Predicting the Risk of Vehicle–Cyclist Crashes in Urban Intersections Using Bayesian Networks

Kranthi Kumar Talluri KranthiKumar.Talluri@th-ab.de
Connected Urban Mobility – Intelligent Traffic Infrastructure Aschaffenburg University of Applied Sciences, Germany

Anders L Madsen anders@hugin.com
HUGIN EXPERT A/S and Department of Computer Science, Aalborg University, Denmark

Galia Weidl Galia.Weidl@th-ab.de
Connected Urban Mobility – Intelligent Traffic Infrastructure Aschaffenburg University of Applied Sciences, Germany

Christer Ahlström christer.ahlstrom@vti.se
Swedish National Road and Transport Research Institute (VTI), Department of Biomedical Engineering, Linköping University, Linköping, Sweden

Johan Olstam johan.olstam@vti.se
Swedish National Road and Transport Research Institute (VTI), Department of Science and Technology, Linköping University, Norrköping, Sweden

Editors: Gustau Camps-Valls, Manuele Leonelli and Gherardo Varando

Abstract

The i4Driving European Union project aims to establish a credible and realistic human road safety baseline for virtual assessment of Connected, Cooperative, and Automated Mobility (CCAM) systems. As part of these efforts, multiple driving simulator experiments have been conducted to collect data to verify the developed human driver models. This paper focuses on the analysis of data from two urban road experiments in the same scenario involving a vehicle and a bicycle. The paper describes the development of a knowledge-driven Bayesian network (BN) to predict the risk of a right-turn maneuver involving a bicycle traveling straight through a four-way urban traffic intersection. The paper also reports on the results of an experimental analysis on data from the two urban road experiments, each involving several hundred driving sequences with a population of different human drivers. The predictive accuracy of the knowledge-driven model is high with a significant prediction horizon. That is, in most cases, the model can predict a high-risk scenario with high certainty in time to warn the driver, allowing a potential collision to be avoided.

Keywords: Bayesian networks, urban traffic collisions, experiments

1. Introduction

Urban intersections are among the most safety-critical environments in road transportation, particularly when interactions involve vulnerable road users such as cyclists and pedestrians. Right-turn conflicts between motor vehicles and cyclists are especially challenging because they often involve covert yield requirements, where drivers must recognise and act on cyclists' right of way despite limited support from salient infrastructure cues, while also involving complex interactions among vehicle dynamics, cyclist behaviour, driver decision-making, and environmental conditions (Kircher et al., 2026). Reliable methods for early

© 2026 K.K. Talluri, A.L. Madsen, G. Weidl, C. Ahlström & J. Olstam.

identification of hazardous situations are therefore essential for the development and assessment of CCAM systems. Traditional traffic-safety analysis frequently relies on surrogate safety measures such as Time-to-Collision (TTC), Post-Encroachment Time (PET), and related conflict indicators. While these measures have proven useful for proactive safety assessment, they often rely on deterministic thresholds with limited uncertainty awareness. Recent work has demonstrated the effectiveness of conflict-based and TTC-based approaches for real-time crash-risk assessment (Arun et al., 2021; Ali et al., 2023; Li et al., 2023). A recent review highlights the growing use of probabilistic and machine-learning methods for traffic safety analysis while emphasizing the need for uncertainty-aware decision support (Sandeep et al., 2026).

Probabilistic graphical models provide a principled framework for reasoning under uncertainty. In particular, Bayesian networks combine expert knowledge and observational data within an interpretable probabilistic model and support efficient inference under incomplete information (Pearl, 1988; Jensen et al., 1990; Kjærulff and Madsen, 2013). Their successful application in decision support, reliability engineering, and human reliability assessment demonstrates their suitability for safety-critical domains (Weidl et al., 2005; Leva and Hansen, 2017; Briwa et al., 2026). Bayesian methods have also been applied to transportation problems such as traffic-conflict prediction and cyclist trajectory estimation (Gao et al., 2021). However, knowledge-driven Bayesian networks for collision-risk prediction in cyclist–vehicle interactions remain relatively unexplored.

The aim of this paper is to investigate whether a knowledge-driven Bayesian network can accurately predict collision risk in urban right-turn cyclist scenarios while providing a sufficient prediction horizon for preventive actions. The proposed model combines driver characteristics, vehicle dynamics, cyclist behaviour, and conflict-point-based temporal indicators. Its performance is evaluated using two driving simulator datasets collected within the Horizon Europe project *i4Driving*[1]: a VTI right-turn cyclist dataset and an independent TUM right-turn cyclist dataset, enabling within-VTI evaluation and assessment of the VTI-derived model on the TUM dataset.

