ORIGINAL PAPER

Weather impacts on various types of road crashes: a quantitative analysis using generalized additive models

Nico Becker[1,2]* , Henning W. Rust[1,2] and Uwe Ulbrich[1]

Abstract

Adverse weather conditions can have different effects on different types of road crashes. We quantify the combined effects of traffic volume and meteorological parameters on hourly probabilities of 78 different crash types using generalized additive models. Using tensor product bases, we model non-linear relationships and combined effects of different meteorological parameters. We evaluate the increase in relative risk of different crash types in case of precipitation, sun glare and high wind speeds. The largest effect of snow is found in case of single-truck crashes, while rain has a larger effect on single-car crashes. Sun glare increases the probability of multi-car crashes, in particular at higher speed limits and in case of rear-end crashes. High wind speeds increase the probability of single-truck crashes and, for all vehicle types, the risk of crashes with objects blown on the road. A comparison of the predictive power of models with and without meteorological variables shows an improvement of scores of up to 24%, which makes the models suitable for applications in real-time traffic management or impact-based warning systems. These could be used by authorities to issue weather-dependent driving restrictions or situation-specific on-board warnings to improve road safety.

Keywords: Road crashes, Crash type, Traffic volume, Weather, Precipitation, Sun glare, Wind speed, Generalized Additive Model, Relative risk, Impact model

1 Introduction

In almost all regions of the world, the road transport system is a key infrastructure most people have to use on a daily basis, in spite of knowing about the pending danger of severe crashes. [1]. In Germany, for example, 300,143 road crashes with injuries and 3,046 fatalities were recorded in 2019 [2]. Various factors can influence the probability of road crashes, including technical or environmental conditions as well as driver behavior. Understanding these influencing factors can help authorities to establish precursory measures like permanent speed limits or improvements of road design. With respect to

variable risk factors like adverse weather conditions, temporary restrictions or warnings can be imposed. To support the identification of measures that help to improve road safety, quantitative knowledge of the relationships between weather and crash probabilities should be provided at a sufficiently high spatial and temporal resolution.

Of course, weather is not only a direct factor for the risk of crashes. It also influences traffic volume, which in turn is one of the main factors related to road crashes; generally, an increasing traffic volume is related to increasing crash rates [3]. Studies addressing crashes at the level of individual road segments generally take traffic volume into account [4]. However, if crashes are analyzed in aggregated form at a regional scale, traffic volume is less frequently considered. The larger the spatial aggregation of crash information, the

*Correspondence: nico.becker@fu-berlin.de

[1] Institute of Meteorology, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6-10, 12165 Berlin, Germany
Full list of author information is available at the end of the article

Becker *et al. European Transport Research Review*   (2022) 14:37

Page 2 of 18

more difficult it is to relate it to measured traffic data, since traffic volume measurements are only available at limited number of locations. A common approach to bypass this problem is to use variables like the hour of the day [5] or the day of the week [6] as a substitute for actual traffic measurements. Furthermore, characteristics of the permanent road environment, such as speed limit, curvature or slope, are important aspects affecting the risk of road crashes. While the effect of speed limits as a measure for risk reduction has often been confirmed [7], the effect of other road characteristics such as slope and curvature have been analyzed less frequently [4].

A large number of studies addresses the effects of different meteorological factors on road safety [3]. A meta analysis of 34 studies addressing the effect precipitation finds an average increase in crash rates of 71% and 84% in case of rain and snowfall, respectively [8]. In terms of crash severity, however, there is a significant reduction under rainy conditions compared to fine weather [9]. The effect of precipitation on crash risk can be different for different types of crashes. For example, the relative risk for single- and multiple-vehicle crashes on Finnish motorways in case of snow is 3.37 and 1.98, respectively, compared to the probability within a random sample [10]. This effect is partly related to single-vehicle run-off-road crashes, which appear to occur more frequently under rain, sleet or snow and in curved road sections [11].

The effect of wind on road safety has not been extensively explored in the literature [3]. In general, the number of road vehicle crashes caused by strong wind is small compared to the total number of crashes [9]. However, wind gusts are shown to increase run-off-road crashes by small but significant amounts of 0.3 to 0.5% [6]. Among different vehicle types, high-sided trucks, vans or buses are most affected by wind [12]. In general, greater recorded wind speeds increase the severity of injuries in single-truck crashes [13].

The effect of sun glare on crashes is only addressed in a few studies. Crash data from signalized crossroads in Tucson, Arizona, show that broad-side and rear-end crashes occur more frequently during glare, but no effect of sun glare on crash severity is found [14]. Injury crashes in Japan indicate that sun glare has an particularly strong impact on pedestrian crashes, bicycle crashes and crashes at crossroads, while there is no indication that the effect of sun glare increases with vehicle speed [15].

Although different studies focus on the effects of specific meteorological parameters on specific crash types, these studies usually differ with respect to region, time period, and methodology, which makes it difficult to compare the results. For a consistent comparison,

it would be useful to apply a single methodological approach to multiple crash types.

In a previous study, a logistic regression model for hourly probabilities of weather-related road crashes was developed at the level of administrative districts in Germany, taking into account the combined effect of precipitation and temperature [5]. Using weather forecast data it was shown that skillful predictions of crash probabilities are possible. However, the model did not explicitly consider traffic volume, but instead assumed a simple diurnal cycle. Furthermore, only weather-related crashes were considered that were classified by the police as being caused by road condition (e.g., slippery road due to water, snow or ice).

The aim of the present study is to extend this model by including observed hourly traffic volume, as well as the effect of precipitation, temperature, sun glare and wind gusts. While previous studies have commonly used traditional weather station data, we derive meteorological predictor variables from gridded radar and reanalysis products. To allow for more flexible functional relationships and combined effects of multiple variables, the classical logistic regression model is replaced by a Generalized Additive Model (GAM) for dichotomous target variables. Models are developed for 78 different crash types in a consistent approach, to compare the weather effects and predictive power of the models for different speeds limits, crash types, road environments and crash severities.

2 Data
2.1 Crash data
A data set with anonymized information from police reports of road crashes in Germany from 2006 to 2017 is used (Source: Research Data Centre of the Federal Statistical Office and Statistical Offices of the Länder, *Statistik der Straßenverkehrsunfälle*, 2006-2017, own calculations). The data set includes severe road crashes, which refers to all crashes with vehicles left unroadworthy, with injuries or fatalities. Crashes related to alcohol consumption of the driver are not included. In total 4,695,687 complete crash reports are available for the period under investigation. The location of the individual crashes is available at the level of administrative districts (*Landkreise*). Because of several territorial reforms during the study period, all crashes are assigned to boundaries of the 401 German administrative districts as they existed in 2017.

Based on the crash reports, we distinguish between different crash characteristics: the type of vehicles involved, the speed limit at the location of the crash, the crash type, the characteristics of the road environment, and the crash severity (see Table 1 for a detailed description of crash characteristics used in this study). It should be

Table 1 Description of crash characteristics used to create the dependent variables of the generalized additive models

Name	Description
Vehicle type	
any	All single- and multi-vehicle crashes and all vehicle types are considered (including bicycles, motorbikes, buses and tractors).
multi-car	Only crashes with two or more passenger cars are considered.
single-car	Only crashes of single passenger cars are considered.
single-truck	Only crashes of single trucks are considered.
Speed limit	
[0, 50)	Only crashes at speed limits below 50 km/h are considered
[50, 70)	Only crashes at speed limits between 50 and 70 km/h are considered
70, 100)	Only crashes at speed limits between 70 and 100 km/h are considered
100, 130)	Only crashes at speed limits between 100 and 130 km/h are considered
[130, *Inf*)	Only crashes at speed limits of 130 km/h and above are considered
Crash type	
run-off-road	A vehicle run off the road.
hit-object	A vehicle collided with an obstacle on the road (e.g. trees, debris, wildlife).
broad-side	Crash with cross traffic or vehicles that turn onto the road from a side road.
head-on	Crash with oncoming traffic, without one of the crash partners intending to turn across the opposite lane.
side-swipe	Crash with another vehicle travelling sideways in the same direction, when driving side by side or when changing lanes.
rear-end	Crash with another vehicle driving ahead or waiting for traffic reasons.
parking	Crash with vehicles stopping or parking at the edge of the carriageway, in the marked parking spaces immediately adjacent to the edge of the carriageway, on footpaths or in parking spaces.
Road environment	
curve	Crash occurred on a curved road section.
descent	Crash occurred on a descending road section.
ascent	Crash occurred on an ascending road section.
t-junction	Crash occurred at a t-junction.
crossroads	Crash occurred at a crossroads.
Crash severity	
no injuries	No injuries or fatalities.
minor injuries	At least one road user with minor injuries, but no serious injuries or fatalities.
serious injuries	At least one road user with serious injuries, but no fatalities.
fatalities	At least one road user died.

noted that the actual driving speed of vehicles may differ from the speed limit used to categorize the crashes. The speed limit should therefore only be interpreted as a rough indicator of the traffic conditions at the location of the crash. In total 78 different crash types with specific characteristics are considered by always combining one of the four vehicle types with one of the other crash characteristics (e. g. single-truck crashes at speed limits between 70 and 100 km/h, or multi-car crashes at crossroads). For each of the resulting 78 crash types an hourly time series of a dichotomous variable is created for all administrative districts, being zero if no crash happened within the hour considered and one otherwise. These hourly time series are used as target variables (dependent variables) in generalized additive models.

