Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2022) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

Transport Research Arena (TRA) Conference

Individual Motorcycling Safety: Creating a safety profile from riding data

Andreas Hula[a,*], Klemens Schwieger[a], Peter Saleh[a]

a Center for Low-Emission Transport, AIT, Austrian Institute of Technology, Giefinggasse 2, 1210 Vienna, Austria

Abstract

Motorcycle riders are vulnerable road users, who suffer fatal accident outcomes at significantly higher rates than car drivers. Lifesaving assistance systems are considerably harder to integrate in the operation of a motorcycle, compared to the operation of a more predictably moving car. Here we present a methodology of estimating an individual rider's classifier of "risky" dynamics from their riding behaviour on several popular motorcycling routes in an experimental set up. Using clustering of common motions and data obtained at known accident sites, as well as an updating regime for the model fit to the individual rider, we are able to identify potential riding risks in an online methodology and determine the driving factors (i.e. the most relevant dynamics for a motion to be classified as risky) in the risk estimate, to potentially base interventions on.

Peer-review under responsibility of the scientific committee of the Transport Research Arena (TRA) Conference
Keywords: Vulnerable Road Users; ADAS; Motorcycle; Risk Model; Vehicle Driving Dynamics; Clustering

1. Introduction

Motorcycling has demonstrated unbroken popularity during the pandemic years and the number of motorcyclists on Europa's roads does not indicate any tendencies for this to end soon. However, motorcycling safety continues to be a challenge (see also EC Report (2020)), compared to cars in particular, as advanced driving assistance systems (ADAS) for motorcyclists are considerably more difficult to design and implement (see Savino et al. (2020)), given the more sensitive control of a motorcycle and the more severe consequences of an accident. Development of such systems is also hampered by the need of potential ADAS to avoid distracting the rider, since this might aggravate any existing critical situation further.

* Corresponding author: Andreas Hula Tel.: +43 664 88390049
E-mail address: andreas.hula@ait.ac.at

2352-1465
Peer-review under responsibility of the scientific committee of the Transport Research Arena (TRA) Conference

Approaches that have been investigated include the effectiveness of helmets (see Tabary et al. (2021)), physical safety gear (for instance Afquir et al. (2020)), risk maps (see Ryder et al. (2016, 2017) for earlier work) and vehicle improvements (see Xiao et al. (2020)). Infrastructure measures of various kinds have been implemented or are being investigated (see Milling et al. (2016)). Yet clearly the share of serious or fatal injuries in motorcyclist accidents is still challenging to reduce.

Our approach to enabling further developments of ADAS for motorcyclists rests on analysing the behaviour of the human rider and comparing the dynamic variables (accelerations and angle velocities/rates) at known risk spots to the values of the dynamic variables displayed most commonly by a given rider. This was made possible by employing a motorcycle equipped with sophisticated sensor systems (dubbed "Motorcycle Probe Vehicle", or "MoProVe", see Ecker and Saleh (2016)) that allows to collect time series of all the variables of interest during test rides. This data was transformed to per-meter values and the variables of each rider were combined in a statistical model (with clusters obtained from the data to represent common motions), to produce a risk estimate on each meter in the data set.

Previous work employed a similar approach to combine risk estimates of multiple riders into a single risk estimate and transformed the joint estimates of the driven tracks into risk maps (see Hula et al. (2019); Schwieger et al. (2020); Hula et al. (2021)). This was done with a focus on including as many as possible known accident sites in as limited as possible risk areas, using estimates across all the obtained data sets. The estimates obtained this way could therefore not have been available to assist riders online or direct/trigger an ADAS.

The current approach presented in this paper instead focuses on fitting a model providing an online risk estimate to the individual rider. We retain several features of the original methodology (clustering of dynamics variables, separation functions to the known accident locations), but instead present an approach to provide risk estimates at each point in time, using only the information obtained during previous rides. We update the clusters of driving dynamics continually during the ride, to potentially reflect changes in driving behaviour and fit new separation models at regular intervals (meaning, in this early study, that we fit new models between designated rides). Using the resulting online estimate, we decompose the factors contributing to a risk warning occurring, to enable an investigation into why a given risk warning occurred. We find that causal attribution of the risk warnings (i.e. which dynamic variables contributed the most to a given risk warning) can be summarized into relatively few patterns. These patterns can be interpreted as certain risk scenarios and we discuss what possible interventions by ADAS could contribute to improving rider control in this case and thus turn a risky situation into a non-risky one.

