European Research Studies Journal
Volume XXIX, Issue 4, 2026

pp. 61-68

Motorcycle Safety Management through Injury Severity Monitoring: Evidence from Poland

Submitted 13/09/26, 1st revision 25/09/26, 2nd revision 10/10/26, accepted 01/11/26

Artur Budzyński[1], Irena Szewczyk[2], Grzegorz Biesok[3], Asen Asenov[4], Velizara Pencheva[5]

Abstract:

Purpose: *This paper examines how national police-reported crash data can support road traffic safety (RTS) management by enabling systematic monitoring of motorcycle injury severity and helping authorities prioritise safety risks, without treating predictive scores as a substitute for clinical triage.*
Design/Methodology/Approach: *The study analyses 83,473 motorcycle riders (excluding moped riders) recorded in Poland's SEWIK police crash registry over the period 2015–2025. Injury severity, classified into five ordered categories, is modelled using the CatBoost multi-class gradient-boosting algorithm, while SHAP (Shapley Additive Explanations) is used to interpret feature associations. A data-quality audit identifies registry-process artefacts, including vehicle production year, protective-equipment status and calendar year. These variables are excluded because their inclusion substantially undermines out-of-time predictive performance around a 2021 recording discontinuity. Model performance is then assessed using walk-forward validation.*
Findings: *Following feature ablation, ordinal agreement remains modest, with a mean quadratic weighted kappa of approximately 0.25 across walk-forward validation. Crash type, posted speed limit, licensing status and impairment exhibit the strongest associations with predicted injury severity. The results highlight the importance of addressing temporal discontinuities and administrative recording artefacts when developing predictive monitoring tools from crash data.*
Practical implications: *Road safety authorities can use the proposed pipeline as a supplementary decision-support tool to monitor injury-severity patterns, identify risk factors warranting further investigation and inform the prioritisation of motorcycle safety*

[1]*Assist. Prof., Silesian University of Technology, Faculty of Transport and Aviation Engineering, Department of Transport Systems, Traffic Engineering and Logistics, Poland, ORCID: 0000-0002-5803-6749, e-mail: artur.budzynski@polsl.pl;*
[2]*Assist. Prof., University of Bielsko-Biala, Department of Management and Transport, Poland, ORCID: 0000-0001-6038-0525, e-mail: iszewczyk@ubb.edu.pl;*
[3]*Assist. Prof., University of Bielsko-Biala, Department of Management and Transport, Poland, ORCID 0000-0002-4987-5823, e-mail: gbiesok@ubb.edu.pl;*
[4]*Prof. PhD, University of Ruse "Angel Kanchev", Faculty of Transport, Bulgaria, ORCID:0000-0002-9770-0403, e-mail: asasenov@uni-ruse.bg;*
[5]*Prof. PhD, University of Ruse "Angel Kanchev", Faculty of Transport, Bulgaria, ORCID:0000-0002-9809-4759, e-mail: vpencheva@uni-ruse.bg;*

assessments. The findings may help practitioners structure dashboard indicators around crash characteristics, speed limits, licensing and impairment, while incorporating routine data-quality checks and temporal validation. Given the modest predictive agreement, outputs should support—not determine—resource allocation and should not be interpreted as individual-level risk assessments or evidence of intervention effectiveness.
Originality/value: *The paper contributes to RTS management by integrating explainable machine learning, temporal validation and explicit data-quality governance into a transparent pipeline for monitoring motorcycle injury severity using national police crash data. It provides a framework for responsible use of predictive analytics in safety management while distinguishing statistical associations from causal effects.*

Keywords: *Road traffic safety management, motorcycle safety, injury severity monitoring, risk prioritization, decision support, explainable machine learning, Poland.*

JEL classification: *R41, C53, C55, I18.*

Paper type: *Research article.*

Acknowledgements: *Co-financed by the Minister of Science under the "Regional Excellence Initiative" Poland. The authors also thank the Opinion and Analysis Department of the Road Traffic Bureau at the General Police Headquarters of Poland for providing access to SEWIK data for research.*