The contributions of this paper are: (1) the development of a knowledge-driven Bayesian network for collision-risk prediction in right-turn cyclist conflicts, (2) an experimental evaluation using two simulator datasets, and (3) an analysis of prediction horizon and evaluation of the VTI-derived risk threshold on the independent TUM dataset.

The remainder of the paper is organized as follows. Section 2 presents the case study and datasets, Section 3 describes the methodology, Section 4 presents the results, Section 5 discusses the findings and limitations, and Section 6 concludes the paper.

2. Application and Datasets

2.1. Case Study: Right-turn Cyclist Scenario

The case study considered in this paper involves an urban intersection where an ego vehicle performs a right turn while a cyclist continues straight ahead with priority. Such situations are safety-critical because the cyclist may remain partially occluded or unnoticed until shortly before the conflict point (CP). The objective is therefore to identify potentially

1. https://i4driving.eu/

hazardous situations early enough to support preventive actions. To model this interaction, a BN combines contextual information, vehicle dynamics, predicted trajectories, and driver characteristics such as age and driving style. The resulting BN estimates collision risk and supports interpretable reasoning about the factors contributing to unsafe situations.



(*a*) Situation features (*b*) Right turn scenario

Figure 1: Right-turn cyclist scenario: (a) interaction features and conflict area; (b) ego vehicle turning right while the cyclist proceeds straight.

The model is evaluated using driving simulator datasets collected independently at the Swedish National Road and Transport Research Institute (VTI) and the Technical University of Munich (TUM). Figure 1 illustrates the situation features (adopted from (Weidl et al., 2013) for the right-turn scenario and the conflict area defined by the intersection of the vehicle and cyclist trajectories (Ahlström et al., 2025). Risk assessment is based on the CP, representing the possible pair-wise intersection of movement trajectories, the time-to-collision (TTC), and the time-to-conflict-point (TTCP), which form the basis for the feature extraction described in the following section.

2.2. Datasets and Sequence Preprocessing

For the analysis in this paper, subsets of driving simulator data were selected from VTI in Linköping, Sweden, and TUM in Garching, Germany. The VTI subset represents a four-way yield-controlled intersection where the ego vehicle turns right while a cyclist proceeds straight in a dedicated bicycle lane. The dataset includes 44 participants and, after preprocessing, 260 unique sequences were obtained. The TUM subset represents a similar right-turn scenario at a four-way intersection, where the ego vehicle turns right while a cyclist with priority proceeds straight. The dataset includes 33 participants and, after preprocessing, 127 sequences were identified.

Both datasets were filtered and reorganized to contain the same set of kinematic variables. These include time, ego and bicycle distances to the CP, and lateral and longitudinal velocities and accelerations of both road users. Each sequence represents one ego vehicle–cyclist interaction and consists of consecutive timesteps describing the evolution of that interaction. Each data row in the sequence corresponds to one timestep. The distances to the CP are represented by `ego_to_cp` and `bike_to_cp`. Within each sequence, a road user is considered to have crossed the CP at the timestep where the absolute distance to the CP is smallest. Distances are positive while approaching the CP and negative after crossing it.

An important difference between the experimental designs concerns the visibility of the cyclist before the interaction. In the VTI experiment, the cyclist emerged from a side street after the ego vehicle had passed the relevant approach point and was therefore not visible during the approach to the intersection. This created a covert conflict scenario, in which the cyclist could easily remain undetected unless the driver actively checked the rear-view and side mirrors or performed an over-the-shoulder glance before initiating the turn. In the TUM experiment, the cyclist was visible ahead of the ego vehicle throughout the approach, making the interaction overt rather than covert and enabling the driver to reduce speed and yield behind the cyclist before reaching the CP.

3. Methodology

3.1. Preliminaries

A Bayesian network $N = (G, \mathcal{P})$ is an efficient representation of the joint probability distribution $P(\mathcal{X})$ over variables $\mathcal{X}$ represented one-to-one by the vertices V of the directed acyclic graph $G = (V, E)$, where E denotes the directed edges. The joint probability distribution factorizes $P(\mathcal{X}) = \prod_{X \in \mathcal{X}} P(X \mid \pi(X))$, where $\pi(X)$ are the parents of X in G, see, e.g., (Kjærulff and Madsen, 2013). We consider Bayesian networks with both discrete and continuous chance variables. Continuous chance variables are assumed to follow a Conditional linear Gaussian distribution (CG), see, e.g., (Lauritzen and Jensen, 2001). Probabilistic inference in N is the task of computing all marginals given a set of evidence $\epsilon = \epsilon_i, \ldots, \epsilon_n$ where ϵ_i is the instantiation of a variable X as a state or values x, i.e., $X = x$. Probabilistic inference can, for example, be performed in a junction tree using HUGIN propagation (Jensen et al., 1990).

