Exposure-Normalized Bicycle Crash Risk Along Berlin Routes

Eric Berger [*1] **Edward Eichhorn** [*2] **Liaisan Faidrakhmanova** [*3] **Luise Grasl** [*4] **Tobias Schnarr** [*5]

Abstract

We plan to investigate route-level bicycle crash risk in Berlin by addressing a key limitation in current safety analyses: raw crash counts don't distinguish between dangerous roads and simply busy roads. Our approach will combine police-recorded crash data from the German *Unfallatlas* with a city-wide dataset of measured bicycle volumes to calculate a normalized risk metric at the street-segment level. We aim to estimate this risk conditional on contextual factors (e.g., time of day and weather) to capture temporal variability. The core idea is to identify which segments have high crash rates per cyclist rather than just high absolute crash numbers. We will then explore how to aggregate these segment-level risks into route-level scores that could inform safer cycling navigation. Building on this, we aim to score a user-specified route and, where possible, recommend alternative routes with lower estimated risk while maintaining comparable convenience (e.g., distance or travel time).

1. Introduction

Cycling safety analyses often rely on raw crash counts, which can conflate danger with demand: streets that attract many cyclists accumulate more incidents even if per-rider risk is low (Lücken, 2018). This obscures where infrastructure is intrinsically risky and limits both targeted interventions and everyday route choice, especially in dense networks like Berlin's (Uijtdewilligen et al., 2024). We address this by estimating bicycle crash risk per unit exposure on the street network and by propagating segment-level risk to route-level scores suitable for navigation.

Our study combines police-recorded crashes from the German *Unfallatlas* (Berlin subset) (Statistische Ämter des Bundes und der Länder, 2025) with a city-wide dataset of measured bicycle volumes at the street-segment level (Kaiser, 2025). We compute a normalized segment-level risk metric and aim to estimate this risk conditional on contextual factors (e.g., time of day and weather) to capture temporal variability. We then aggregate segment scores to evaluate arbitrary routes and recommend alternatives that reduce estimated risk while maintaining comparable convenience (distance or travel time), aligning with emerging work on safety-aware routing (Wage et al., 2022).

Our work makes the following contributions: (i) a reproducible pipeline for exposure-normalized, context-conditional risk at the segment scale using measured volumes rather than inferred exposure; and (ii) a route-scoring and recommendation procedure that balances risk reduction against convenience.

The paper is organized as follows: we review related work in Section 2, describe the data and methods in Section 3, present results in Section 4, and discuss and conclude in Section 5.

2. Related Work

To avoid conflating danger with demand, the literature normalizes crash risk by cyclist exposure (Lücken, 2018). City-scale studies therefore normalize by exposure and show that finer temporal resolution improves inference while noting under-reporting in police records (Uijtdewilligen et al., 2024). To obtain exposure, some work extrapolates city-wide volumes from sparse counters using learning-based models and multi-source features, showing that short campaigns improve accuracy at new locations (Kaiser et al., 2025a). More recent efforts release street-segment datasets of measured bicycle volumes, enabling downstream safety analyses without building an exposure model (Kaiser et al., 2025b). At the network level, studies define risk as crashes per exposure (often annualized), discuss link snapping and intersection allocation, and integrate risk into routing with convenience constraints (Wage et al., 2022). In Berlin,

[*]Equal contribution [1]Matrikelnummer 12345678, MSc Machine Learning [2]Matrikelnummer 12345678, MSc Computer Science [3]Matrikelnummer 12345678, MSc Media Informatics [4]Matrikelnummer 12345678, MSc Medical Informatics [5]Matrikelnummer 12345678, MSc Medical Informatics. Correspondence to: Initials1 <first1.last1@uni-tuebingen.de>, Initials2 <first2.last2@uni-tuebingen.de>, Initials3 <first3.last3@uni-tuebingen.de>, Initials4 <first4.last4@uni-tuebingen.de>, Initials5 <first4.last4@uni-tuebingen.de>.

Project report for the "Data Literacy" course at the University of Tübingen, Winter 2025/26 (Module ML4201). Style template based on the ICML style files 2025.

case analyses highlight intersection concentration and caution that infrastructure comparisons require controlling for exposure (Medeiros et al., 2021). Building on exposure-normalized analysis (Uijtdewilligen et al., 2024), measured volumes (contrasting extrapolation) (Kaiser et al., 2025a;b), and route-level integration (Wage et al., 2022), we estimate context-conditional segment risk for Berlin and aggregate it to routes, consistent with local evidence on exposure control (Medeiros et al., 2021).

3. Data and Methods

4. Results

5. Discussion & Conclusion

References

Kaiser, S. K. Data from: Spatio-temporal graph neural network for urban spaces: Interpolating citywide traffic volume, May 2025. URL https://doi.org/10.5281/zenodo.15332147.

Kaiser, S. K., Klein, N., and Kaack, L. H. From counting stations to city-wide estimates: data-driven bicycle volume extrapolation. *Environmental Data Science*, 4:e13, 2025a. doi: 10.1017/eds.2025.5.

Kaiser, S. K., Rodrigues, F., Azevedo, C. L., and Kaack, L. H. Spatio-temporal graph neural network for urban spaces: Interpolating citywide traffic volume, 2025b. URL https://arxiv.org/abs/2505.06292.

Lücken, L. On the variation of the crash risk with the total number of bicyclists. *European Transport Research Review*, 10(2):33, 2018. doi: 10.1186/s12544-018-0305-9. URL https://doi.org/10.1186/s12544-018-0305-9.

Medeiros, R. M., Bojic, I., and Jammot-Paillet, Q. Spatiotemporal variation in bicycle road crashes and traffic volume in berlin: Implications for future research, planning, and network design. *Future Transportation*, 1(3):686–706, 2021. ISSN 2673-7590. doi: 10.3390/futuretransp1030037. URL https://www.mdpi.com/2673-7590/1/3/37.

Statistische Ämter des Bundes und der Länder. Unfallatlas. https://unfallatlas.statistikportal.de/, 2025. Interaktive Kartenanwendung zu Straßenverkehrsunfällen mit Personenschaden in Deutschland.

Uijtdewilligen, T., Ulak, M. B., Wijlhuizen, G. J., Bijleveld, F., Geurs, K. T., and Dijkstra, A. Examining the crash risk factors associated with cycling by considering spatial and temporal disaggregation of exposure: Findings from four dutch cities. *Journal of Transportation Safety & Security*, 16(9):945–971, 2024. doi: 10.1080/19439962.2023.2273547. URL https://doi.org/10.1080/19439962.2023.2273547.

Wage, O., Bienzeisler, L., and Sester, M. Risk analysis of cycling accidents using a traffic demand model. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B4-2022:427–434, 2022. doi: 10.5194/isprs-archives-XLIII-B4-2022-427-2022. URL https://isprs-archives.copernicus.org/articles/XLIII-B4-2022/427/2022/.