Declaration: *Generative artificial intelligence tools were employed during the preparation of this work for language proofreading and style improvement. The author remains fully responsible for the final text and its scientific integrity, having thoroughly reviewed and revised the generated content afterward.*

1. Introduction

Road traffic deaths and serious injuries remain a strategic concern for European transport policy and for organisations that interact with the road system (European Commission, n.d.; World Health Organization, 2023). Motorcycle riders are a vulnerable road-user group whose exposure, kinetic energy, and mixed-traffic contexts differ from passenger-car outcomes (Vlahogianni *et al.,* 2012). Within road traffic safety management, the practical question for authorities is less whether a model can detect every fatality and more which operating contexts should enter a monitoring dashboard and where registry quality should be improved first.

National crash registries support large-sample severity studies when vehicles and participants can be linked consistently. In Poland, SEWIK underpins recent interpretable severity work on cyclists framed for sustainable road safety (Budzyński and Cieśla, 2026). Crash-data quality reviews show that incomplete or inconsistently

coded fields limit what statistical and machine-learning models can support (Imprialou and Quddus, 2019). Recording practice is therefore part of the management problem, not a technical footnote.

Motorcycle injury severity has a long econometric and statistical tradition (Savolainen and Mannering, 2007; Savolainen *et al.,* 2011; Eluru *et al.,* 2008), later emphasising unobserved heterogeneity and temporal instability of parameters (Mannering *et al.,* 2016; Mannering, 2018). Tree-based ensembles with SHAP have been applied to powered two-wheeler crashes in several countries (Wei *et al.,* 2023; Wisutwattanasak *et al.,* 2024). Evidence that links motorcycle severity monitoring explicitly to RTS management practice, including registry-artefact audit, remains thinner.

This study isolates motorcycle riders (MOTO; mopeds excluded), extends the window through 2025, audits registry-process covariates, estimates a CatBoost severity model, and reports associational SHAP for safety-management monitoring. Fuller spatial leave-one-region protocols and extensive baseline horse-races are reserved for subsequent technical work.

Contributions: (i) an RTS-management framing of motorcycle injury severity monitoring; (ii) an audited national sample (N = 83,473); (iii) transparent non-causal association ranking for dashboards; (iv) an explicit warning against using the scores as fatal-case classifiers.

2. Theoretical Background

ISO 39001 specifies requirements for a road traffic safety (RTS) management system so that an organisation can reduce death and serious injury related to crashes it can influence (ISO, 2012). The logic is managerial: policy and objectives, identification of RTS performance factors, action plans, monitoring, and continual improvement. Severity monitoring from administrative data fits the performance evaluation and risk/opportunity steps: it does not replace enforcement or engineering measures, but it structures which contexts deserve attention.

European research on road-safety decision support systems distinguishes risk factors (contributors to crash occurrence or injury severity) from measures (interventions that prevent crashes or mitigate injury), and links both to stakeholder choices (Yannis *et al.,* 2016). In that framing, a registry model is useful when it ranks observable contexts (crash type, posted limit band, licensing flag) that managers can monitor, not when it claims individual-level triage.

If fill rates of key fields break between years, machine-learning scores can track the registry period rather than road risk. For RTS management, that is a control failure: year-to-year trends in model scores become misleading (Imprialou and Quddus, 2019; Mannering, 2018).

Auditing and ablating process artefacts is therefore a governance step before dashboard use. Gradient boosting with SHAP (Prokhorenkova *et al.,* 2018; Lundberg and Lee, 2017) can surface which coded factors move predicted severity. For management use, explanations must stay associational: they support prioritisation and hypothesis generation, not causal claims about helmets, travel speed, or licensing enforcement without a separate identification design.