2.2 Traffic data

The German Federal Highway Research Institute (*Bundesanstalt für Straßenwesen*, BASt) operates a traffic measurement network on federal highways (*Autobahn*) and federal roads (*Bundesstraßen*). Federal highways usually have two or three lanes per direction, partly without speed limits, while federal roads usually have one lane per direction and general speed limits of up to 100 km/h. At about 2,000 traffic counting stations the hourly number of vehicles is registered. The data set provides separate counts for different vehicle types. In this study the total vehicle counts are used, as well as the counts of passenger cars and trucks.

Hourly count data of 1,400 traffic measurement stations, which contain at least five years of data between 2006 and 2017, are used in this study. Missing data in the traffic count time series are filled using Poisson

regression models for weather-related variations of hourly traffic counts developed in a previous study [16].

The hourly traffic counts of each station are rescaled so that 0 and 1 correspond to the average daily minimum and maximum hourly traffic count, respectively. This makes the data at different traffic stations comparable and suitable for an application as a predictor variable in our modeling approach. Note, that values below 0 and above 1 can occur at individual hours as the reference for rescaling is an average minimun/maximun.

Because a single traffic measurement station might not be representative for a whole administrative district, for each district the five traffic stations closest to the district center are identified and for each hourly time step the mean of the rescaled traffic volume of these five stations is computed. This is done separately for federal road and highway stations.

2.3 Radar-based precipitation data

Precipitation values are derived from the RADOLAN data set [17], provided by the German Meteorological Service, which contains hourly precipitation sums on a spatial grid with a spatial resolution of 1 km for the area of Germany. RADOLAN combines radar reflectivity, measured by the 16 C-band Doppler radars of the German weather radar network, and ground-based precipitation gauge measurements. As from radar reflectivity we cannot directly infer the precipitation amount at the ground, observations from rain gauges are used to calibrate the precipitation amounts estimated from the radar reflectivity in an online-procedure. Thus, RADOLAN combines the benefits of high spatial resolution of the radar network with the accuracy of gauge-based precipitation measurements.

For each administrative district, all RADOLAN grid points within the district boundaries are identified and for each hourly time step the average precipitation of all identified grid points is computed. These hourly precipitation estimates at district level are subsequently used as a predictor variable.

2.4 Reanalysis data

The fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) global atmospheric reanalysis (ERA5) is a synthesis of various heterogeneous meteorological observational data and atmospheric model simulations, which is produced using a fixed version of the numerical weather forecasting model and data assimilation scheme [18]. ERA5 contains different atmospheric and surface variables on a global grid with a spatial resolution of 30 km at an hourly temporal resolution. The advantage of ERA5 over station-based observations is the spatial and temporal homogeneity. However, it should be

noted that local station measurements can deviate from the gridded ERA5 values.

For each administrative district all ERA5 grid points within the district boundaries are identified and for each hourly time step the district average surface temperature, total cloud cover and maximum hourly wind gust is computed and subsequently used as a predictor variable.

3 Methods

3.1 Generalized additive models

The probability p for a certain event to occur can be described with a logistic linear model

$$\log\left(\frac{p}{1-p}\right) = \alpha + \mathbf{X_i}\vec{\beta} \tag{1}$$

with l predictor variables (or independent variables) $\mathbf{X_i} = (X_{i1}, ...X_{il})$, where $\vec{\beta} = (\beta_1, ..., \beta_l)$ are the corresponding model parameters, α is the intercept and n is the number of available observations. The logistic regression model is a powerful tool for modeling the effects of predictor variables on event probabilities. However, if the functional relationship between predictor variables and probability is complex, or if non-linear interactions between different continuous predictor variables have to be taken into account, finding an appropriate transformation of the predictor variables can be cumbersome. In generalized additive models [19] the concept of generalized linear models is extended by adding smooth functions of predictor variables to the linear term of the equation, so that

$$\log\left(\frac{p}{1-p}\right) = \alpha + \mathbf{X_i}\vec{\beta} + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + ... , \tag{2}$$

where the f_j are smooth functions of the predictor variables x_k. Specifying relationship between predictor and target variable (dependent variable) in terms of smooth functions makes generalized additive models more flexible than generalized linear models.

The predictor variables can contribute to the model as additive effects, like $f_1(x_{x_1})$ and $f_2(x_{x_2})$ in Eq. 2, for example. In this case, the effect of x_1 on the target variable is independent from the value of x_2. The smooth function f can be written as

$$f(x) = \sum_{j=1}^{J} b_j(x)\beta_j, \tag{3}$$

where $b_j(x)$ is the jth of some *basis functions* and β_j are some unknown parameters, which must be estimated. The basis functions are usually based in some way on splines. Commonly used smoothers in generalized

additive models are cubic regression splines which are also used in the present study.

The assumption of additive effects is a quite restrictive case of the more general function of two variables $f(x_1, x_2)$ [19]. Eq. 3 can be generalized to allow smooth functions of any number of predictor variables using tensor product bases. For a smooth function of two predictor variables, for example, we can write

$$f(x_1, x_2) = \sum_{i=1}^{I} \sum_{l=1}^{L} \delta_{il} a_i(x_1) c_l(x_2) \tag{4}$$

where δ_{il} are the parameters, and the $a_i(x_1)$ and $c_l(x_2)$ are the basis functions.

3.2 Model setup

For each of the 78 crash types introduced above two different models are developed: First, model $\mathcal{M}_{\mathrm{met}}$

$$\log\left(\frac{p_a}{1-p_a}\right) = \alpha + \mathrm{Yr} + f_1(\bar{p}_{a,d}) + f_2(\mathrm{Trf}) + f_2(\mathrm{Wnd})$$
$$+ f_4(\mathrm{Tmp}, \mathrm{Prc}) + f_5(\mathrm{Cld}, \mathrm{Elv}) \tag{5}$$

to describe the effects of meteorological and non-meteorological variables on crash probability p_a (see Tab 2 for a description of the variables in the predictor). Second, model $\mathcal{M}_{\mathrm{nomet}}$

$$\log\left(\frac{p_a}{1-p_a}\right) = \alpha + \mathrm{Yr} + f_1(\bar{p}_{a,d}) + f_2(\mathrm{Trf}), \tag{6}$$

with only non-meteorological terms in the predictor as a reference model.

There are different possible approaches to distinguish between administrative districts. For example, a separate model could be built for each district. However, this is difficult for sparsely populated districts or rare crash types, because there are not enough crashes in the time series to establish robust relationships between crash probability and the different meteorological parameters. Instead, a single model including all districts is built for each of the 78 crash types. We distinguish between the districts by using the time averaged crash probability of each district $\bar{p}_{a,d}$ as a predictor variable.

The rescaled hourly traffic volume Trf is based on hourly counts of all vehicles, cars or trucks, according to the crash type considered in the model. When considering speed limits above 100 km/h, only traffic counts at highway stations are used, since on federal roads speed limits larger than 100 km/h are rare.

The interaction term for temperature and precipitation in Eq. 5 allows for a different effect of precipitation on crash probability for different temperatures. This is important, since we cannot directly distinguish between rain and snow using the radar-based precipitation estimates.

Furthermore, cloud cover and sun elevation is included as an interaction term to allow for different effects of cloud cover at different elevation angles. This is important to capture potential effects of sun glare.

For each of the the 78 crash types, we use a bootstrap approach and estimate the model parameters of $\mathcal{M}_{\mathrm{met}}$ 100 times, each time drawing randomly 10,000,000 of the 42,153,120 available observations (with replacement). This allows us to estimate confidence intervals for the analysis of the functional relationships and to estimate, if values of relative risk increase can be regarded as statistically significant, as described below.

