili

Il confronto tra infrastrutture ciclabili individua un gradiente osservato nell'intensità degli eventi di guida brusca fra le tipologie, condizionato ai controlli inclusi: esposizione, limite di velocità, numero di corsie e larghezza stradale. Le corsie delimitate solo da segnaletica sono associate a un punteggio superiore del 19,5% e le configurazioni promiscue con veicoli a un punteggio superiore dell'11,5%, entrambe rispetto alle piste separate.

Queste associazioni sono coerenti con la direzione suggerita dalla letteratura più ampia sulla sicurezza ciclistica, ma non stabiliscono da sole una graduatoria causale delle tipologie di infrastruttura. Il tipo di pista/corsia è fortemente correlato con la gerarchia stradale, il contesto di quartiere e i pattern di domanda: a Milano le piste separate tendono a trovarsi su strade più larghe e di classe superiore o lungo greenway periferiche, mentre le configurazioni solo segnaletica e promiscue sono più frequenti su strade minori nei quartieri densi. La specificazione OLS a sei predittori raggiunge $R^2 = 0,394$, lasciando quindi una sostanziale eterogeneità non modellata.

4.3. Implicazioni per gli interventi di sicurezza urbana

Nel complesso, i pattern riportati sostengono interventi di sicurezza specifici per il contesto anziché uniformi. Le strade che combinano grandi sezioni trasversali, visuali aperte e campi visivi non ostruiti emergono come candidate per restringimenti delle corsie, elementi visivi o fisici di moderazione del traffico e maggiore definizione dei margini stradali. I pattern temporali suggeriscono che controlli modulati per fascia oraria o una diversa temporizzazione semaforica possano essere appropriati sulle strade che concentrano il rischio per veicolo nelle ore notturne o nel weekend. Il gradiente relativo alle infrastrutture ciclabili è coerente con la priorità alla separazione fisica dove possibile e, dove la separazione fisica non è praticabile, con la preferenza per configurazioni promiscue esplicite rispetto alla sola delimitazione con vernice, subordinatamente a ulteriori verifiche causali.

4.4. Limiti

L'analisi è osservazionale e non può stabilire causalità. Variabili confondenti non misurate – intensità dei controlli, coordinamento semaforico, contesto socio-economico, volumi ciclistici, mix degli usi del suolo – possono spiegare parte dei pattern osservati. Lo snapshot OSM è fissato al 31 dicembre 2023, mentre le date di acquisizione di Google Street View variano tra luoghi; fattori transitori come meteo, lavori stradali o eventi speciali non sono modellati. Il conteggio dei probe usato come proxy del volume di traffico, pur derivando da un campione grande e rappresentativo, resta una misura semplificata dell'esposizione.

La flotta telematica combina due famiglie di dispositivi con caratteristiche di rumore differenti: i dispositivi Targa registrano gli eventi tramite accelerometro, mentre i MetaSystem li ricavano dalle variazioni di velocità misurate dal GNSS su una finestra di un secondo. Le due distribuzioni sono state armonizzate sulle soglie di severità Targa, ma non è stata eseguita un'analisi di sensibilità limitata ai soli eventi Targa né introdotto un effetto fisso del dispositivo nelle regressioni. Con l'analisi attuale non è quindi possibile quantificare quanto del pattern spaziale osservato rifletta la distribuzione spaziale dei tipi di dispositivo nella flotta.

Infine, per il caso di studio sulle infrastrutture ciclabili non sono stati applicati effetti fissi di quartiere né una stratificazione per classe stradale. I coefficienti riportati devono quindi essere letti come associazioni all'interno di questo specifico campione, non come una graduatoria causale dei tipi di progettazione; qualunque implicazione di policy infrastrutturale richiede ulteriori indagini causali.

4.5. Direzioni future

Dai limiti discussi emergono quattro direzioni per la ricerca futura. Primo, l'aggiunta di ulteriori predittori – meteo, luce diurna, segnali stradali rilevati, volumi pedonali – migliorerebbe l'interpretabilità del modello e catturerebbe effetti transitori. Secondo, una stima più avanzata dell'esposizione, per esempio la ricostruzione basata su CatBoost proposta da TomTom (2025), che integra conteggi dei probe con velocità a flusso libero, numero di corsie, classe funzionale della strada e attributi socio-demografici, rafforzerebbe il denominatore della metrica di tasso. Terzo, collegare il dataset degli eventi bruschi agli incidenti registrati dalla polizia – cosa non possibile in questo studio per vincoli di accesso ai dati – consentirebbe di validare l'indicatore surrogato rispetto agli esiti con feriti. Quarto, replicare l'analisi in altre città europee permetterebbe di valutarne direttamente la trasferibilità.

