13th International Motorcycle Conference Cologne
September 2020

Avoiding Motorcycle Accidents by Motorcycle Risk Mapping

Klemens Schwieger, Andreas Hula, Peter Saleh; AIT – Austrian Institute of Technology GmbH
Horst Ecker, Manfred Neumann; TU Wien

Abstract:

The risk of suffering a motorcycle accident is still difficult to quantify given the different types of individual riding styles of motorcyclists. Every year, nearly 100 motorcycle accidents occur on Austrian roads [1], but the rare occurrence of accidents at the same spot makes it difficult to locate risky spots. To tackle these challenges, test riders were gathering data while driving on popular motorcycle routes with a motorcycle equipped with the latest sensor technology in order to collect individual driving dynamics data. The test vehicle used (MoProVe, Motorcycle Probe Vehicle) was a modern KTM motorcycle that is equipped with cameras, geo-position antennas (GPS, GLONASS), inertial measurement units (IMUs) and with access to the motorcycle's internal data acquisition (CAN bus).

In order to achieve the identification of risky spots from driving dynamics, the project viaMotorrad (funded by the Austrian Road Safety Fund - VSF) [2] started in 2015 with the aim of developing tools and methods for improving motorcycle safety in the national road network. The test vehicle MoProVe was developed by the Vienna University of Technology and the AIT Austrian Institute of Technology in order to obtain the necessary driving dynamics data. It was the focus of the project to enable a preventive / predictive approach to motorcycle safety based on the dynamic driving data instead of the conventional, reactive approach.

This work presents a central part of the final project results of viaMotorrad: A method for identifying accident hot spots with risky driving dynamics parameters through machine learning. The developed approach uses a combination of unsupervised and supervised learning methods in machine learning to distinguish the dynamics at known accident sites from the most common driving dynamics of every single driver in a pool of 6 drivers. The individual risk estimates are then combined into a common risk estimate for the pool of test drivers in the experiment. Using the joint estimate, maps of risk spots for 6 popular motorcycle routes in the Austrian Alps were derived. The risk spots on these maps were further divided into three groups depending on how high the individual risk warnings were in their vicinity. Therefore, these maps may include a priority checklist for implementing security measures on popular routes.

1. Introduction

As advanced driver assistance systems (ADAS) in cars get ever more sophisticated, devising approaches to improve motorcycle security remains challenging. The more delicate issue of controlling a motorcycle in complex driving situations is difficult, as interventions into the steering process or other automatic adjustments of the driving process are not necessarily welcomed by motorcycle riders. Hence another approach to this may be to recognize patterns among several drivers' dynamics at known accident spots (e.g. a big data approach) and determine dangerous spots via an analysis of (individual) dynamics, combined into to a (population-wise) estimate of local accident risk spots (see earlier work in [3,4,5]). These estimates can in turn be used to address issues of road infrastructure improvements for the safety of motorcyclists or perhaps inform onboard warning systems of upcoming spots of notable dynamics.

2. Measurement Systems

To be able to carry out the measurements, a motorcycle was provided by KTM [6]. This motorcycle was then equipped with high quality sensors (see Fig. 1), to gather the precise movements (e.g. speeds, angular velocities, accelerations) of the KTM 1290 during the test drives. All measurements were undertaken in normal daytime traffic on the routes of interest. The motorcycle was equipped with additional measurement systems [7,8]. Those were comprised of a separate data logger each, IMUs (Inertial Measurement Unit), other sensors and a connection to the CAN-bus. For a detailed description, see Schwieger et al, 2018 [3].