3.3 Relative risk

The crash probability under adverse meteorological conditions $p_{a,adv}$ can be compared to the crash probability under

Table 2 Independent variables used in generalized additive models

Variable	Description
Yr	Categorical variable with a category for each year from 2006 to 2017 to capture long term temporal changes in crash probability due to external factors like improved safety features of vehicles.
p_a	Probability of at least one occurrence of a specific crash type within a one hour time interval in a certain district with an average crash probability $\bar{p}_{a,d}$.
$\bar{p}_{a,d}$	Average hourly crash probability in an administrative district.
Trf	Average hourly traffic volume of the five traffic measurement stations closest to the district centre. Traffic volume is rescaled, so that 0 and 1 correspond to the average daily minimum and maximum traffic volume at a traffic station, respectively.
Wnd	Hourly maximum wind gusts, averaged over all ERA5 grid cells within disctrict boundaries.
Tmp	Hourly surface temperatures, averaged over all ERA5 grid cells within disctrict boundaries.
Prc	Hourly precipitation sum, averaged over all RADOLAN grid cells within disctrict boundaries.
Cld	Hourly total cloud cover, averaged over all ERA5 grid cells within disctrict boundaries.
Elv	Angle of sun elevation above the horizon, where positive and negative values correspond to the sun beeing located above and below the horizon, respectively.

meteorological reference conditions $p_{a,ref}$ using the measure of *Relative Risk*

$$RR = \frac{p_{a,adv}}{p_{a,ref}} \tag{7}$$

and *Relative Risk Increase*

$$RRI = 1 - \frac{p_{a,adv}}{p_{a,ref}}. \tag{8}$$

The RRI is computed for winter precipitation, summer precipitation, sun glare and extreme wind speeds using Eq. 5 (see Table 3 for parameter settings).

As described above, for each of the 78 crash types the model $\mathcal{M}_{met}$ is fitted 100 times in a bootstrap approach. For each of the 100 model versions RRI values are computed and the averages of these RRI values are presented in the results section. If more than 95 of the 100 models show a positive or negative RRI, we conclude that the RRI is significantly positive or negative, respectively.

3.4 Model performance
The area under the receiver operating characteristic curve (AUC) is a measure of the ability of a model to discriminate between events and non-events (see Additional file 1 for details and additional metrics for testing the validity of the models). The AUC ranges between 0.5 and 1, which compares to random guessing and perfect discrimination, respectively.

A skill score *SS* is a relative measure of how much a prediction *S* outperforms a reference prediction S_r, defined as

$$SS = (S - S_r)(S_p - S_r)^{-1}, \tag{9}$$

where S_p is the score of a perfect prediction. In this study we compute the AUC Skill Score (AUCSS), which compares the AUC of model $\mathcal{M}_{met}$ (Eq. 5) with meteorological predictor variables to the AUC of model $\mathcal{M}_{nomet}$ (Eq. 6) without meteorological predictor variables.

We compute the AUCSS in a two-fold cross validation approach. The available data is split randomly into a training and a testing data set. The training data is used for estimating the model parameters, while testing data is used for computing the AUCSS. This is repeated after switching the testing and training data and the resulting AUCSS values are averaged.

4 Results
4.1 Average crash probability
Prior to the analysis of the statistical models, the average probability that at least one road crash occurs within one hour in a German administrative district is computed for each of the 78 crash types considered (Fig. 1). If all crashes are considered without distinguishing between specific crash characteristics, the hourly probability is 9.487%. Probabilities are lower if computed for more specific vehicle types. For example, the probability for single-car or a single-truck crashes is 1.441% and 0.111%, respectively. The lower crash probability for trucks can at least partly be attributed to a lower number of trucks on the roads and to a lower vulnerability of trucks due to their structural characteristics.

Crashes are further classified in terms of the speed limit at the crash location, crash type, road environment and crash severity. For certain crash types probabilities are relatively low, in particular in case of some sub-types of single-truck crashes (hit-object crashes, crashes at crossroads and with fatalities), where the probability is 0.002. This should be kept in mind when interpreting the RRI values for these crash types.

Table 3 Setup of meteorological parameters for calculation of relative risk increase for different meteorological conditions. $p_{a,adv}$ and $p_{a,ref}$ is the crash probability under adverse and reference conditions, respectively. In all cases, Yr = 2017, Trf = 1 and $\bar{p}_{a,d}$ is set to the median value of all districts

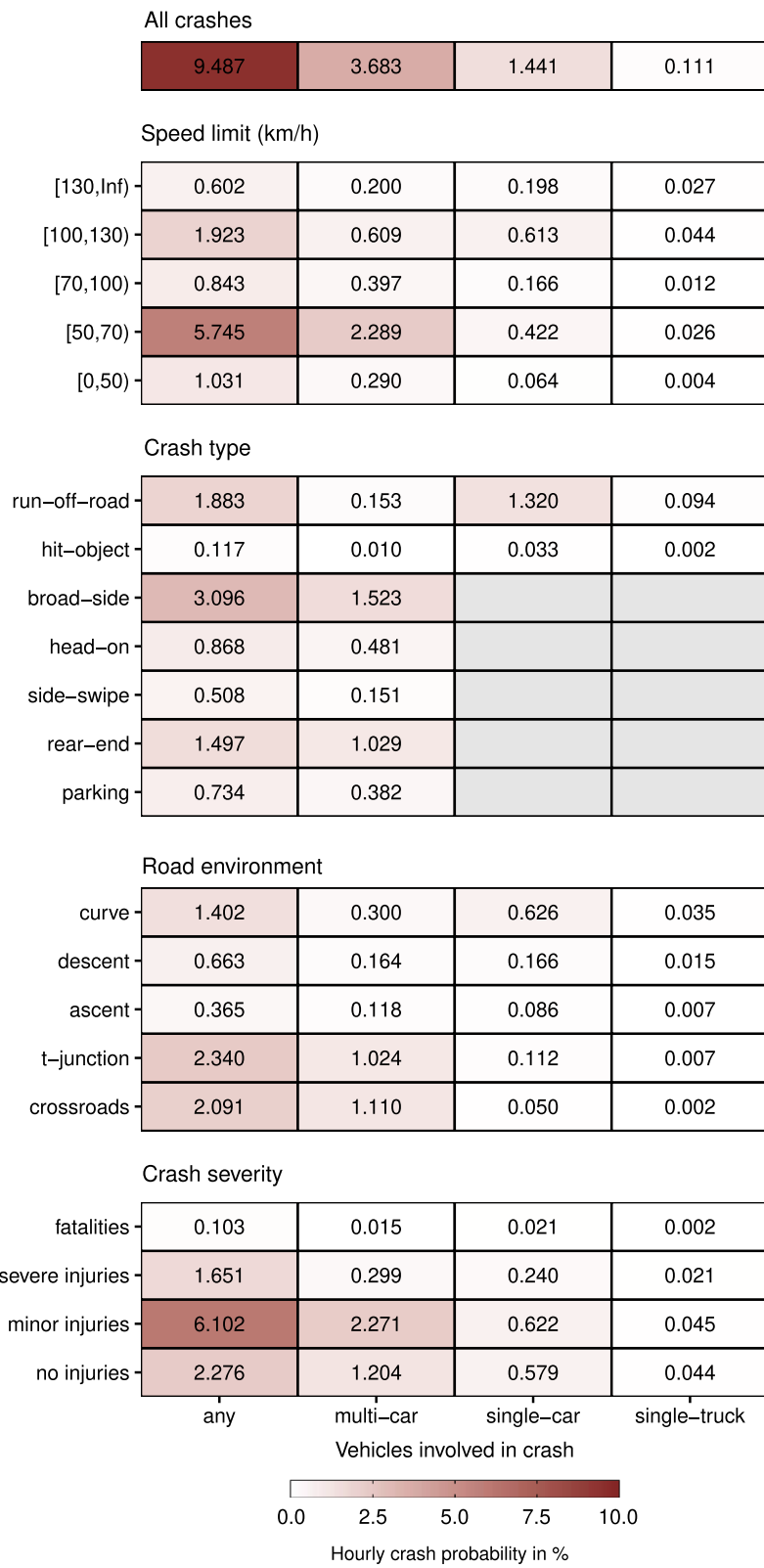| Weather condition	Probability	Tmp	Prc	Cld	Elv	Wnd
Winter precipitation	$p_{a,adv}$	−3℃	1 mm/h	100%	30°	10 m/s
	$p_{a,ref}$	+3℃	0 mm/h	100%	30°	10 m/s
Summer precipitation	$p_{a,adv}$	15℃	1 mm/h	100%	30°	10 m/s
	$p_{a,ref}$	15℃	0 mm/h	100%	30°	10 m/s
Low sun, no clouds	$p_{a,adv}$	15℃	0 mm/h	0%	20°	10 m/s
	$p_{a,ref}$	15℃	0 mm/h	100%	20°	10 m/s
Extreme wind	$p_{a,adv}$	15℃	0 mm/h	100%	30°	25 m/s
	$p_{a,ref}$	15℃	0 mm/h	100%	30°	5 m/s

Fig. 1 Average hourly probabilities for 78 different crash types in German administrative districts

4.2 Functional relationships

Within the frame of this paper a detailed discussion of models for all 78 crash types is infeasible. Instead, as an example, we discuss the functional relationships between the occurrence probability of multi-car rear-end crashes and the different predictor variables. For this discussion, we focus on one predictor term in Eq. 5 at a time. Modeled crash probability is then plotted against this term, while keeping the predictor variables of the other terms constant at selected values.