Thus, this study presents a computational approach to online risk warnings for motorcyclists based on individual factors (clusters of the riding dynamics, individual dynamics at known accident spots) and provides a model of how an on-board system could repeatedly adjust to the rider's behaviour. Such a system could keep learning without sharing data with remote systems or be an element for the systematic study or rider behavioural types and learning in riders.

2. Methods

2.1. Test Vehicle - MoProVe

The Motorcycle Probe Vehicle used in this study was a KTM 1290 Super Adventure (generously provided by KTM, KTM (2022)) with access to CAN-Bus data (ISO (2015)) and several additional measurement and data logging systems (see VBOX (2022); Debus (2022)). Several Geopositioning systems (GPS, GLONASS) were available, to improve the quality of localisation data during the ride. More details on the measurement setup can be found in Schwieger et al. (2018).

2.2. Data Collection and Preparation

We recruited 5 riders to carry out multiple data collection rides on several popular motorcycling tracks in Austria. All riders were experienced motorcyclists and gave informed consent at the start of the study. We collected data using our motorcycle probe vehicle (MoProVe, see Ecker and Saleh (2016)) on 6 popular motorcycling tracks. In total we had 2121 kilometers of dynamic variables data available for analysis. The number of valid data collection rides per track and rider can be seen in Table 1.

Table 1. Number of included rides per rider on each of the 6 tracks of interest.

Rider	Track 1	Track 2	Track 3	Track 4	Track 5	Track 6
Rider 1	10	0	8	12	0	4
Rider 2	8	6	11	12	10	4
Rider 3	7	10	10	12	4	4
Rider 4	6	4	0	0	10	6
Rider 5	0	13	0	0	6	0

Accident data on the 6 studied tracks was obtained from Statistik Austria (Statistik Austria (2022)) for the years 2012 to 2015 and single rider motorcycling accidents as well as collisions with oncoming vehicles (i.e. motorcyclists potentially leaning into the opposite road side) were included as known accident locations (denoted Acc) in the analysis. The collected data was transformed to per-meter-data by partners at TU-Vienna in an earlier project. Indicators to document when riders were following another vehicle (thus limiting the quality of dynamic variables data) were provided by TU-Vienna. Data on which riders followed another vehicle or data for which less than 3 satellites were shown as available, were excluded from the estimation (clustering and separation procedures). Data was smoothed using a rolling average of 5 meters for this analysis.

2.3. Model Components

The flow of processing steps used in this work is displayed in Fig. 1.