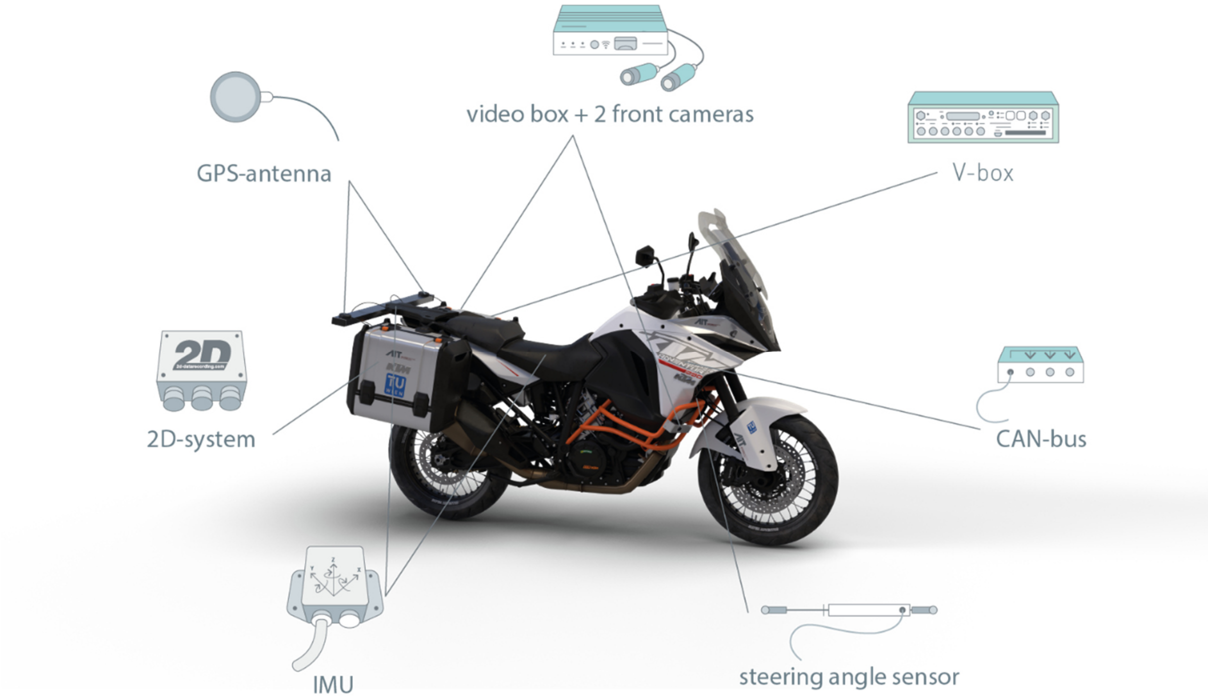

Figure 1: KTM 1290 Super Adventure equipped and instrumented as a Motorcycle Probe Vehicle (MoProVe)

3. Data

Data collection took place on 6 selected tracks with 6 experienced drivers using the probe vehicle to collect data. If a rider tested on a given tack, they would drive several times on that track, in both directions. Data was annotated to record special events (such as following a slow vehicle or overtaking). All rides took place during daylight hours and under generally favourable weather conditions.

Measured time series were postprocessed, to obtain “per meter” values of the observed variables, using values at the start of each track meter and providing a means for spatial comparison between the different rides on the same track.

Accident site data on these 6 tracks for single driver accidents and frontal collisions was obtained for the years 2012 to 2015.

4. Data Preparation

For the purpose of analysis, data was only included when 2 criteria were fulfilled: At least four satellite signals were received and we required our test-riders to not be following another vehicle e.g. to drive freely (there was a separate indicator, that specified when the MoProVe was driving behind another vehicle).

Our analysis was based on several measures of driving dynamics: The vehicle speed, X-, Y- and Z- Accelerations, Roll-Rate, Yaw-Rate and Pitch-Rate, measured by more than one system (IMUs and CAN Bus) for additional information. Where meaningful, we separated the positive and negative values of variables (e.g. left-turn or right-turns of the vehicle, accelerations or decelerations) into separate time series, in order to allow for separated statistical weights in our model.

Dynamics variables were smoothed with a rolling average window, (see Fig. 2 below). We used the function “rollmean” found in the package “zoo” of the statistics language “R” [9, 10] and shifted the resulting series of rollmeans in such a way, that each new value would be an average of only values that happened before it. E.g. we chose a rollmean that was “predictable” in the sense of only using available information at the actual point in time. The window used in this study was 60 meters.