3. Materials and Methods

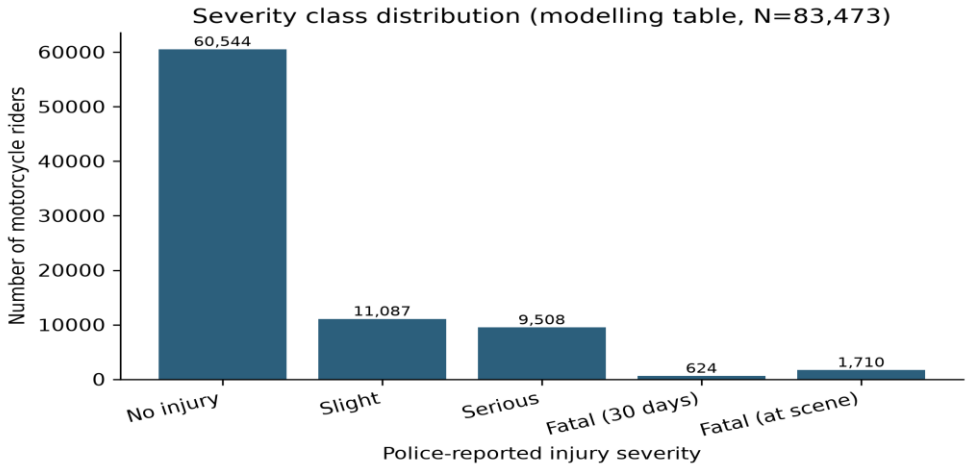
Primary source: SEWIK, Poland's national police road-crash registry. Harmonised vehicle–participant extracts for 2015–2025. Primary sample: motorcycle riders (role = driver) with vehicle type in {motorcycle, motorcycle (other), motorcycle ≤125 cm³}. Mopeds and quads excluded. After deduplicating repeated (event, rider) keys in the raw join (86,169 → 83,473), the modelling table matches the derived participant panel (Table 1). Surplus duplicates are recording artefacts (higher fatal share among surplus rows), not a severity-based exclusion rule.

Table 1. *Modelling-table severity counts for MOTO riders (SEWIK, Poland, 2015–2025).*

Indicator	**N**
MOTO riders in modelling table	83,473
No injury	60,544
Slight	11,087
Serious	9,508
Fatal within 30 days	624
Fatal at scene	1,710

Source: *Own study.*

Figure 1. *Police-reported injury severity distribution for MOTO riders in the modelling table (SEWIK Poland, 2015–2025; N = 83,473).*



Source: *Own study.*

Outcome: police-reported ordered severity in five classes (no injury, slight, serious, fatal within 30 days, fatal at scene). Both fatal codes denote death; separation follows SEWIK reporting practice. Predictors for manuscript models exclude four fields with registry-leak or temporal-shift risk: vehicle production year and its missingness indicator (missingness falls from about 95% in 2015–2020 to about 20–25% from 2021), the protective-equipment flag (rare affirmative coding concentrated among injured/killed in audits), and calendar year.

Remaining fields include motorcycle type, demographics and licensing, impairment and unprotected-riding flags, fault and partner vehicle group, posted speed limit (km/h at the location as recorded; not measured travel speed), intersection and surface, crash type and cause fields, lighting, simplified weather and geometry, and road type.

Learner: CatBoost multi-class gradient boosting (MultiClass loss; balanced class weights). Primary ordinal metric: quadratic weighted kappa (QWK); macro-F1 and balanced accuracy as complements. Evaluation reported here: expanding-window walk-forward for test years 2019–2025 (train on all earlier years).

Hyperparameters: depth 8, learning rate 0.05, L2 leaf regularisation 5, at most 180 rounds with early stopping on a stratified train-only validation slice. Explainability: global mean absolute SHAP on the last temporal fold; signed contribution to predicted expected severity obtained by transforming multi-class class-margin SHAP toward the model's expected severity score (associational, not causal). Helmet counterfactuals are out of scope.

4. Results

With the ablated feature set, walk-forward mean QWK ≈ 0.25 (test years 2019–2025; yearly QWK roughly 0.22–0.27). Macro-F1 remains about 0.25–0.28. Precision for fatal within 30 days on the last fold is near zero: the model is not a fatal-case detector and must not be used for triage or screening.