An increase in traffic volume leads to a non-linear but monotonous increase in rear-end crash probability, if all other parameters are held constant (Fig. 2a). This is a reasonable behavior, which indicates that the spatially aggregated traffic data can adequately represent the effects of traffic volume on crash probability at district level.

To account for different basic properties in different administrative districts, such as the total number of vehicles per district, we have included the average hourly crash probability $\bar{p}_{a,d}$ in the model. In districts with larger $\bar{p}_{a,d}$ the model shows that also the crash probability p_a under specific meteorological conditions is larger (see Fig. 2b). The relationship between $\bar{p}_{a,d}$ and p_a is approximately piecewise linear with a change in slope at around $\bar{p}_{a,d} = 0.1$. Note that values of $\bar{p}_{a,d} > 0.1$ are only reached in a few highly populated districts.

Rear-end crash probability shows a small variability in time after controlling for traffic volume and meteorological factors (Fig. 2c). A systematic trend in rear-end crash probability is not evident. However, it should be noted that other crash types show a decreasing trend, which could be attributed to advanced safety features of cars, for example (not shown).

An increase in wind speed leads only to a small increase in probability of rear-end crashes (Fig. 2d). If the confidence intervals are taken into account, this increase is not significant. This is not surprising, since one would not expect a large impact of wind speed on this crash type, but the following section will reveal wind speed impacts on other crash types.

There is a clear impact of summer precipitation (Tmp = 15°C) on rear-end crash probability (indicated by color and line type in Fig. 2a and b). If the district average hourly precipitation changes from 0 to 1 mm/h, crash probability increases by a factor of 1.548, which corresponds to an RRI of 54.8%. A further increase in precipitation to 2 and 3 mm/h leads to smaller increases in crash probability, but it is still significant with respect to the confidence intervals.

At negative temperatures the increase in rear-end crash probability is even stronger in case of increasing precipitation, which is shown in a visualization of the combined effect (interaction) of precipitation and surface temperature (Fig. 2e). For example, if precipitation changes from 0 mm/h to 1 mm/h at -3°C, crash probability increases by 80.1%. The varying effect of precipitation depending on temperature can be described by including the two variables as an interaction term in the generalized additive model. One could expect a sharper increase in crash probability around 0°C, however, the functional relationship shows a rather smooth transition between positive and negative temperatures. This can be attributed to different sources of uncertainty, which are represented by the smoothing term. For example, the actual road surface temperatures at specific locations within a district can differ from the aggregated surface temperatures based on the gridded ERA5 data used in the model.

Finally, the visualization of the combined effect of cloud cover and sun elevation angle shows that there are two areas with increased probabilities of rear-end crashes (Fig. 2f). First, increased probabilities occur at positive sun elevation angles combined with low cloud cover. This could be attributed to a distraction of drivers due to sun glare, leading to a larger number of rear-end crashes. Second, an increasing probability is also observed with increasingly negative elevation angles, indicating that the sun is below the horizon. This increase could be attributed to reduced visibility under low-light conditions during night time. Note that probabilities are computed assuming a constant traffic volume. At night times one can expect traffic volume (and thus also crash probability) to decrease, which counteracts the increase in probability due to low light conditions.

4.3 Increase in crash probability due to weather conditions

4.3.1 Winter precipitation

For the 78 crash types, we compute the *Relative Risk Increase* (RRI) for hours with winter precipitation with respect to hours without winter precipitation as described in the methods section. Under the selected conditions (Table 3), precipitation is most likely snowfall or freezing rain. In general, single-vehicle crashes (i.e. single-car and single-truck crashes) show a larger RRI compared to multi-vehicle crashes (i.e. multi-car crashes and crashes including all vehicle types; Fig. 3). Single-truck crashes show the largest RRI of 872.9%.

A higher (or even no) speed limit at the location of the crash leads to a larger RRI under conditions with winter precipitation. In case of single-truck crashes the RRI increases to 1521.7% at speed limits of 130 km/h. However, it should be noted that generally the maximum speed of trucks above 3.5 t is limited to 80 km/h.

Run-off-road and head-on crashes also show a relatively large RRI under conditions with winter
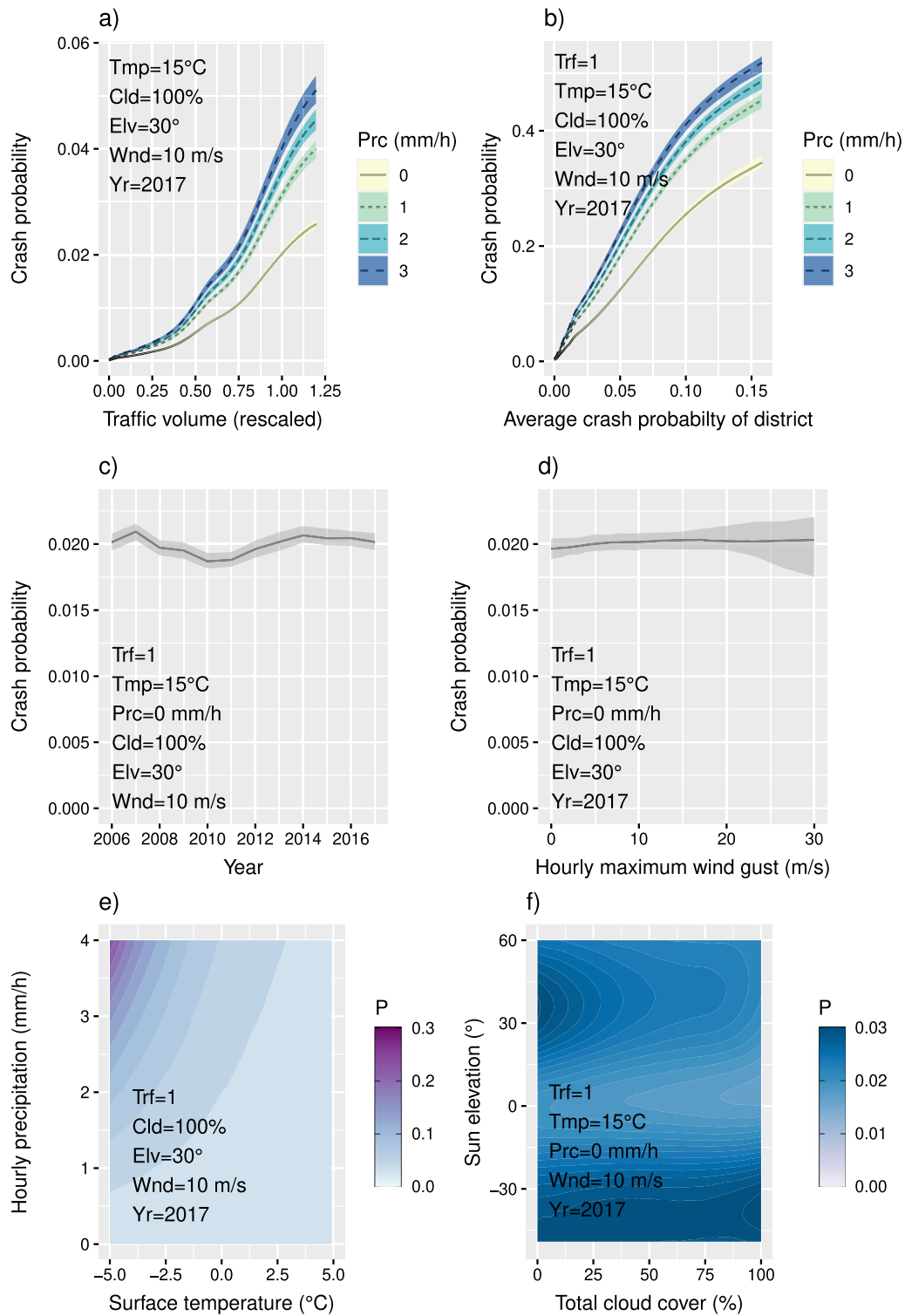
Fig. 2 Functional relationships between predictor variables and the hourly probability of muli-car rear-end crashes estimated by a generalied linear model. 95% confidence intervals (shaded areas) are estimated from 100 models fitted with randomly drawn training data