The rolling average serves two purposes: Firstly, compared to raw data, changes happen in a more consistent and steady manner. Secondly, since the value at any given time point is an average of several values that happened before, the value of the rollmean at a given meter encodes a kind of autoregressive approach to studying the vehicle dynamics. This was desirable, as we assume the dynamics leading up to a certain point would be highly relevant in whether an accident might be recorded at the investigated location.

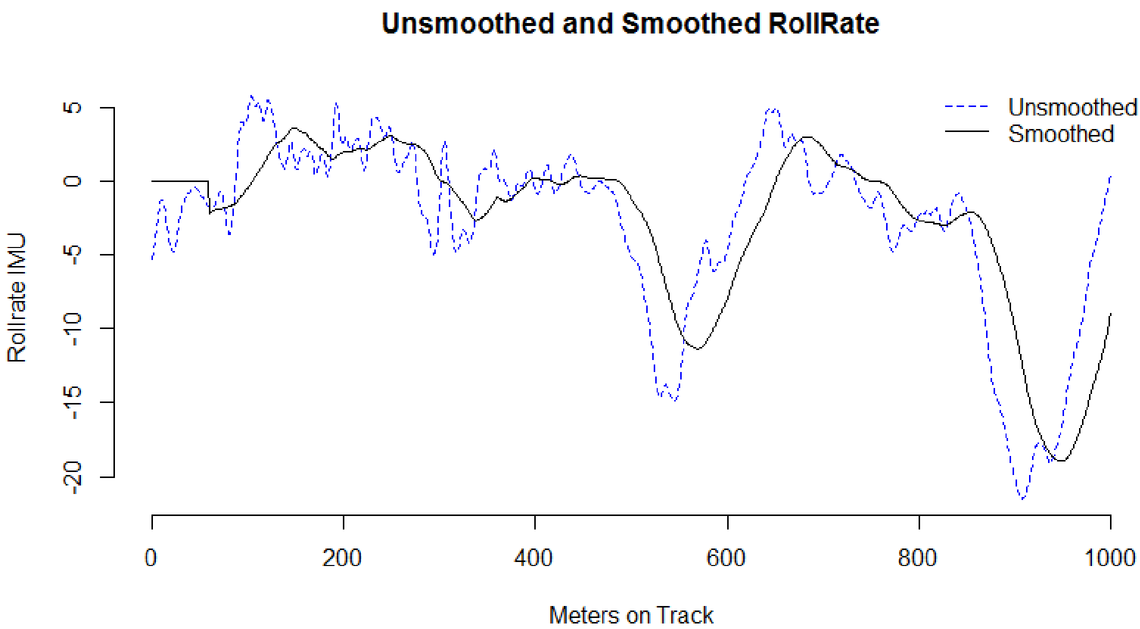

Figure 2: Example of the effect of smoothing and shifting the data (Rollrate in this case) of a single driver. We show only the first 1000 meters of the given ride. The “lag” between the rollmean and the actual data is well visible and this serves to ensure that critical dynamics affect the likelihood of the next few meters after their occurrence still (e.g. the potential recorded accident locations).

Finally, the smoothed curves served as basis to calculate approximate derivatives (namely first differences) between their meter values and those derivatives were smoothed again. In addition to this, we determined the effective curvature of the driven path (defined as the current speed divided by the current Yaw-Rate) as well as accelerations uphill and downhill and treated them as independent variables of interest.