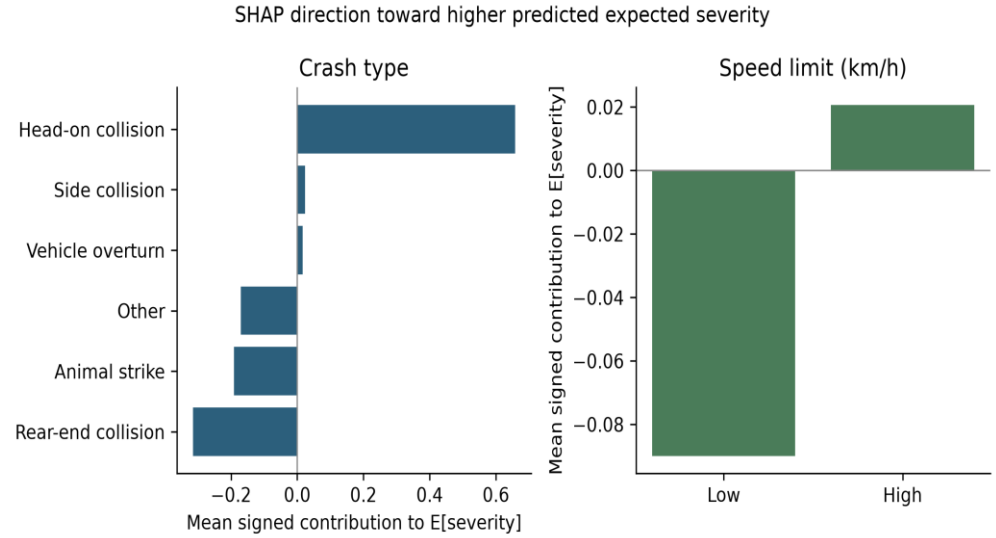
When production year, protective-equipment, and calendar year are retained, walk-forward QWK collapses toward chance in 2021–2022 (about 0.05 versus about 0.22–0.24 without them). That collapse is the managerial result on data quality: fill-rate breaks must be audited before interpreting machine-learning severity scores over time.

On the 2025 fold, built-in importance and mean absolute SHAP agree that leading non-artefact associations are crash type, posted speed limit, licensing entitlement, and impairment. Signed contributions to predicted expected severity (2025 subsample) show, in particular: head-on collisions associate toward higher predicted expected severity and rear-end collisions toward lower; the higher posted-limit band associates toward higher predicted expected severity (median split on discrete limit

values); no licence associates toward higher predicted expected severity than recorded entitlement; recorded alcohol or other psychoactive flags associate toward higher predicted expected severity where coded (many cases remain unknown or not tested).

These patterns align directionally with international motorcycle severity evidence on speed context and crash configuration (Billah et al., 2023; Wisutwattanasak *et al.,* 2024), but remain registry associations under the fitted model (Figure 2).

Figure 2. *Signed contribution to predicted expected severity (from MultiClass class-margin SHAP) for a leading categorical feature (left) and posted speed limit (right; low vs high median split) on a subsample of the 2025 walk-forward test fold.*



Source: *Own study.*

5. Discussion

Under an ISO 39001-style RTS logic (ISO, 2012), the practical outputs for motorcycle safety management are: (1) prioritise crash-type and posted-limit contexts in safety dashboards; (2) treat licensing and impairment flags as high-value fields for cross-agency follow-up and data-completeness campaigns; (3) do not deploy these scores as fatal-event classifiers; (4) invest in registry quality (production year, protective equipment) before any causal protective-equipment evaluation.

The design extends national SEWIK severity modelling from cyclist work (Budzyński and Cieśla, 2026) to motorcycles with an explicit artefact audit and an RTS-management reading of the results. Relative to classical severity methods (Savolainen *et al.,* 2011), we keep a five-class police outcome and a monitoring-

oriented interpretation. Limitations include police selection and residual coding breaks; MultiClass loss without ordinal distances; modest QWK; no validated helmet counterfactual; and young ages retained without legal-eligibility filters (age under 18: 5.1% of N; may include unlicensed or underage riding or coding anomalies). Findings are not causal effects of posted limit, travel speed, licensing, or protective equipment.

6. Conclusions

Motorcycle rider injury severity can be monitored for road traffic safety management when registry-process artefacts are audited and models are read as association rankings (Poland, SEWIK 2015–2025; N = 83,473). Walk-forward ordinal agreement remains modest (mean QWK ≈ 0.25). Crash type, posted speed limit, licensing, and impairment are the leading associations for dashboard prioritisation. Deeper spatial protocols and full baseline comparisons are reserved for subsequent technical work.