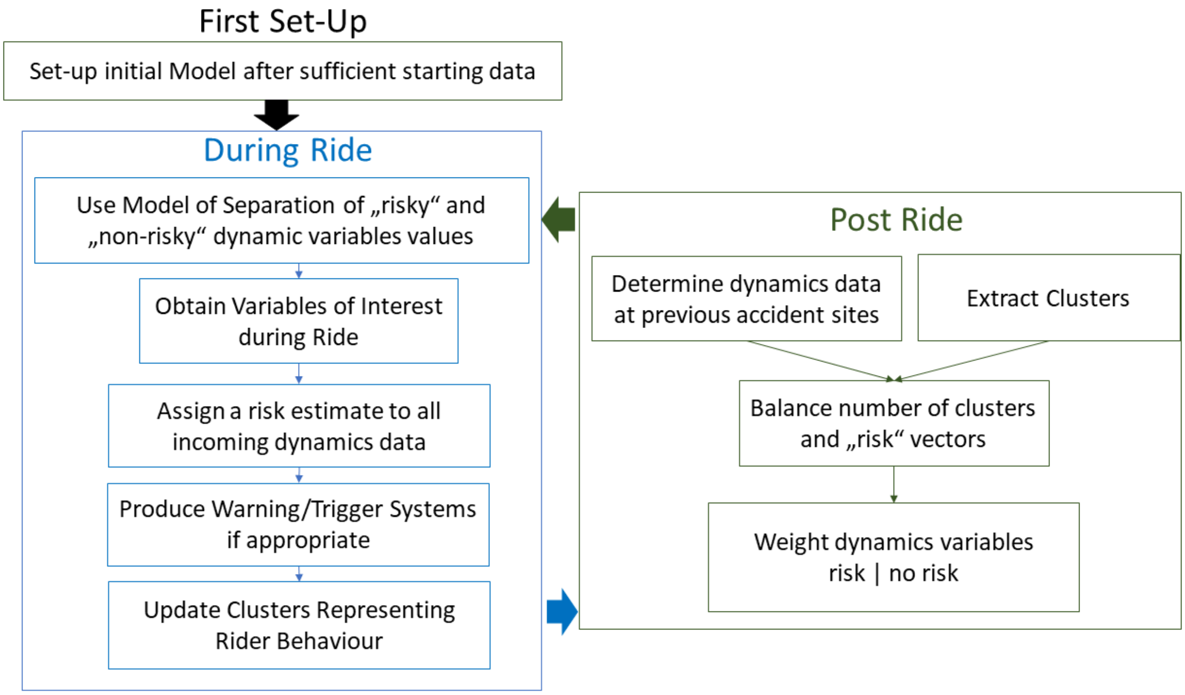

Fig. 1. Methodology flow for the updating risk estimate. After an initial data collection phase (7 rides to get enough clusters to fit a model with 24 parameters), the method produces risk estimates online during the ride, enabling an interaction with assistance systems and updates cluster centres online. After a defined data collection period, it also adds clusters, data collected near accident spots and determines a new separation model, which can then be using during the next rides again.

We review the components of the model, which are based an earlier published model Hula et al. (2021), but go beyond it in several aspects. The variables used are reviewed in Table 2.

Table 2. Variables used in the risk model.

Variable Name - Symbol	Dimension
Yaw-Rate - j	degrees/s
Pitch-Rate - p	degrees/s
Roll-Rate - r	degrees/s
x-acceleration - x	m/s2
y-acceleration - y	m/s2
z-acceleration - z	m/s2

For each variable V we used positive $V_+ = max(V, 0)$ and negative values $V_- = max(-V, 0)$ separately, as well as first differences (m denoting a meter on the track here) $dV(m) = |V(m) - V(m-1)|$ (we are interested in the amount of change, rather than the sign, for the higher derivatives), to form a full data structure, consisting of $(V+, V-, dV+, dV-)$ for each variable V. Thus, our data structure DS consisted of a large matrix with 24 columns (4 quantities derived from each of the 6 variables) per driver. We show this structure in equation (1).

$$DS = (j_+, p_+, r_+, x_+, y_+, z_+, j_-, p_-, r_-, x_-, y_-, z_-, \ldots \quad dj_+, dp_+, dr_+, dx_+, dy_+, dz_+, dj_-, dp_-, dr_-, dx_-, dy_-, dz_-). \tag{1}$$

Clustering (k-means clustering, see Bishop (2006); McLachlan (1988); Duda et al. (2001)) was applied to each ride in the data set to obtain 5 clusters each and valid dynamic variables data at a previous accident location was used to fit the separation model, with the cluster centres denoted as cen. The linear model (see Bishop (2006); Draper and Smith (1998)) for the risk value L to separate cluster centres and data at accident spots is described by the model formula in equation (2).

$$L = \sum_V a_V V_+ + \sum_V b_V V_- + \sum_V c_V dV_+ + \sum_V d_V dV_- + \varepsilon \tag{2}$$

with coefficients $a_V, b_V, c_V, d_V \in \mathbb{R}$ and a normally distributed error term ε. A reference value was derived from the mean value of the separation function at the known accident spots $\lambda = mean(L(Acc))$. Model values at a given meter m that were above λ i.e. $(L(m) > \lambda)$ were defined as “risky”, those below as “non-risky”.

All implementations were carried out in R (R Core Team (2022)), using in particular the underlying “stats” package and the package “zoo” (see Zeileis and Grothendieck (2005)).

3. Results

Our main result is the implementation of the online risk estimation with constant updating and an investigation of causes for risk warnings to occur and their interpretation.

3.1. The Updating Methodology

We employ k-means clustering to the dynamic variables data to obtain, as representatives of the most common motions of a rider, the means of the clusters (5 clusters per ride). Clusters were also subjected to an ongoing updating mechanism of their cluster centres: each incoming data point was associated to the nearest cluster centre (in the Euclidean distance) and the cluster centre component values, denoted as cen, were updated according to equation (3).

$$cen_{ds}^{i+1} = \frac{i}{i+1} cen_{ds}^{i} + \frac{1}{i+1} ds \tag{3}$$

where i denotes the number of elements in the respective cluster observed before the update and $ds \in DS$denotes a *component* of the newly observed data point, associated with this data point. Following every ride, the updated clusters and the accident locations on which valid data was obtained, were used to calculate a new separation between cluster centres and dynamic variables at known accident locations.