3.2. Bayesian Network Model

Figure 2 shows the DAG structure G of the Bayesian network $N = (G, \mathcal{P})$ developed to predict collision risk (by the Boolean variable denoted *Collision_Risk*) involving a vehicle and a bicycle. The Bayesian network developed in HUGIN (Madsen et al., 2005) has a total of 18 chance variables consisting of five continuous variables and 13 discrete variables. The colors in Figure 2 are used for visual organization of the network and do not affect the probabilistic model.

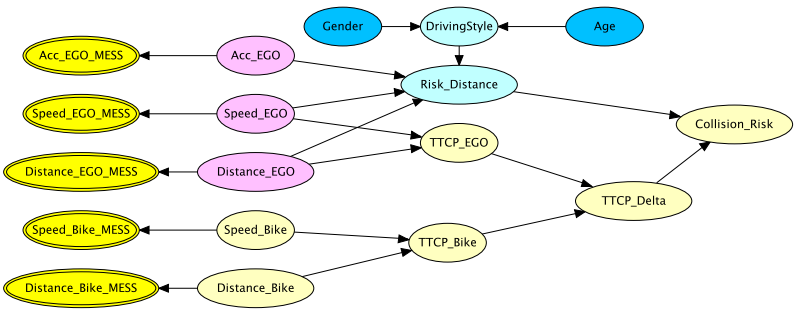


Figure 2: The DAG G of the Bayesian network for collision risk prediction.

The DAG G was determined using subject matter expert knowledge and physical relations of the problem domain. The DAG G has two different parts that, in combination, determine the risk of collision: 1) driving style and distance of the ego vehicle to the CP, and 2) the difference in time to collision between the ego vehicle and the bike.

The distance variables are defined relative to the CP, e.g., *Distance_Bike* is the distance of the bike to the CP. The units of the distance, speed, and acceleration variables are m, m/s, and m/s^2, respectively, while the unit of time is s where m is meters and s is seconds. *TTCP_Delta* specifies the time to CP (TTCP) and it is computed as the difference between *TTCP_EGO* and *TTCP_Bike* that both depend on speed and distance.

All variables except the *_MESS* variables used to capture the observed continuous properties are discretized. For instance, *Age* is discretized into three different age groups, taking the grouping of the participants into consideration. The approach of, e.g., (Weidl et al., 2005), was used to model observed physical entities such as acceleration, speed, and distance. The speed of the bike is modeled with two variables, where *Speed_Bike* represents the true speed of the bike and *Speed_Bike_MESS* represents the measured speed of the bike. This approach has two advantages. First, it captures uncertainty or variation in the measurement. Second, entering an observation on *Speed_Bike_MESS* results in a probability distribution over the discretized states of *Speed_Bike*, thereby reducing the effect of sharp discretization boundaries. The discretization of variables takes into account the physical limitations of bike and vehicle such as maximum speed and acceleration.

Table 1: Conditional probability tables for variables *DrivingStyle* and *Speed_Bike_MESS*.

Gender	Woman			Man		
Age	18–25	25–65	65–∞	18–25	25–65	65–∞
defensive	0.875	0.9	0.92	0.6	0.7	0.9
offensive	0.125	0.1	0.08	0.4	0.3	0.1

(a) $P(\textit{DrivingStyle} \mid \textit{Gender}, \textit{Age})$

Speed_Bike	0–1	1–2	2–3	3–4	4–5	5–∞
Mean	0.5	1.5	2.5	3.5	4.5	5.5
Variance	0.0625	0.0625	0.0625	0.0625	0.0625	0.0625

(b) $P(\textit{Speed_Bike_MESS} \mid \textit{Speed_Bike})$

The CPTs for variables are mainly specified using formulas reflecting physical relations and expert assessment using an anchoring approach similar to (Leva and Hansen, 2017) and (Briwa et al., 2026). Table 1 shows the conditional probability tables of variables *DrivingStyle* Table 1(a) and *Speed_Bike_MESS* Table 1(b), respectively. The *DrivingStyle* variable represents two assumed driving-style states: defensive and offensive, representing relatively cautious and relatively assertive driving behaviour, respectively. Its conditional probability distribution given *Gender* and *Age* was specified by the authors based on expert assumptions used for this model. In particular, the model assigns different probabilities to the two driving-style states across age and gender groups; these relationships should be interpreted as modeling assumptions rather than empirical findings from the datasets.