5. Deriving individual Reference Motions from Clustering

A well-established clustering method (k-nearest neighbours, kNN) [11, 12], implemented in the R package “mclust” [13], was used to determine the 8 vectors (cluster centers) of our processed dynamics variables around which the observed data of an individual driver could best be grouped for each single ride, according to the kNN algorithm. These clusters served as the “reference motions” of each driver, assuming them to be non-dangerous, as drivers would on average drive in ways that they could control.
The data of each driver was aggregated into a large time series over all tracks and rides and the dynamics values at each known accident spot were extracted for every single ride.
Finally, we set as target values on the dynamics of the known accident spots the value 1 and for the dynamics of the cluster centers the value 0. Thus, a linear regression was run for each driver, to find the weights (regression coefficients) on the dynamic variables, that would best separate the “reference motions” from the known accident spots (similar to a single layer perceptron [14]). From these separation values, an individual threshold was calculated by taking the average of the linear separation function values on the known accident spots, corresponding to the assumption that the accident spots should be “dangerous on average”.
The weights/coefficients resulting from the linear separation of “reference motions” and accident spots, can then be used to assign a “Dynamical Value” to each meter on the track. These values can then be compared to the individual threshold mentioned above, to determine when the evaluation of the driving dynamics reach the accident-spot-like domain (see Fig. 3 below for an illustration).

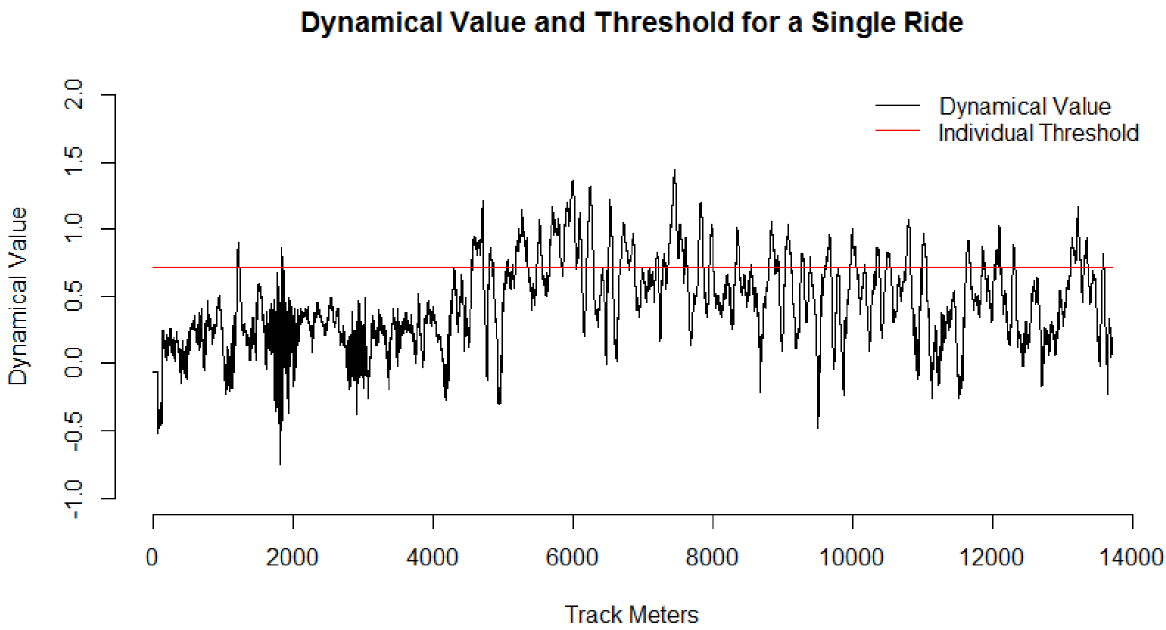

Figure 3: The measured dynamics on each meter of the track can be transformed into a value (“Dynamical Value”) that encodes whether dynamics might be in a similar risk domain as known accident spots. An individual threshold, derived from the Dynamic Values at known accident spots, allows to identify domains of interest for safety studies.

6. Building an Overlay of Threshold Exceedances

Each meter on the track is assigned a value of zero and then incremented by one, for every ride with the dynamical value above the driver's threshold. E.g. by counting how many threshold exceeding events occurred at this spot. This is done jointly for all drivers on the given track e.g. their counts are summed up and on each meter the counts are then divided by the number of valid measurements on that meter (e.g. the maximal number of possible counts), thus giving a proportion of occurred threshold exceeding events to potential threshold exceeding events. Since the dynamical value fluctuates quite a bit, this map of threshold exceeding event proportions is again smoothed (with a window of 60 meters), to provide a consistent structure of potential accident spot like dynamics locations. An illustration can be found in Fig. 4 below.