All crashes

any	multi-car	single-car	single-truck
177.3*	160.4*	309.6*	872.9*

Speed limit (km/h)

	any	multi-car	single-car	single-truck
[130,Inf)	563.9*	251.8*	1114.9*	1521.7*
[100,130)	300.2*	302.2*	349.6*	599.4*
[70,100)	191.9*	136.4*	269.5*	471.5*
[50,70)	109.8*	121.5*	206.3*	467.5*
[0,50)	46.2*	62.9*	53.7*	136.0*

Crash type

	any	multi-car	single-car	single-truck
run–off–road	304.6*	290.3*	327.7*	977.6*
hit–object	–2.8	–3.2	30.1	–0.4
broad–side	46.3*	58.4*		
head–on	482.0*	519.1*		
side–swipe	121.5*	127.4*		
rear–end	92.3*	86.4*		
parking	64.3*	52.2*		

Road environment

	any	multi-car	single-car	single-truck
curve	317.4*	508.6*	244.5*	565.7*
descent	445.5*	531.6*	474.1*	892.6*
ascent	231.1*	222.4*	238.7*	825.4*
t–junction	80.7*	82.3*	149.4*	235.4*
crossroads	61.9*	58.1*	117.4*	194.2*

Crash severity

	any	multi-car	single-car	single-truck
fatalities	15.4	30.7	–11.4	16.3
severe injuries	66.1*	122.6*	90.6*	236.4*
minor injuries	135.0*	136.0*	269.1*	565.8*
no injuries	300.9*	198.1*	491.8*	1238.4*

Vehicles involved in crash

RRI in %   –2000 –1000 0 1000 2000

Fig. 3 Relative risk increase (RRI) of crash probabilities in situations with precipitation and negative temperature (Tmp = −3 °C and Prc = 1 mm/h) compared to situations without precipitation and positive temperatures (Tmp = +3 °C and Prc = 0 mm/h). Significant changes (i. e. more than 95 of 100 models fitted with randomly drawn training data show the same direction of change) are indicated with an asterisk

precipitation, indicating that vehicles tend to leave their lane due to slippery road conditions. The RRI for other crash types are smaller, but in most cases positive and significant. When comparing different road environments, RRI values are largest in curves and descending road segments. Single-truck crashes also show a strong increase on ascending roads segments. If the RRI under winter precipitation conditions is computed separately for crashes with different severities, it is evident that less severe crashes show larger RRI values compared to more severe crashes.

4.3.2 Summer precipitation

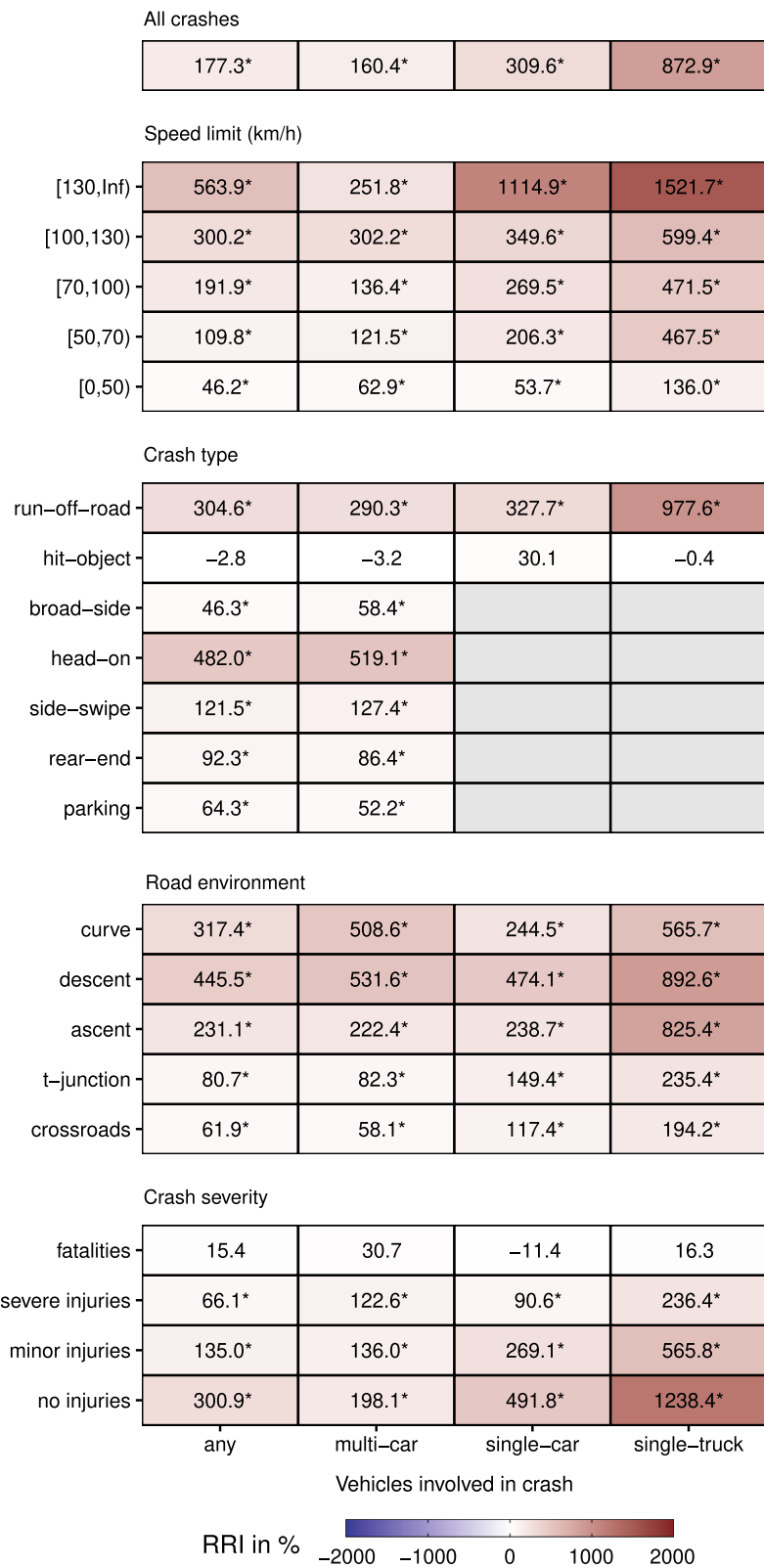Analogously to winter precipitation, we investigate the effect of summer precipitation (most likely rainfall). The RRI is computed for hours with summer precipitation with respect to hours without summer precipitation. In case of summer precipitation, crash probability increases for all crash types (Fig. 4). However, the increase is generally smaller than in case of winter precipitation. Similar to winter precipitation, summer precipitation leads to larger RRI values in case of single-vehicle crashes, at higher driving speeds, as well as in case of run-off-road crashes and in curves, descents and ascents. Probabilities of less severe crashes increase more than those of severer crashes. While in case of winter precipitation RRI values for single-truck crashes were generally larger than single-car crashes, in case of summer precipitation the opposite is true. The largest RRI of 536.2% is found in case of single-car crashes at speed limits of 130 km/h and above.

4.3.3 Low sun elevation at cloud-free conditions

To evaluate the effect of sun glare on crash probability, we compute the RRI for hours with low sun elevation angle and cloud-free conditions with respect to hours with low sun elevation angle and full cloud cover. The RRI of different crash types under low sun and cloud-free conditions range from -35% to +52% (Fig. 5). In case of multi-vehicle crashes, probabilities increase for most crash types. The largest RRI values of multi-car crashes crashes occur at speed limits of 130 km/h and more (+52.5%) and in case of rear-end crashes (+43.2%). These increases could be attributable to sun glare, which could lead to reduced visibility and increased reaction times of drivers, which is particularly dangerous at high driving speeds and in dense traffic.