Author Contributions: *Conceptualization, methodology, software, formal analysis, investigation, data curation, original draft, supervision, and project administration: A.B. Review and editing: A.B., I.S., B.G., A.A. and V.P. All authors have read and agreed to the published version of the manuscript.*
Conflicts of Interest: *The authors declare no conflicts of interest.*
Data Availability: *Aggregated tables and the feature inventory are available from the corresponding author upon reasonable request, subject to SEWIK access rules. Unit-level police records are not redistributed.*

References:

Billah, K., Sharif, H.O., Dessouky, S. 2023. Bivariate-logit-based severity analysis for motorcycle crashes in Texas, 2017-2021. Sustainability, 15(13), 10377. https://doi.org/10.3390/su151310377.

Budzyński, A., Cieśla, M. 2026. Explainable machine learning for cyclist injury severity in bicycle–vehicle crashes in Poland: Association patterns and implications for sustainable road safety. Sustainability, 18(11), 5501. https://doi.org/10.3390/su18115501.

Eluru, N., Bhat, C.R., Hensher, D.A. 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accident Analysis & Prevention, 40(3), pp. 1033-1054. https://doi.org/10.1016/j.aap.2007.11.010.

European Commission. n.d. EU road safety policy framework / Vision Zero materials. Brussels: European Commission.

Imprialou, M., Quddus, M. 2019. Crash data quality for road safety research: Current state and future directions. Accident Analysis & Prevention, 130, pp. 84-90. https://doi.org/10.1016/j.aap.2017.02.022.

ISO. 2012. ISO 39001:2012 Road traffic safety (RTS) management systems — Requirements with guidance for use. Geneva: International Organization for Standardization.

Lundberg, S.M., Lee, S.I. 2017. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30 (NeurIPS 2017).

Mannering, F. 2018. Temporal instability and the analysis of highway accident data. Analytic Methods in Accident Research, 17, pp. 1-13. https://doi.org/10.1016/j.amar.2017.10.002.

Mannering, F.L., Shankar, V., Bhat, C.R. 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Analytic Methods in Accident Research, 11, pp. 1-16. https://doi.org/10.1016/j.amar.2016.04.001.

Prokhorenkova, L., Gulin, A., Gusev, G., Vorobev, A., Dorogush, A.V. 2018. CatBoost: unbiased boosting with categorical features. In: Advances in Neural Information Processing Systems 31 (NeurIPS 2018).

Savolainen, P., Mannering, F. 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Accident Analysis & Prevention, 39(5), pp. 955-963. https://doi.org/10.1016/j.aap.2006.12.016.

Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A. 2011. The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. Accident Analysis & Prevention, 43(5), pp. 1666-1676. https://doi.org/10.1016/j.aap.2011.03.025.

Vlahogianni, E.I., Yannis, G., Golias, J.C. 2012. Overview of critical risk factors in Power-Two-Wheeler safety. Accident Analysis & Prevention, 49, pp. 12-22. https://doi.org/10.1016/j.aap.2012.04.009.

Wei, T., Zhu, T., Lin, M., Liu, H. 2023. Predicting and factor analysis of rider injury severity in two-wheeled motorcycle and vehicle crash accidents based on an interpretable machine learning framework. Traffic Injury Prevention. https://doi.org/10.1080/15389588.2023.2284111.

Wisutwattanasak, P., Se, C., Champahom, T., Kasemsri, R. et al. 2024. Factors affecting single and multivehicle motorcycle crashes: Insights from day and night analysis using XGBoost-SHAP algorithm. Big Data and Cognitive Computing, 8(10), 128. https://doi.org/10.3390/bdcc8100128.

World Health Organization. 2023. Global status report on road safety 2023. Geneva: WHO.

Yannis, G. et al. 2016. Developing the European road safety decision support system. Injury Prevention, 22(Suppl 2), A30. https://doi.org/10.1136/injuryprev-2016-042156.80.