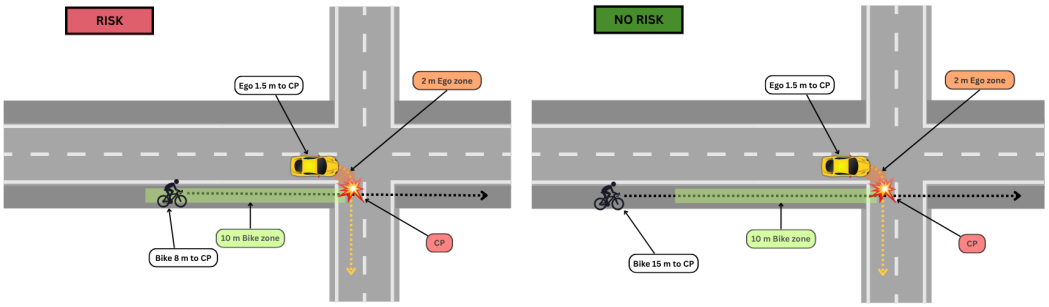
The conditional distribution of *Speed_Bike_MESS* given *Speed_Bike* represents measurement uncertainty using a Normal distribution centered at the midpoint of the corresponding *Speed_Bike* interval. The variance is defined by the interval width, with a coefficient of variation of 10%.

The identified variables are based on the properties measured in the data. Other variables can be included later, e.g., glance (where the driver looked and when), eye movements, and workload. There are significant challenges in taking variables such as glance into consideration, as a driver may look in the right direction but not see the bike or pedestrian due to, for example, occlusion.

3.3. Data labels: Manual Classification Rule

Each sequence is classified as either risk or no_risk using a deterministic geometric rule based on the distances of the ego vehicle and cyclist to the CP. A sequence is labeled as risk if, when the ego vehicle first enters the 2 m zone ($0 <$ ego_to_cp $<= 2$) around the CP, the cyclist is within 10 m of the CP ($0 <$ bike_to_cp $<= 10$); otherwise, it is labeled as no_risk. This is referred to as the 2 m / 10 m rule.

Sequences are trimmed to the ego_to_cp distance window of [25, -4] to ensure consistency. The 25 m mark is the starting point for analyzing right-turn-with-cyclist scenarios, and -4 m marks when the ego vehicle has crossed the CP. Rows outside this range are excluded, as the focus is on assessing risk near the CP. The manual rule examines the first row in which the ego enters the 2 m zone and checks whether the bicyclist is within 10 m of the CP at that timestep. This approach uses a single-row check, as the moment the ego enters the conflict geometry is the key point of interest.



(*a*) Risk scenario (*b*) No risk scenario

Figure 3: The manual classification rule for the right turn cyclist conflict.

Figure 3 illustrates two example scenarios for the rule-based classification. In both cases, the ego enters the 2 m zone at a distance of 1.5 m from the CP. In the left example, the cyclist is 8 m from the CP at the same timestep, meeting the manual rule and classifying the scenario as risk. In the right example, the cyclist is 15 m away, so the scenario is classified as no_risk. The 2 m / 10 m thresholds were selected for specific reasons. The 2 m ego threshold accounts for typical car dimensions, ensuring the vehicle's bumper is considered at the conflict point. The 10 m threshold reflects the typical stopping distance for cyclists before crossing in urban lanes. Further analysis of this rule-based classification is provided in Section 4.

3.4. Threshold Selection: Evaluation Protocol

The BN produces a collision-risk probability at each timestep of a sequence. To convert these continuous probabilities into binary risk/no_risk decisions, an operating threshold is required. For each sequence, the maximum risk probability within the trimmed *ego_to_cp* distance window of [25, -4] m is selected and compared with this threshold. Sequences with a maximum probability above the threshold are classified as risk, while the remaining sequences are classified as no_risk.

The operating threshold is determined using Youden's J method. Performance is estimated using 100 iterations of stratified 70/30 train/test splitting, where stratification preserves the proportion of risk and no_risk sequences. In each iteration, the threshold is selected on the training set by maximizing Youden's J statistic, ($J = sensitivity + specificity - 1$), on the receiver operating characteristic (ROC) curve. The selected threshold is then applied to the corresponding test set, and the confusion matrix, accuracy, precision, recall, and F1 score are recorded.

Section 4 presents two evaluations. The Within-VTI evaluation approach uses the VTI dataset for stratified train/test splitting, operating-threshold selection, classification, and prediction-horizon analysis. The VTI-to-TUM evaluation approach determines the operating threshold using the complete VTI dataset and applies it, without re-tuning, to the TUM dataset.

4. Experimental Results

4.1. Manual Classification Outcomes

Applying the manual rule defined in Section 3.3 to the 260 VTI sequences resulted in 37 risk and 223 no_risk sequences. This imbalance reflects a structural feature of the right-turn-with-cyclist scenario, in which most sequences ended without a near-miss in the simulator. To address this and reduce the impact of major no_risk scenarios, stratified sampling with Youden's J rule is used in this work as an evaluation protocol.