Since the set of passed accident locations and the set of cluster centres can vary considerably in size in an online approach (in particular, some tracks have only few accident locations, some a lot more than the 5 clusters included per ride) we employed a rebalancing mechanism, to ensure the two sets used in the separation model were of comparable size. If the set of cluster centres contained cn elements and the set of accident locations contained ln elements, then if $cn > ln$ we used $k_{cn} = cn/ln$ rounded to an entire number and put the set of accident locations k_{cn} times repeated into data for the regression step to ensure risky and nonrisky dynamics would receive approximately the same weight when calculating a separation. Conversely, if $ln > cn$ we used a similar factor to increase the weight on cluster centres.

The resulting model was used to determine a new separation value and risky estimates were produced according to equation 3 above.

The implementation on a Intel(R) Core(TM) i7-7600U CPU running at 2.80GHz was able to evaluate the risk estimate at a worst case 0.08 seconds, so that an online provision of risk estimates/real time evaluation would be possible, given that the implementation could still be subjected to a number of optimizations. Updating clusters for each ride and fit a new risk model takes at most 2 minutes and 20 seconds, which is not a problem as this would be supposed to happen between rides.

3.2. Responsibilities for Risk Warnings

If we let ds be a component (column value) of any element in our data structure DS and r_{ds} its model coefficient in equation 2, then, since the employed model is linear, we can directly associate responsibilities R_{ds} for each data component at a “risky” spot according to equation (4).

$$R_{ds} = \frac{r_{ds}ds}{\sum_V (a_V V_+)_+ + \sum_V (b_V V_-)_+ + \sum_V (c_V dV_+)_+ + \sum_V (d_V dV_-)_+}. \tag{4}$$

In other words, we use the positive contributions to the risk estimate and normalize them according to their size. For this study we find a first characterisation of risk types, classified by which component is dominant in the responsibilities (see Fig. 2 for examples at previous accident sites).

Risk type 1 is characterized by a largest contribution of the jaw-rate j i.e. (j_+, j_-, dj_+, dj_-). Risk type 2 is characterized by a largest contribution of the pitch-rate p, specifically (p_+, p_-). If the largest contribution stems from the roll-rate r, namely (r_+, r_-) this signifies the risk type 3. Finally, risk type 4 stems from a dominant y-acceleration y i.e. (y_+, y_-, dy_+, dy_-). It can be associated to a substantial y-drift during a curve. Results on the share of risky meters attributable to these types among all ridden meters per driver can be found in Table 3.

Table 3. Share (Sum to 1) of Risk Types per Rider

Rider	Risk Type 1	Risk Type 2	Risk Type 3	Risk Type 4	Other
Rider 1	0.71	0.07	0.01	0.19	0.02
Rider 2	0.73	0.07	0.0	0.16	0.04
Rider 3	0.58	0.09	0.0	0.29	0.03
Rider 4	0.7	0.14	0.0	0.12	0.04
Rider 5	0.54	0.22	0.15	0.09	0

3.3. Recommendations

We note that of the typical risk profiles we identified, the biggest shares can be associated to sharp left or right turns, signified in the yaw-rate. This is not surprising, as the vast majority of single vehicle accidents for motorcyclists happen in curves. However, the advantage of our approach is the insight into which components are currently responsible for heightened risk values and making this information available to ADAS consequently. We also note that the limit for “risky” dynamics will have to be adjusted in the future, since the requirements for risk mapping and system interventions are quite different (in the latter, the “risky” classification should be used more sparingly).