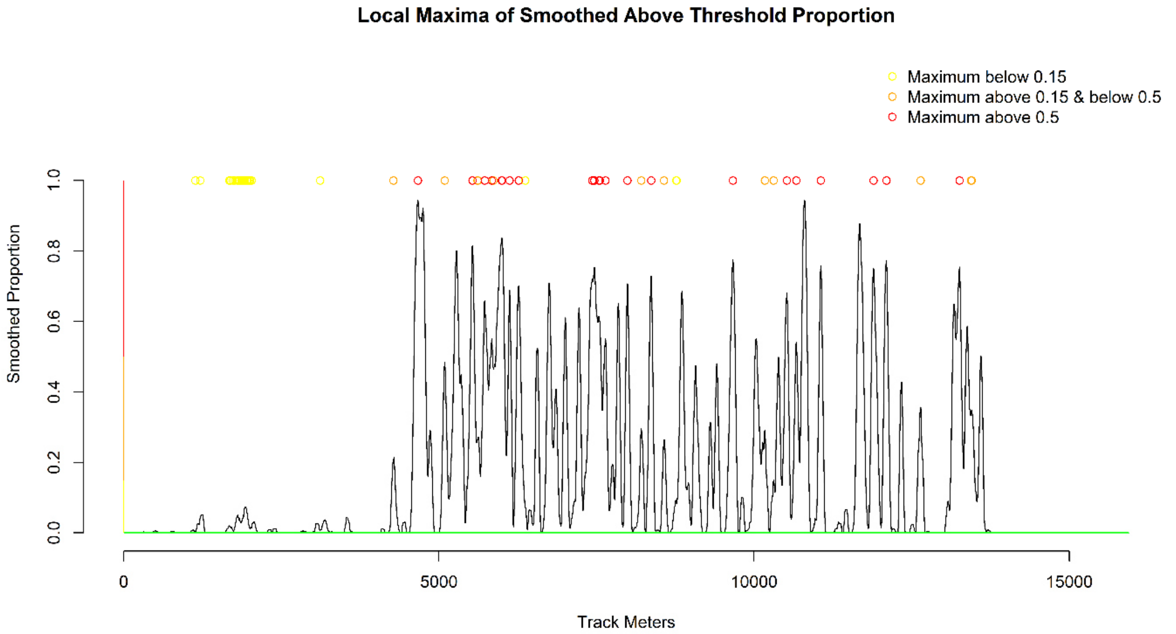

Figure 4: Outline of the smoothed (via rollmean) proportion of threshold exceeding events on the Kalte Kuchl track (over all drivers and rides). Above we marked the local maxima of the resulting curve. Only those local maxima are then used to estimate the potential risk spots and their surrounding area. The maxima are colour coded as "low" (yellow, 0 to 15 percent), "middle" (orange, 15 percent to 50 percent) and "high" (red, 50 percent and above) stress locations. The colour green encodes meters with no local maximum nearby.

The local maxima of the resulting curve serve as the spots of "local dynamical stress" and around them an area of 120 meters before and after them is labelled their "area" of influence (lead up and follow up area, accounting for potential ambiguity of accident spot localization). These maxima and their areas are colour coded, depending on the proportion of above-threshold observations. Yellow maxima reach less than 15 percent of the possible threshold exceeding events and are local maxima of low stress. Orange maxima and areas reach at least 15 percent and up to 50 percent and can provide a serious indication of accident risk. Red maxima and areas reach at least 50 percent and serve as definite indicators of challenging driving dynamics. The focus on maxima was decided in order to have estimates for local peaks within the "yellow" area, without immediately "colouring" the whole track and just focus on the local maxima instead.
Based on the Kalte Kuchl track, we present an exemplary map, as would be the result of this method, in Fig. 5 below.