For the TUM dataset, the manual rule classified all 127 sequences as no_risk. This outcome is consistent with the different TUM experimental design described in Section 2.2, where the cyclist appeared ahead of the ego vehicle and typically passed the CP before the ego vehicle entered the 2 m zone.

4.2. Risk-Probability Distributions

The BN's risk probability assessment for a single sequence in VTI data is illustrated before the overall manual rule-based classification applied to 260 sequences in VTI data and 127 sequences in the TUM dataset is presented.

For the selected sequence, the risk probability for each time frame is plotted alongside the underlying ego and bike kinematics in Figure 4. The probability of risk increases as the ego and bike approach CP and the ego's speed decreases. The risk probability exceeds the operating threshold (0.226) and reaches a maximum value of approximately 0.70 for this sequence when both the ego and the bike are near the CP simultaneously. The BN model produces continuous-valued risk probabilities, enabling the risk signal to be computed or thresholded at any time.

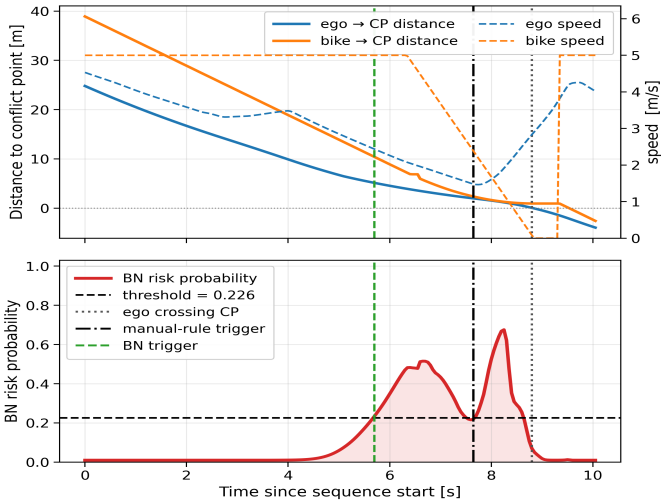
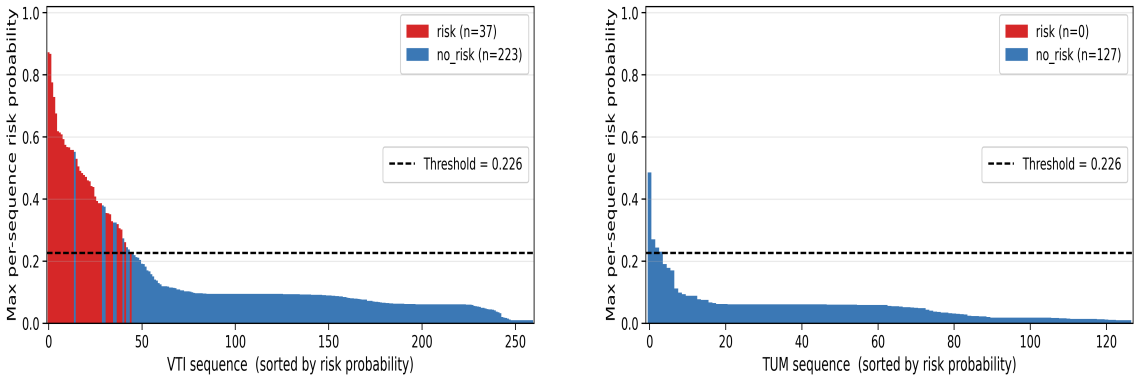


Figure 4: Example BN risk estimate for ego–cyclist kinematics.

The operating threshold shown in Figure 5 is determined using Youden's J method as described in Section 3.4. Sequences classified as risk are shown in red and generally have probability values above the threshold. No_risk sequences are shown in blue, with probabilities below the threshold. However, there is a slight overlap, as some no_risk sequences exceed the threshold. This overlap indicates the presence of True Negatives (TN) and False Positives (FP), which are discussed below.



(*a*) VTI sub-scenarios (*b*) TUM sub-scenarios

Figure 5: Maximum BN risk probability comparison for VTI and TUM sub-scenarios.

Figure 5(*a*) shows the maximum risk probability for each sequence in the VTI dataset, sorted in descending order and classified using the 2 m / 10 m manual rule. Similar to Figure 5(*a*), Figure 5(*b*) presents the maximum risk probability for each sequence in TUM, arranged in descending order. When classified using the manual 2 m / 10 m rule, the TUM data is largely dominated by the no_risk class. Since all TUM sequences are labeled as

no_risk under the manual rule, the distribution primarily reflects the behavior of the BN risk estimate on the independent TUM dataset.