In case of single-car and single-truck crashes, RRI values are negative in most cases. However, in case of single-trucks these decreases are mostly not significant. These decreases of crash probability of single-vehicle crashes under low sun and cloud-free conditions could be due to the fact that we have not taken time lagged effects of precipitation into the design of our models. If

within the hour of the crash there was no rain, the road surface is more likely to be dry under cloud-free conditions (e.g. due to evaporation effects due to sunshine) and more likely to be wet under cloudy conditions (e.g. due to possible precipitation at previous time steps). A higher likelihood for a dry road (with higher surface friction) consequently leads to the observed reductions of single-car crash probabilities, which is particularly large in case of curves (-35.3%).

4.3.4 Extreme wind speeds

To evaluate the effect of extreme wind speeds on crash probability, the RRI at hours with high wind speeds is computed with respect to hours with low wind speeds (Fig. 6). In general, the RRI values in case of extreme wind speeds are relatively small, compared to the effects of the other meteorological parameters analyzed above, and mostly not significant, except for single-truck crashes. Those show a significant RRI of 104.9%, which is in line with other studies showing that trucks are particularly vulnerable to high wind speeds [12]. RRI values of single-truck crashes are largest at high speed limits between 100 and 130 km/h. Since the maximum speed of trucks is limited to 80 km/h, this effect could be explained by the assumption that highways with such high speed limits often run through open rural terrain and are particularly exposed to high wind speeds.

What stands out are the hit-object crashes, which increase in probability by 413.6% for single-car crashes and by 789.8% for single-truck crashes. These crashes can be attributed to crashes with broken tree branches, debris, or other objects, which are blown onto the road by strong winds.

4.4 Cross-validation results

The predictive power of the model $\mathcal{M}_{met}$ and whether meteorological terms in the predictor improve the predictions compared to a model without these ($\mathcal{M}_{nomet}$) is analyzed in a two-fold cross-validation experiment using the AUC and AUCSS. The AUC values for the 78 different crash types mainly range between 0.7 and 0.85. Values between 0.7 and 0.8 correspond to an acceptable discrimitation, values above 0.8 correspond to an excellent discrimination [20].

For all crash types (except for single-truck crashes with fatalities), the AUCSS values are positive, indicating an improvement of the models ability to discriminate between time steps with and without crashes due to the meteorological predictor variables (Fig. 7). The AUCSS ranges between relatively low values of 0.44% in case of single-car crashes with fatalities to high values of 24.21% in case of single-car crashes at locations with high speed limits of 130 km/h and above. In general, AUCSS values

All crashes

any	multi-car	single-car	single-truck
40.5*	56.0*	151.5*	101.6*

Speed limit (km/h)

	any	multi-car	single-car	single-truck
[130,Inf)	173.9*	84.3*	536.2*	121.0*
[100,130)	40.7*	51.0*	127.0*	65.9*
[70,100)	56.9*	48.9*	169.5*	116.0*
[50,70)	34.1*	58.4*	98.4*	73.9*
[0,50)	22.9*	45.6*	66.2*	41.1*

Crash type

	any	multi-car	single-car	single-truck
run–off–road	105.6*	129.0*	164.8*	123.2*
hit–object	3.8	4.2	21.6*	3.8
broad–side	29.8*	46.1*		
head–on	50.9*	85.3*		
side–swipe	16.1*	58.1*		
rear–end	46.0*	55.4*		
parking	25.8*	35.3*		

Road environment

	any	multi-car	single-car	single-truck
curve	95.9*	132.9*	202.1*	155.6*
descent	64.3*	99.8*	229.8*	109.4*
ascent	67.0*	64.5*	232.2*	150.0*
t–junction	32.6*	48.1*	83.2*	32.3*
crossroads	41.6*	49.6*	75.0*	39.7

Crash severity

	any	multi-car	single-car	single-truck
fatalities	6.9*	21.8*	22.9*	11.0
severe injuries	14.9*	48.0*	63.8*	41.2*
minor injuries	36.1*	57.0*	135.0*	74.4*
no injuries	92.0*	59.9*	257.2*	149.3*

Vehicles involved in crash

RRI in %   −1000 −500   0   500 1000

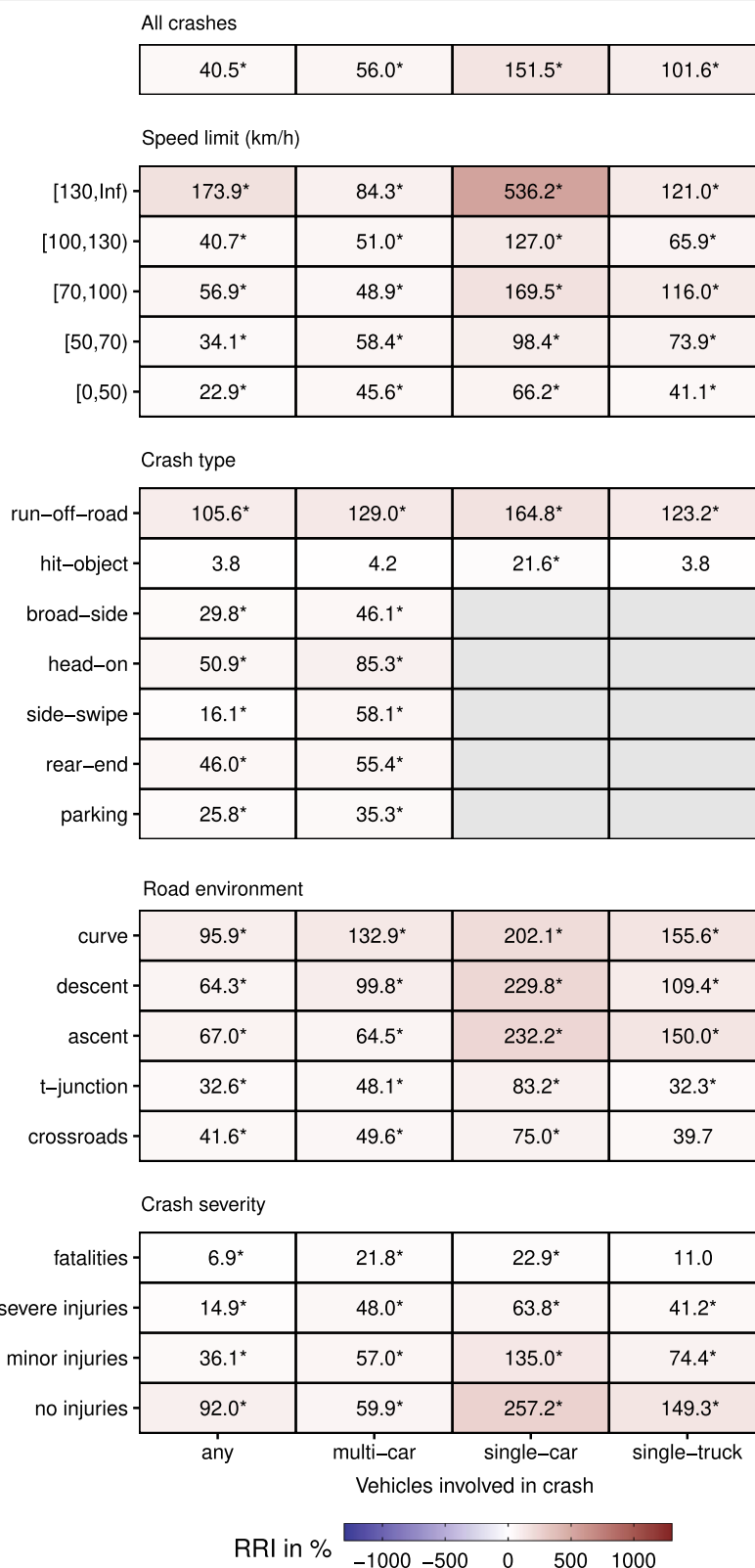**Fig. 4** Relative risk increase (RRI) of crash probabilities in situations with precipitation and positive temperature (Tmp $= 15\,°$C and Prc $= 1$mm/h) compared to situations without precipitation and positive temperatures (Tmp $= 15\,°$C and Prc $= 0$ mm/h). Significant changes (i. e. more than 95 of 100 models fitted with randomly drawn training data show the same direction of change) are indicated with an asterisk

All crashes

22.8*	25.6*	−21.2*	−8.0

Speed limit (km/h)

	any	multi–car	single–car	single–truck
[130,Inf)	29.2*	51.5*	−23.2*	−2.8
[100,130)	20.9*	29.8*	−22.0*	−6.0
[70,100)	25.2*	25.6*	−28.3*	−9.7
[50,70)	26.2*	23.6*	−14.3*	−1.2
[0,50)	23.5*	18.3*	−5.2	−2.1