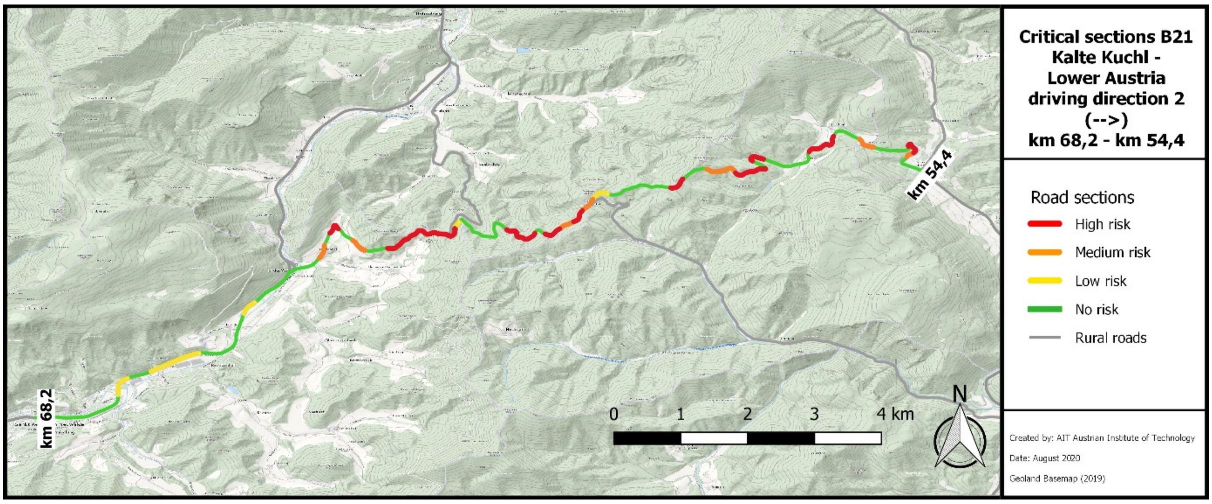

Figure 5: Example of a resulting map for the Kalte Kuchl. Spots in this case are the ones shown in Fig. 4.

We consider an accident to be “hit” by our model, if its location is in one of the coloured areas described above. For the Kalte Kuchl track mentioned here, this yields a “Hit” share of 68 percent, to be compared with a total share of 52 percent of the track being in any of the coloured areas. The Kalte Kuchl is an extreme example, given the high number of known accident spots and the many small curves on the track. On the totality of all 6 tracks we have 38 percent in any coloured area and 66 percent of accidents hit.

7. Discussion and Summary

We have analysed data collected in 2017 by 6 motorcycle drivers on 6 reference tracks, covering a total of approximately 3000 kilometers driven in almost 200 test-rides on the MoProVe. Driving dynamics were quantified in terms of speed, accelerations (X-, Y-, Z-) and angle changes (Roll-Rate, Yaw-Rate, Pitch-Rate) and per-meter values were provided. These dynamics data were postprocessed via smoothing (rolling average), separation by sign and approximate derivatives (first differences). The values at known accident spots were compared to the values around which the dynamical values were most likely to cluster (kNN cluster centers) via a high dimensional linear separation model. The resulting function on the driving dynamics allowed to provide individual estimates, the “dynamic value” for each driver. It also enabled us to define a threshold for when driving dynamics reach a domain of similar individual stress compared to known accident sites. Combining the estimates of multiple drivers, it was possible to define a map of areas around local maxima of the number of threshold exceeding events and these areas contained 66 percent of the considered accident sites.
This methodology combines unsupervised approaches, such as clustering (kNN), with supervised approaches (such as the linear separation) and forms a group effect (the overlay of threshold exceeding events) at points of interest. This has the advantage of stabilizing and summarizing the estimates of the very different individual drivers, while still characterizing drivers individually.