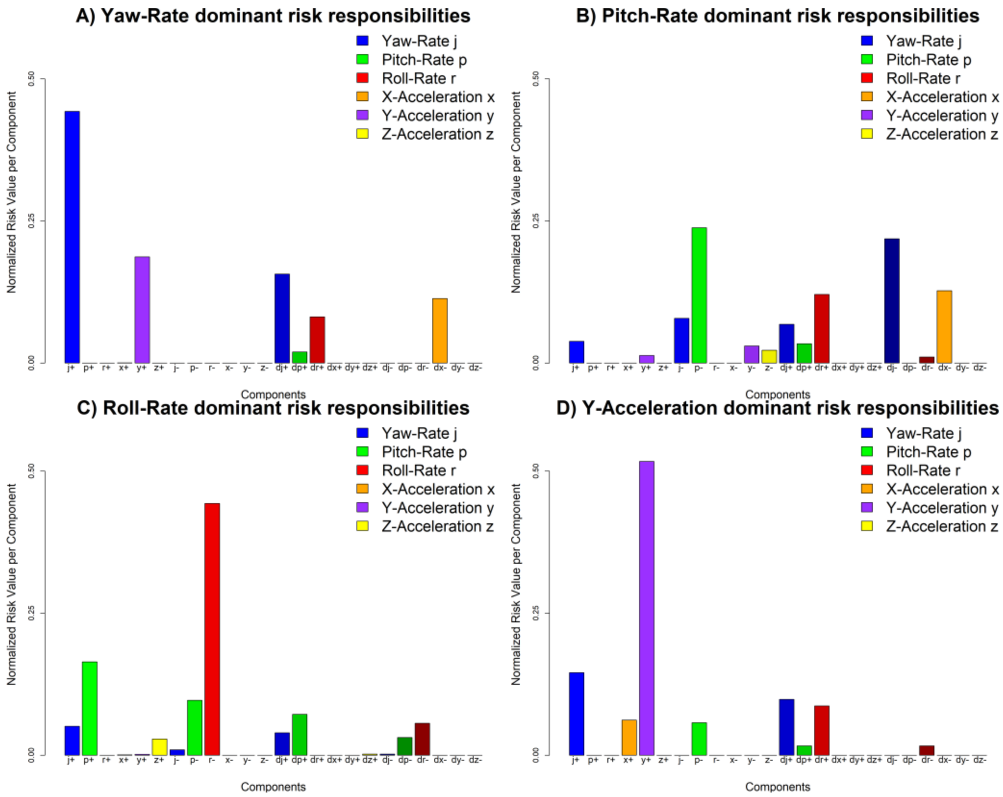

Fig. 2. Examples of risk responsibilities (according to equation 4 applied to all 24 components) for a particular rider. From top left to bottom right: a) Common profile for a curve with j being the dominant contributor. b) Pitch-rate p as most responsible factor. c) Common profile for a curve with r being the dominant responsibility component. d) y-Acceleration y as largest factor for several historical accident spots.

An intervention that might help with making the yaw-rate based risky situations less dangerous is a yaw-rate and risk value guided traction control, which thus might trigger in relation to how the driver is typically riding curves. In particular, sudden jumps in the risk estimate or values close to the “risky” λ might indicate the need for the system to act. Such an individual trigger would profit from the updating methodology in that it would take into account developing changes in how the individual driver is behaving.

Major contributions by the pitch rate can be seen in several profiles. This might indicate the need to “soften” the motions of the motorcycle i.e. increase motility of the wheels in the direction of the z-Axis. Also, this type of warning might be used to find instances of road infrastructure in need of repair i.e. the warning could be used to guide and support road safety inspections.

The risk profiles showing a substantial effect in the rollrate can often be associated to serpentine curves, as well as the risk of leaning into oncoming traffic. An intervention into the driving behaviour in this situation is difficult to suggest, however it might be helpful to prepare systems for stabilizing and braking in the event of a sudden return to the upright position (i.e. a last second attempt to dodge), as well as a risk value based stabilisation (i.e. an individualized motorcycle stability control).

Finally, if the main contribution to the risk value stems from y-acceleration, this might indicate skidding rather than a sharp turn. Activation of traction control systems based on y_+ might support the rider’s control in such situations.

4. Discussion

4.1. Summary and Outlook

We devised an updating methodology to continuously fit a model of separating risky from non-risky dynamics, based on individual motorcycle riders features (clusters of driving dynamics) and dynamic variables at known accident spots. We investigated the factors contributing to a given spot being classified as risky and found 4 most frequent "risky profiles" to base potential interventions on. We discuss recommendations on how to utilize the responsibilities of the different components in the risk warning, to adjust future ADAS to individual driving behaviour.

The employed methodology is based on earlier work which focused on forming risk maps. We note that for further development towards usage in individual warnings in comparison to risk map making, it may make sense to instead assume slightly higher thresholds for "risk" (than the mean we used here), since the similarity in driving dynamics to known accident spots does not always imply immediate danger. In fact, an additional requirement for producing an internal warning could be if the risk value were to change sharply upwards in a short time frame. This could be a way forward here and will be investigated.