Crash type

	any	multi–car	single–car	single–truck
run–off–road	−7.2*	−5.6*	−23.4*	−13.0*
hit–object	43.2*	18.1	27.6*	38.1
broad–side	30.5*	21.2*		
head–on	26.1*	12.4*		
side–swipe	30.1*	18.2*		
rear–end	45.3*	43.2*		
parking	26.6*	21.7*		

Road environment

	any	multi–car	single–car	single–truck
curve	−1.1	−1.5	−35.3*	−17.2*
descent	11.6*	11.3*	−27.6*	−4.9
ascent	23.7*	26.7*	−22.5*	9.3
t–junction	31.0*	24.5*	−16.0*	−9.4
crossroads	33.6*	24.5*	−6.6	18.9

Crash severity

	any	multi–car	single–car	single–truck
fatalities	35.9*	4.7	−7.8	3.1
severe injuries	36.6*	22.4*	−10.6*	7.7
minor injuries	25.9*	28.7*	−18.0*	−4.7
no injuries	11.1*	20.0*	−34.7*	−18.7*
	any	multi–car	single–car	single–truck

Vehicles involved in crash

RRI in %   −100 −50 0 50 100

Fig. 5 Relative risk increase (RRI) of crash probabilities in situations with low sun elevation angle and cloud free conditions (Elv = 20° and Cld = 0%) compared to situations with low sun elevation angle and clouded conditions (Elv = 20° and Cld = 100%). Significant changes (i. e. more than 95 of 100 models fitted with randomly drawn training data show the same direction of change) are indicated with an asterisk
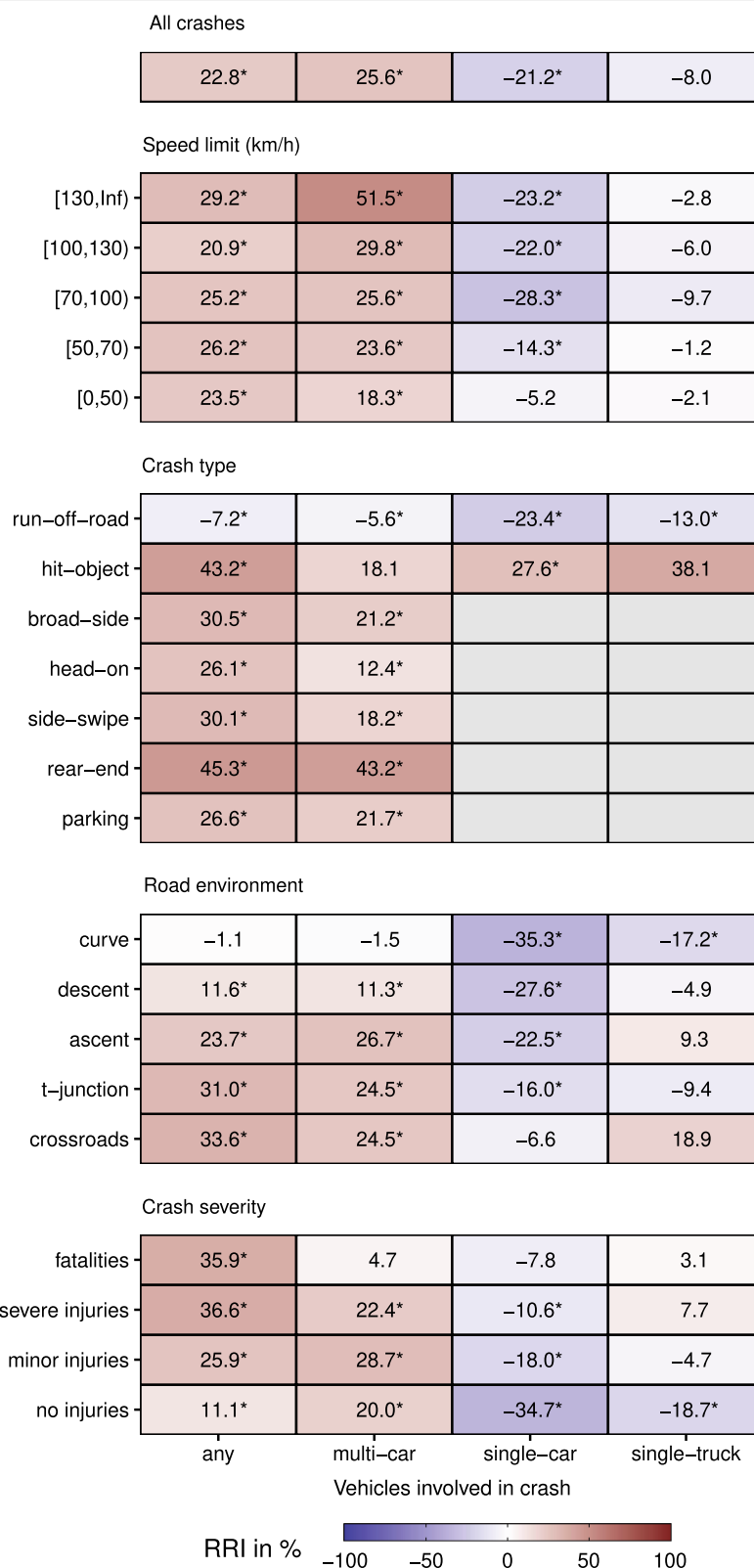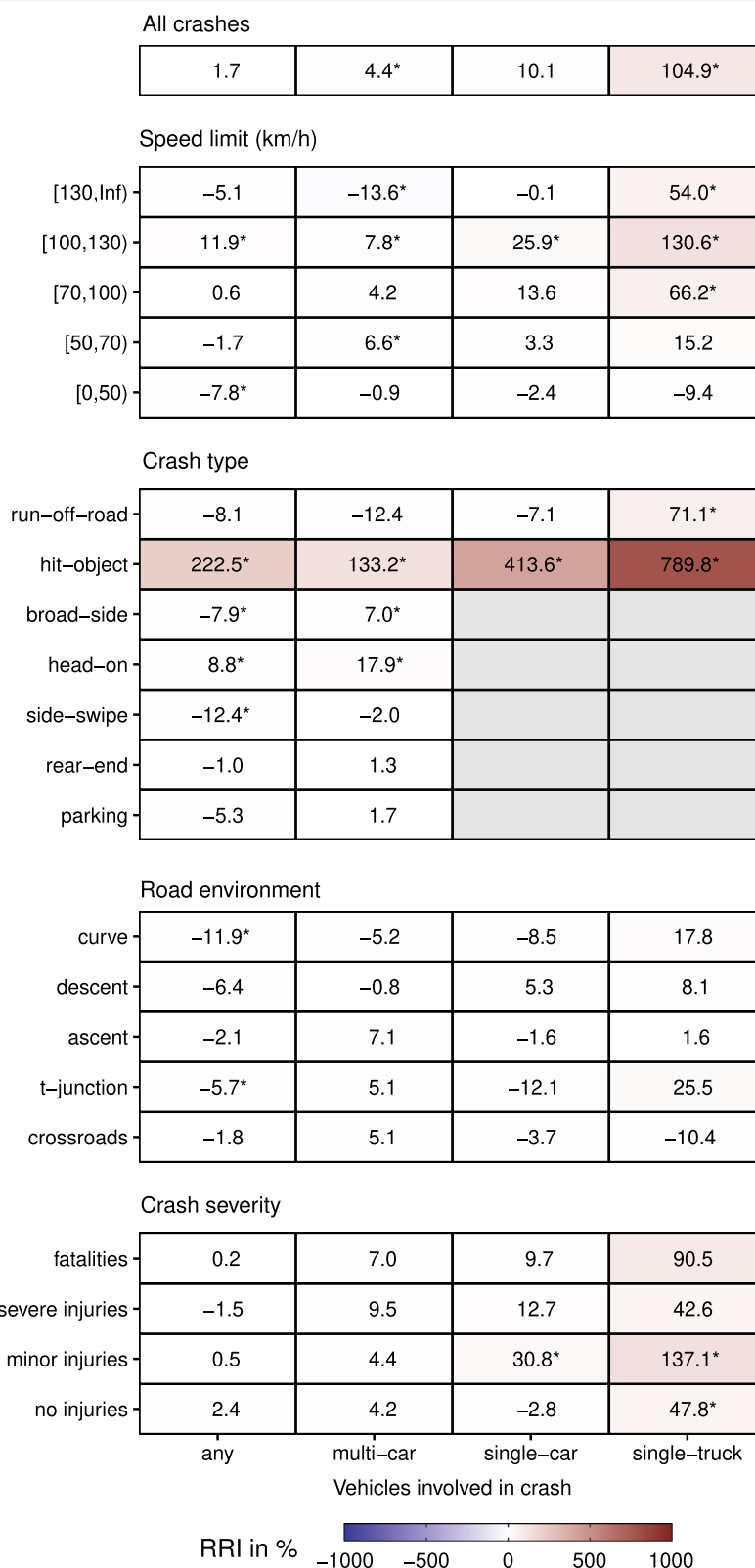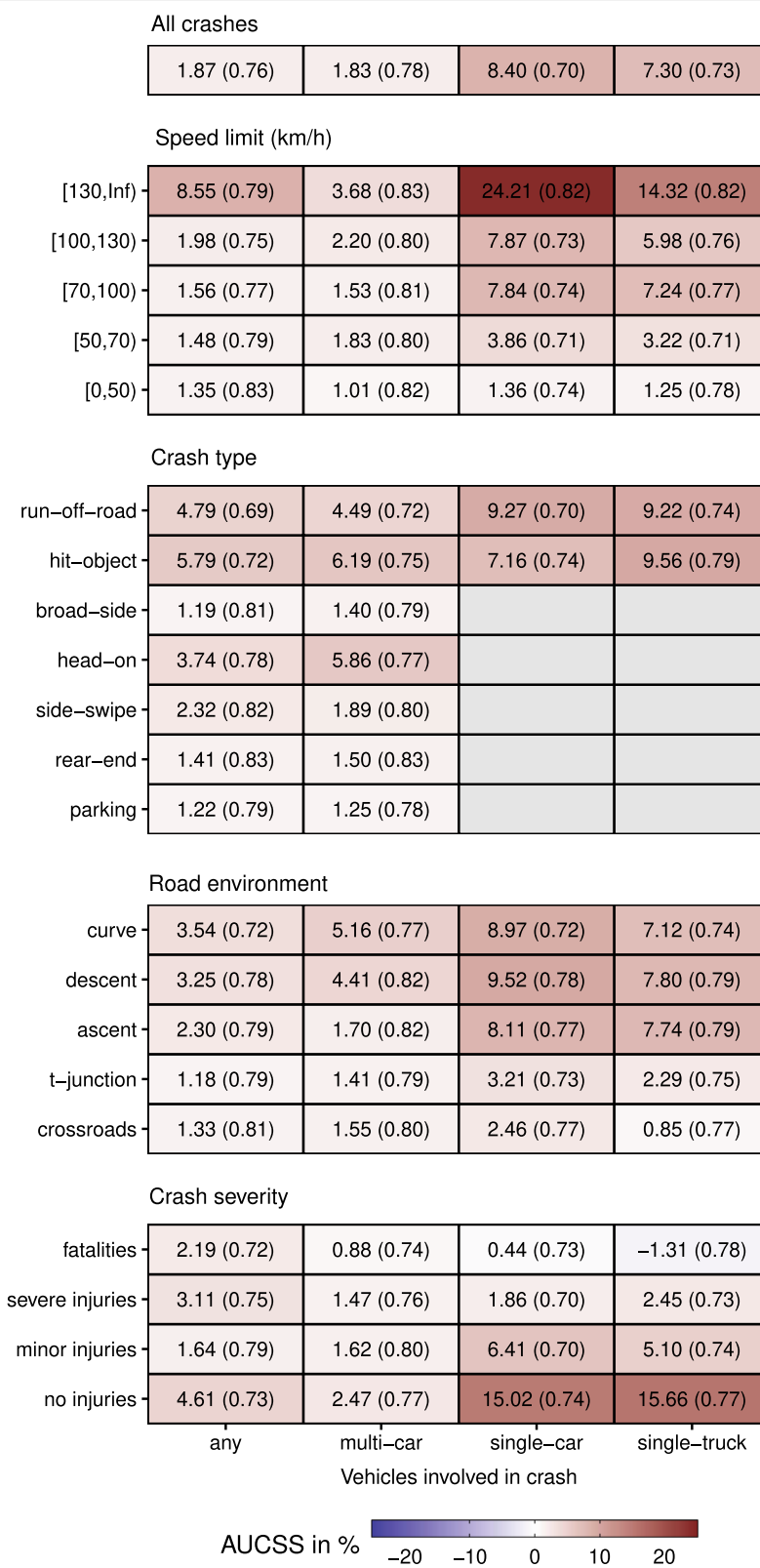
All crashes