4.3. Threshold Evaluation on VTI

This section presents the results of the Within-VTI evaluation described in Section 3.4. Table 2 lists the accuracy, recall, precision, and F1 score for Within-VTI approach. High recall and F1 scores indicate that the manually classified sequences were accurately recovered. The low standard deviation further demonstrates the stability of the selected operating threshold.

Table 2: Within-VTI threshold evaluation: 100 stratified 70/30 cross-validation iterations.

Metrics	Mean ± Std
Accuracy	0.969 ± 0.016
Precision	0.847 ± 0.075
Recall	0.963 ± 0.067
F1 score	0.898 ± 0.051

Manual rule-based classification may introduce implicit threshold fitting. To address this, all sequences are re-labeled using various combinations, as detailed in Tables 3 and 4. For each combination, quantitative assessments are performed. In Tables 3, bike_to_cp is fixed at 10 m while ego_to_cp varies from 0.5 m to 5 m. F1 scores are measured for each combination after 100 iterations of stratified cross-validation. As ego_to_cp increases, the F1 score decreases because the decision horizon moves further from the CP. For VTI data, the risk count drops, indicating the bike often does not reach the 10 m zone when ego_to_cp is set high. For the TUM dataset, the number of sequences classified as risk increases as the ego-distance threshold increases. A larger ego threshold evaluates the interaction earlier, when the cyclist, positioned ahead of the ego vehicle in the TUM experiment, is more likely to remain close to the CP.

Table 3: Case 1: bike_to_cp threshold fixed at 10 m, ego_to_cp threshold varied.

Ego threshold [m]	**VTI risk count**	**VTI F1**	**TUM risk count**
0.5	40	0.832 ± 0.051	0
1	41	0.858 ± 0.050	0
2	**37**	$\mathbf{0.877 \pm 0.048}$	**0**
3	28	0.867 ± 0.080	1
4	21	0.612 ± 0.132	3
5	25	0.238 ± 0.092	12

To ensure fairness, the next analysis fixes ego_to_cp at 2 m and varies bike_to_cp from 6 m to 14 m. Table 4 presents the risk counts and F1 scores. As the risk horizon increases, both risk count and F1 score rise, capturing more sequences as risk. For TUM, risk counts remain zero because the bike has already passed the CP safely when ego_to_cp is 2 m. Tables 3 and 4 show that the 2 m / 10 m combination of ego_to_cp and bike_to_cp provides balanced F1 scores and results compared to other combinations.

Table 4: Case 2: ego_to_cp threshold fixed at 2 m, bike_to_cp threshold varied.

Bike threshold [m]	VTI risk count	VTI F1	TUM risk count
6	19	0.667 ± 0.081	0
8	33	0.866 ± 0.067	0
10	**37**	$\mathbf{0.877 \pm 0.048}$	**0**
12	41	0.823 ± 0.051	0
14	49	0.758 ± 0.073	0

4.4. Prediction-Horizon Analysis

The aim of developing a risk analysis model is not only to classify sequences correctly but also to flag future conflicts as early as possible. This section defines the prediction horizon for evaluating the BN model's performance and for comparing it with manual rule-based decisions. The prediction horizon is calculated as the signed difference between the time the ego first enters the 2 m zone (as defined by the manual rule) and the time when the BN risk probability exceeds the operating threshold. Risk detection is considered positive if it occurs earlier than the manual rule's decision frame, and negative if the BN prediction is delayed. These results are recorded for 37 sequences identified as risks in the VTI data, and the corresponding prediction horizon is shown in Figure 6.

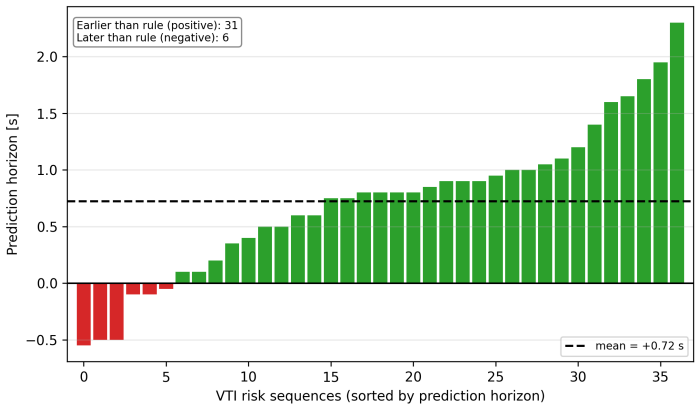


Figure 6: Prediction Horizon for each risk sequence.

In Figure 6, the BN model flags 31 out of 37 sequences as risks earlier than the baseline, with an average lead time of approximately 0.72 seconds. This demonstrates the model's robustness in predicting future risk. Of the remaining sequences, the BN model flags 6 later than the 2 m / 10 m manual rule. The slower predictions occur due to the rapid deceleration of the ego vehicle near CP, which causes the rule-based approach to detect risk earlier.