A future direction for the frequent updating approach might be, to enhance the analysis of the process of learning and changing behaviour in an individual rider's actions. An interesting extension of this approach could be the assessment of rider capabilities, i.e. the quality of their curve rides and their vehicle control through the repeated clustering and updating. Based on a model of scoring rider experience/control the maximum speed/engine power might be made accessible or kept limited, to ensure rider safety.

We note that in our experimental set-up, we have used the existing separation of test rides to fit separate clusters. In order to employ this method in practice, criteria would have to be defined on when the updating system registers a "concluded ride". A number of approaches could be feasible, in particular defining starts and ending of rides by a cut-off criterion, defining that after a given pause in active riding or a certain number of meters driven, a "new ride" is used for clustering and recalibrating of the model.

4.2. Limitations

Current implementational limitations map out the path for further improvements of the method: The most important limitation is the size of the current data. While we have more than 2000 kilometers driven of data at our disposal, the number of riders (5) and tracks (6) is in need of expansion. We are working to add more data to the analysis soon, using a recently set-up new MoProVe.

The comparison to known accident sites requires a reliable localisation and access to reference accident spots. The feasibility of such an approach for wider use (i.e. for an onboard system to have access to this information and the quality of the onboard localisation relative to it) needs to be explored. Alternatively, risky dynamics could be defined by using a measure of similarity between different driver types, however, this is ongoing work.

With the present method in particular, the set of cluster centres used currently holds quite a number of redundant (i.e. close to pre-existing) clusters. A natural improvement would be to reduce the number of clusters repeatedly after the expansion following a few rides. This would ensure a compact representation of driving manoeuvres and thus an improved separation. A possible means to achieve this could be a threshold for merging clusters based on proximity of their centres, potentially including the estimated variances within the clusters or the use of model selection criteria during the repeated clustering phase.

The separation model employed is a linear separation model, therefore nonlinear models or easy to update approaches with regards to changes in cluster structure need to be explored. We note that between yaw-rate, roll-rate and y-acceleration, various types of confusion can occur i.e. high values of all three might occur in similar situations, making the causal attribution difficult. We will keep exploring the data with various models based on fewer factors, to see if all 6 are in fact necessary to characterize risk spot like dynamics.

The present investigation used per meter data, but the onboard systems would provide values in the time domain only. In principle the approach should be well transferable (since the values themselves would be on a similar scale), but we have yet to investigate this in practice, as the calculation of the differences will require some adjustments.

Acknowledgements

The authors gratefully acknowledge funding of the preceding project by the Austrian road safety fund “VSF” (Verkehrssicherheitsfond) under grant number BMVIT-199.940/0002-IV/IVVS/2016. KTM generously provided the motorcycle used to build the MoProVe system. Collaborators at TU Vienna Horst Ecker and Manfred Neumann prepared the per meter values and annotated the data in the earlier joint project. We also thank Matthias Hahn, Stephan Wittman and Alexander Genser for their help with the earlier project data.

References

European Comission Report (2020). Mobility and transport: Road Safety: Making roads safer for motorcycles and mopeds .

Savino, G., Lot, R., Massaro, M., Rizzi, M., Symeonidis, I., Will, S. and Brown, J. (2020). Active safety systems for powered two-wheelers: A systematic review. Traffic Injury Prevention, 21 , 78–86, doi:10.1080/15389588.2019.1700408., PMID: 31914321.

Tabary, M., Ahmadi, S., Amirzade-Iranaq, M. H., Shojaei, M., Sohrabi Asl, M., Ghodsi, Z., Azarhomayoun, A., Ansari-Moghaddam, A., Atlasi, R., Araghi, F., Shafieian, M., Heydari, S. T., Sharif-Alhoseini, M., O’Reilly, G. and Rahimi-Movaghar, V. (2021). The effectiveness of different types of motorcycle helmets – A scoping review. Accident Analysis and Prevention, 154 , 106065. URL: https://www.sciencedirect.com/science/article/pii/S0001457521000968. doi:https://doi.org/10.1016/j.aap.2021.106065.

Afquir, S., Melot, A., Ndiaye, A., Hammad, E., Martin, J.-L. and Arnoux, P.-J. (2020). Descriptive analysis of the effect of back protector on the prevention of vertebral and thoracolumbar injuries in serious motorcycle accident. Accident Analysis and Prevention, 135 , 105331. URL: https://www.sciencedirect.com/science/article/pii/S000145751930106X. doi:https://doi.org/10.1016/j.aap.2019.105331.