Ulteriori passi naturali includono una validazione formale della pipeline di map-matching su un insieme di dati etichettato manualmente e tenuto separato, un'analisi di sensibilità rispetto ai pesi di severità e alla famiglia di dispositivi usata per registrare gli eventi e l'impiego di modelli multiscala o gerarchici che consentano relazioni spaziali non stazionarie. Questi sviluppi potrebbero trasformare alcune delle associazioni qui osservate in affermazioni causali più difendibili.

Contributi degli autori (CRediT)

Andrea La Grotteria: concettualizzazione, cura dei dati, analisi formale, metodologia, software, visualizzazione, stesura della bozza originale, revisione ed editing. Paolo Santi: concettualizzazione, supervisione, metodologia, revisione ed editing. Titus Venverloo: concettualizzazione, metodologia, revisione ed editing. Umberto Fugiglando: concettualizzazione, acquisizione dei finanziamenti, revisione ed editing. Carlo Ratti: concettualizzazione, acquisizione dei finanziamenti, revisione ed editing.

Riferimenti bibliografici

I riferimenti sono mantenuti nella lingua e forma bibliografica dell'originale.

- Araldi, Alessandro and Giovanni Fusco (2019). “From the street to the metropolitan region: Pedestrian perspective in urban fabric analysis”. *Environment and Planning B: Urban Analytics and City Science* 46.7, pp. 1243–1263.
- Boeing, Geoff (2025). “Modeling and analyzing urban networks and amenities with OSMnx”. *Geographical Analysis*.
- Cabrera-Arnau, Carmen, Rafael Prieto Curiel, and Steven Richard Bishop (2020). “Uncovering the behaviour of road accidents in urban areas”. *Royal Society Open Science* 7.4, p. 191739.
- CINEA (2025). *EU Road Safety: Towards Vision Zero*.
- Comune di Bologna (2023). *Bologna diventa Città 30*.
- Eurostat (2025). *EU road fatalities decreased by 1.3% in 2023*.

- Fan, Zhuangyuan et al. (2023). “Urban visual intelligence: Uncovering hidden city profiles with street view images”. *Proceedings of the National Academy of Sciences* 120.27, e2220417120.
- Fleischmann, Martin (2019). “momepy: Urban Morphology Measuring Toolkit”. *Journal of Open Source Software* 4.43, p. 1807. doi: 10.21105/joss.01807.
- Greibe, Poul (2003). “Accident prediction models for urban roads”. *Accident Analysis & Prevention* 35.2, pp. 273–285.
- Iranmanesh, Aminreza, Can Kara, and Tuğsad Tülbentçi (2024). “Mapping the relationship between traffic accidents, road network configuration, and urban land use”. *International Journal of Injury Control and Safety Promotion* 31.4, pp. 672–685.
- Jain, Jitesh et al. (2022). “OneFormer: One Transformer to Rule Universal Image Segmentation”. *arXiv preprint arXiv:2211.06220*.
- Li, Yongchang et al. (2022). “Street View Imagery (SVI) in the built environment: A theoretical and systematic review”. *Buildings* 12.8, p. 1167.
- Ng, K.S., W.T. Hung, and W.G. Wong (2001). “Effects of urban design on road accidents in Hong Kong”. *WIT Transactions on The Built Environment* 52.
- Sadeghi-Bazargani, Homayoun and Mohammad Saadati (2016). “Speed management strategies; a systematic review”. *Bulletin of Emergency & Trauma* 4.3, p. 126.
- Silva, Iván and Jose Eugenio Naranjo (2020). “A systematic methodology to evaluate prediction models for driving style classification”. *Sensors* 20.6, p. 1692.
- TomTom (2025). *Historical Traffic Volumes*. Combines probe data and machine learning to estimate accurate traffic volumes network-wide.
- Vorko-Jovic, Ariana, Josipa Kern, and Zrinka Biloglav (2006). “Risk factors in urban road traffic accidents”. *Journal of Safety Research* 37.1, pp. 93–98.
- Yannis, George and Eva Michelaraki (2024). “Review of city-wide 30 km/h speed limit benefits in Europe”. *Sustainability* 16.11, p. 4382.
- Yao, Yao et al. (2019). “Exploring the relationship between risk perception, speed limit credibility and speed limit compliance”. *Transportation Research Part F: Traffic Psychology and Behaviour* 62, pp. 575–586.
- Ziakopoulos, Apostolos (2021). “Spatial analysis of harsh driving behavior events in urban networks using high-resolution smartphone and geometric data”. *Accident Analysis & Prevention* 157, p. 106189.
- Ziakopoulos, Apostolos et al. (2022). “Spatial predictions of harsh driving events using statistical and machine learning methods”. *Safety Science* 150, p. 105722.