The resultant map may be used to guide road safety efforts and study the highlighted spots in detail to determine, whether actual threats to safety can be noted there and infrastructure improvement might mitigate the effects of crashes in these spots. The colour coded system can serve to prioritize investigations and measures for a road safety inspection on the analysed tracks.
The obtained estimates become generalizable due to the overlay of multiple driver’s estimates and the individual thresholds and clusters, which allow to adjust for personal driving styles. Therefore, a systematic study of motorcycle safety on common tracks based on riding dynamics becomes feasible by this approach.

The current approach of using clustering on the individual riding styles and separations between typical driving dynamics and driving dynamics at known risk spots leading to overlays of multiple limit transgressions for a final estimate of accident spot like dynamics has been filed as a European patent [15] and will be the focus of further work.

8. Limitations and Further Work

The used methods are still quite simple and it may well be that other approaches, such as support vector machines or deep learning, might yield yet better results for the separation.
On the other hand, interpretability is already a challenge, as the combination of individual threshold exceeding events makes it hard to discern the individual causes of threshold exceedance. Further work will focus on creating interpretable criteria for risk spots, based on the described risk map approach.
The individual threshold estimates might be refined further, to define personal safety profiles of drivers that would be interested to use such a functionality and provide online feedback to the drivers.

Finally, the use of only the extrema of the threshold exceeding events overlay may be replaced by types of hazard maps, that would instead colour the whole area above a certain threshold, rather than the areas around extrema. This approach will require a refinement of the used criteria, as the estimates of which points to choose will have to be able to still hit known accident spots, without resorting to colouring large areas around each above-threshold point, since this would otherwise include a far too big proportion of the investigated track.

Acknowledgements

This project resulted from a research grant of the VSF Austrian Road Safety Fund [2]. The authors gratefully acknowledge the funding from VSF.

We thank KTM Austria [6] which provided the motorcycle used in this study.

The authors would like to thank (in alphabetic order) Matthias Hahn, Till J. Kniffka, Johann Schindele and Peter Unterkreuter for their significant support in instrumenting, adapting and riding MoProVe.

References

[1] Statistik Austria 2017; [https://www.statistik.at] visited August 01, 2020.

[2] VSF Austrian Road Safety Fund, [https://www.bmk.gv.at/themen/verkehr/strasse/verkehrssicherheit/vsf.html] visited August 01, 2020

[3] Schwieger K., Saleh P., Hula A.; Ecker H., Neumann M., viaMotorrad – Can motorcycle safety be measured? ,12th International Motorcycle Conference Cologne. October 2018

[4] Hula A., Schwieger K., Saleh P., Neumann M., Ecker H., Objectifying and Predicting Motorcycle Accident Risk through Riding Dynamics, Proceedings of the International Technical Conference on the Enhanced Safety of Vehicles (ESV), 2019. June 10-13, Eindhoven, Netherlands

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

[6] KTM Sportmotorcycle GmbH; [http://www.ktm.com/] visited August 01, 2020.

[7] 2d-datarecording, 2D Debus & Diebold Messsysteme; [http://2d-datarecording.com/produkte/] visited August 01, 2020.

[8] VBOX automotive, VBOX 3i Dual Antenna | 100Hz Vehicle Dynamics Measurement; [https://www.vboxautomotive.co.uk/index.php/en/products/data-loggers/vbox-3i-dual-antenna] visited August 01, 2020.

[9] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL https://www.R-project.org/.

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

[11] Ling, R. F. “On the theory and construction of k-clusters". The Computer Journal, 1972. 15 (4): 326–332

[12] Bishop, C.M. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg, 2006

[13] Scrucca L., Fop M., Murphy T. B., Raftery A. E. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models The R Journal, 2016, 8/1, pp. 205-233

[14] Rosenblatt, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, Cornell Aeronautical Laboratory, Psychological Review, 1958, v65, No. 6, pp. 386–408. doi:10.1037/h0042519

[15] Hula, A., METHOD FOR DETERMINATION OF ACCIDENTAL HOTSPOTS FOR MOTORISED TWO-WHEELED VEHICLES, European Patent EP3691960A1, Filed 09.December.2019, Published 12.August.2020