4.5. VTI-to-TUM Evaluation

The VTI-derived operating threshold is evaluated on the independent TUM dataset. Using all VTI sequences, Youden's J gives a threshold of 0.226, which is applied to the 127 TUM sequences without re-tuning. Four sequences exceed the threshold. Because all TUM

sequences are labeled as no_risk by the manual 2 m / 10 m rule, sensitivity, recall, precision, and F1 score cannot be evaluated on this dataset. Therefore, they are not considered stable for assessment and are excluded from Table 5.

Table 5: VTI-to-TUM evaluation using the VTI-derived threshold without re-tuning.

Dataset	n	Threshold	TP	FP	TN	FN	Accuracy	FAR	Specificity
VTI (threshold fit)	260	0.226	37	8	215	0	0.969	0.036	0.964
TUM (transfer)	127	0.226	0	4	123	0	0.969	0.031	0.969

Because only a small number of TUM sub-scenarios cross the operating threshold, accuracy and specificity are high on both datasets, as shown in Table 5. On TUM, the VTI-derived threshold results in 4 BN-positive/manual-negative cases and 123 true negatives, corresponding to a false-alarm rate of 0.031. Since the TUM dataset contains no positive cases, these results characterize false-alarm behavior.

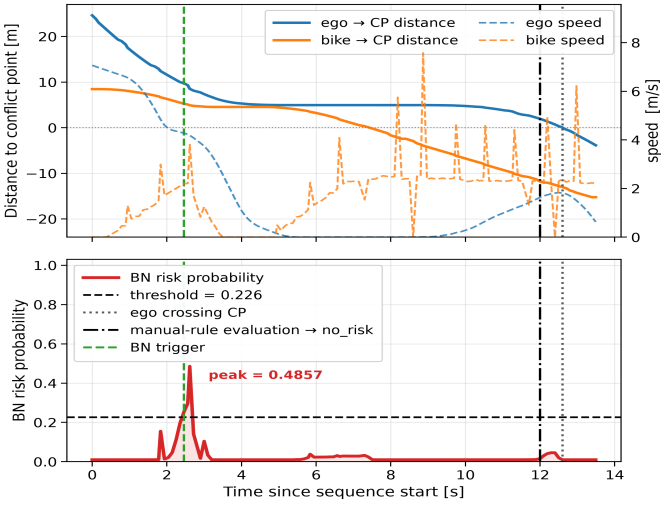


Figure 7: Example of BN-positive/manual-negative TUM sequence.

To further examine the four FPs (i.e., BN-positive/manual-negative cases), Figure 7 illustrates one representative TUM sequence. The BN exceeds the threshold early in the interaction, while both road users are still approaching the CP, as it jointly considers multiple kinematic variables, including distance, speed, and acceleration. These variables together result in a risk probability that reaches a maximum of 0.486. In contrast, the manual rule is evaluated later at a single point, after the cyclist has passed the CP, and therefore labels the sequence as no_risk. This illustrates the different temporal characteristics of the continuous BN assessment and the single-point manual rule.

5. Discussion

A strength of this study is that the manual rule-based classification and the BN's risk probability calculations are independent, which makes their risk assessment comparison

more reliable. The BN does not rely on the binary classification from the manual rules, supporting the value of the BN model. The primary difference between these methods is how they assess each sequence. The manual rule-based approach checks a single moment to determine if the bicyclist is close when the ego vehicle first enters the conflict zone. In contrast, the BN evaluates all relevant kinematic data, such as speed, acceleration, and the distance between vehicles throughout the entire sequence, assigning risk probabilities at each timestep. This comprehensive approach enables a more thorough assessment of risk over the full motion.

Within-VTI evaluation showed that the BN identifies risk cases with high agreement with the manual classification and predicts risk on average 0.73 s before the manual rule. The VTI-derived threshold was subsequently applied without re-tuning to the TUM dataset. However, because TUM contains no positive cases according to the manual rule, this evaluation cannot assess sensitivity and should instead be interpreted as an assessment of false-alarm behavior on an independent dataset. The four BN-positive/manual-negative cases illustrate situations where the continuous BN assessment and the single-point manual rule can disagree. Additionally, both datasets are from simulators, where drivers are aware that there is no real danger, which may lead to calmer braking and reactions than in real traffic. Future work should expand the analysis to more diverse datasets and conflict scenarios, including additional cyclist–vehicle conflict scenarios.