Ryder, B., Gahr, B., Dahlinger, A. (2016). An In-Vehicle Information System Providing Accident Hotspot Warnings. In Twenty-Fourth European Conference on Information Systems (ECIS), Istanbul, Turkey. URL: https://doi.org/10.3929/ethz-b-000121549.

Ryder, B., Gahr, B., Egolf, P., Dahlinger, A., Wortmann, F. (2017). Preventing traffic accidents with in-vehicle decision support systems - The impact of accident hotspot warnings on driver behaviour. Decision Support Systems, 99 , 64 – 74. URL: http://www.sciencedirect.com/science/article/pii/S0167923617300829. doi:https://doi.org/10.1016/j.dss.2017.05.004

Xiao, Z., Wang, L., Mo, F., Lv, X. and Yang, C. (2020). Influences of impact scenarios and vehicle front-end design on head injury risk of motorcyclist. Accident Analysis and Prevention, 145 , 105697. URL: https://www.sciencedirect.com/science/article/pii/ S0001457519301356. doi:https://doi.org/10.1016/j.aap.2020.105697.

Milling, D., Affum, J., Chong, L. and Taylor, S. (2016) Infrastructure improvements to reduce motorcycle casualties. Austroads Ltd., Sydney, Australia, https://www.amda.org.au/images/docs/AP-R515-16_Infrastructure_Improvements_to_Reduce_Motorcycle_ Casualties.pdf

Ecker, H. and Saleh, P. (2016). MoProVe - A Probe Vehicle for Traffic Accident Research, EVU Proceedings 25th Annual Congress, EDIS Editing Centre of University of Zilina, ISBN: 978-80-554-1260-3; S. 29 – 41

Hula, A., Schwieger, K., Saleh, P., Neumann, M., Ecker, H. (2019). Objectifying and Predicting Motorcycle Accident Risk through Riding Dynamics. In Proceedings of the 26th International Technical Conference on the Enhanced Safety of Vehicles (ESV).

Schwieger, K., Hula, A., Saleh, P., Ecker, H., Neumann, M. (2020). Avoiding Motorcycle Accidents by Motorcycle Risk Mapping. In Proceedings of the 13th International Motorcycle Conference Cologne.

Hula, A., Furnsinn, F., Schwieger, K., Saleh, P., Neumann, M., Ecker, H. (2021) Deriving a joint risk estimate from dynamic data collected at¨ motorcycle rides, Accident Analysis and Prevention, 159, 2021, 106297, ISSN 0001-4575, URL: https://doi.org/10.1016/j.aap.2021.106297

KTM Sportmotorcycle GmbH (2022), URL: https://www.ktm.com/en-us/models/travel/ktm-1290-super-adventures2022.html

International Organization for Standardization (2015). ISO 11898-1:2015 Road vehicles — Controller area network (CAN) — Part 1: Data link layer and physical signalling. URL: https://www.iso.org/standard/63648.html

VBOX automotive (2022), VBOX 3i Dual Antenna — 100Hz Vehicle Dynamics Measurement, URL :https://www.vboxautomotive.co.uk/index.php/en/products/data-loggers/vb3i

Debus and Diebold Messsysteme (2022), 2d-datarecording, URL:https://2d-datarecording.com/en/

Schwieger, K., Saleh, P., Hula, A., Ecker, H., Neumann, M. (2018). ViaMotorrad – Can motorcycle safety be measured? In Proceedings of the 12th International Motorcycle Conference Cologne.

Statistik Austria (2022), URL: https://www.statistik.at

Bishop, C. (2006). Pattern Recognition and Machine Learning. Springer Berlin Heidelberg-New York

McLachlan, G.J., (1988), Mixture Models: inference and applications to clustering, Statistics: Textbooks and Monographs, Dekker

Duda, R.O., Hart, P.E. and Stork, D.G. (2001), Pattern Classification, Wiley, New York

Draper, N.R. and Smith, H. (1998) Applied Regression Analysis, Wiley, New York

R: A language and environment for statistical computing. URL: https://www.R-project.org.

Zeileis, A. and Grothendieck, G. (2005). zoo: S3 Infrastructure for Regular and Irregular Time Series. Journal of Statistical Software, 14(6),1-27. doi:10.18637/jss.v014.i06