6. Conclusion

This paper presented a knowledge-driven Bayesian network for collision-risk prediction in an urban right-turn cyclist scenario. The proposed model combines driver characteristics, vehicle dynamics, cyclist behaviour, and conflict-point-based temporal indicators within an interpretable probabilistic framework. Unlike deterministic rule-based approaches, the BN explicitly represents uncertainty and supports probabilistic reasoning throughout the evolution of a traffic situation. The model was evaluated using two independent driving simulator datasets collected as part of the Horizon Europe *i4Driving* project. Experimental results on the VTI dataset show strong agreement with the manually defined risk classifications. In addition, the BN predicts hazardous situations before the corresponding rule-based trigger in most cases, providing an average prediction horizon of approximately 0.73 s. This additional warning time can be valuable for driver assistance systems and future CCAM applications. Applying the VTI-derived threshold to the TUM dataset yielded a low false-alarm rate, although the absence of positive TUM cases prevents evaluation of sensitivity and thus limits conclusions about cross-dataset generalization. The interpretable structure of the BN enables transparent reasoning about the influence of driver behaviour, vehicle dynamics, and cyclist motion on collision risk.

Future work will focus on extending the model to a Dynamic Bayesian Network in order to explicitly capture temporal dependencies between consecutive observations. Additional directions include incorporating driver monitoring information such as gaze behaviour, attention allocation, and workload, as well as evaluating the approach on a broader range of traffic scenarios and real-world driving data. Such extensions may further improve early risk prediction and strengthen the role of probabilistic graphical models in the safety assessment of future CCAM systems.

Acknowledgments

This study has been partially funded by the European Commission Horizon Europe project i4Driving (Grant Agreement No. 101076165). The authors would also like to thank Maya Sekeran and Tanja Niels from the Technical University of Munich (TUM) for their support in the design, execution, and documentation of the TUM driving simulator experiments, which contributed to the data used in this study.

References

Christer Ahlström, Katja Kircher, Fredrik Johansson, Anders Andersson, and Johan Olstam. Driver attention in urban intersections when crossing paths with cyclists. *Accident Analysis & Prevention*, 223:108276, 2025. ISSN 0001-4575. doi: https://doi.org/10.1016/j.aap.2025.108276.

Y. Ali, S. Washington, and M. M. Haque. Estimating real-time crash risk at signalized intersections using surrogate safety measures and bayesian extreme value models. *Safety Science*, 164:106181, 2023.

A. Arun, M. M. Haque, and A. Bhaskar. Real-time conflict-based bayesian tobit models for safety evaluation of signalized intersections. *Accident Analysis & Prevention*, 2021.

H. Briwa, A. L. Madsen, and M. C. Leva. Predicting operators reliability for control room alarm management using knowledge-based bayesian networks. *Reliability Engineering & System Safety*, 273:112261, 2026. doi: 10.1016/j.ress.2026.112261.

H. Gao, H. Su, Y. Cai, et al. Trajectory prediction of cyclists based on dynamic bayesian networks and long short-term memory models at urban intersections. *Science China Information Sciences*, 64(12), 2021.

F. V. Jensen, S. L. Lauritzen, and K. G. Olesen. Bayesian updating in causal probabilistic networks by local computations. *Computational Statistics Quarterly*, 4:269–282, 1990.

Katja Kircher, Christer Ahlström, Fredrik Johansson, Anders Andersson, and Johan Olstam. Drivers' neglect of cyclists in right-turns—a matter of rule knowledge rather than a lack of attention. *Accident Analysis & Prevention*, 236:108709, 2026.

U. B. Kjærulff and A. L. Madsen. *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*. Springer, 2nd edition, 2013.

Steffen L Lauritzen and Frank Jensen. Stable local computation with conditional gaussian distributions. *Statistics and Computing*, 11(2):191–203, 2001.

M. C. Leva and P. F. Hansen. Use of Bayesian network for human reliability modelling: Possible benefits and an example of application. *Human Factors and Reliability Engineering for Safety and Security in Critical Infrastructures*, pages 189–223, 2017.

D. Li, C. Fu, T. Sayed, and W. Wang. An integrated machine learning and bayesian spatial modeling framework for real-time traffic conflict prediction. *Accident Analysis & Prevention*, 2023.

A.L. Madsen, F. Jensen, U.B. Kjærulff, and M. Lang. HUGIN - The Tool for Bayesian Networks and Influence Diagrams. *International Journal on Artificial Intelligence Tools 14*, 3:507–543, 2005.

J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, CA, 1988.

S. Sandeep, V. Prasannakumar, et al. Conflict-based models for real-time crash risk assessment: A review. *Vehicles*, 6(3), 2026.

G. Weidl, A. L. Madsen, and S. Israelson. Applications of object-oriented bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes. *Computers & chemical engineering*, 29(9):1996–2009, 2005.

Galia Weidl, Gabi Breuel, and Virat Singhal. Collision risk prediction and warning at road intersections using an object oriented bayesian network. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pages 270–277, 2013.