1.7	4.4*	10.1	104.9*

Speed limit (km/h)

	any	multi-car	single-car	single-truck
[130,Inf)	−5.1	−13.6*	−0.1	54.0*
[100,130)	11.9*	7.8*	25.9*	130.6*
[70,100)	0.6	4.2	13.6	66.2*
[50,70)	−1.7	6.6*	3.3	15.2
[0,50)	−7.8*	−0.9	−2.4	−9.4

Crash type

	any	multi-car	single-car	single-truck
run−off−road	−8.1	−12.4	−7.1	71.1*
hit−object	222.5*	133.2*	413.6*	789.8*
broad−side	−7.9*	7.0*		
head−on	8.8*	17.9*		
side−swipe	−12.4*	−2.0		
rear−end	−1.0	1.3		
parking	−5.3	1.7		

Road environment

	any	multi-car	single-car	single-truck
curve	−11.9*	−5.2	−8.5	17.8
descent	−6.4	−0.8	5.3	8.1
ascent	−2.1	7.1	−1.6	1.6
t−junction	−5.7*	5.1	−12.1	25.5
crossroads	−1.8	5.1	−3.7	−10.4

Crash severity

	any	multi-car	single-car	single-truck
fatalities	0.2	7.0	9.7	90.5
severe injuries	−1.5	9.5	12.7	42.6
minor injuries	0.5	4.4	30.8*	137.1*
no injuries	2.4	4.2	−2.8	47.8*
	any	multi−car	single−car	single−truck

Vehicles involved in crash

RRI in %   −1000   −500   0   500   1000

Fig. 6 Relative risk increase (RRI) of crash probabilities in situations with high wind speeds (Wnd = 25 m/s) compared to situations with low wind speeds (Wnd = 5 m/s) . Significant changes (i. e. more than 95 of 100 models fitted with randomly drawn training data show the same direction of change) are indicated with an asterisk

All crashes

1.87 (0.76)	1.83 (0.78)	8.40 (0.70)	7.30 (0.73)

Speed limit (km/h)

	any	multi-car	single-car	single-truck
[130,Inf)	8.55 (0.79)	3.68 (0.83)	24.21 (0.82)	14.32 (0.82)
[100,130)	1.98 (0.75)	2.20 (0.80)	7.87 (0.73)	5.98 (0.76)
[70,100)	1.56 (0.77)	1.53 (0.81)	7.84 (0.74)	7.24 (0.77)
[50,70)	1.48 (0.79)	1.83 (0.80)	3.86 (0.71)	3.22 (0.71)
[0,50)	1.35 (0.83)	1.01 (0.82)	1.36 (0.74)	1.25 (0.78)

Crash type

	any	multi-car	single-car	single-truck
run-off-road	4.79 (0.69)	4.49 (0.72)	9.27 (0.70)	9.22 (0.74)
hit-object	5.79 (0.72)	6.19 (0.75)	7.16 (0.74)	9.56 (0.79)
broad-side	1.19 (0.81)	1.40 (0.79)		
head-on	3.74 (0.78)	5.86 (0.77)		
side-swipe	2.32 (0.82)	1.89 (0.80)		
rear-end	1.41 (0.83)	1.50 (0.83)		
parking	1.22 (0.79)	1.25 (0.78)		

Road environment

	any	multi-car	single-car	single-truck
curve	3.54 (0.72)	5.16 (0.77)	8.97 (0.72)	7.12 (0.74)
descent	3.25 (0.78)	4.41 (0.82)	9.52 (0.78)	7.80 (0.79)
ascent	2.30 (0.79)	1.70 (0.82)	8.11 (0.77)	7.74 (0.79)
t-junction	1.18 (0.79)	1.41 (0.79)	3.21 (0.73)	2.29 (0.75)
crossroads	1.33 (0.81)	1.55 (0.80)	2.46 (0.77)	0.85 (0.77)

Crash severity

	any	multi-car	single-car	single-truck
fatalities	2.19 (0.72)	0.88 (0.74)	0.44 (0.73)	−1.31 (0.78)
severe injuries	3.11 (0.75)	1.47 (0.76)	1.86 (0.70)	2.45 (0.73)
minor injuries	1.64 (0.79)	1.62 (0.80)	6.41 (0.70)	5.10 (0.74)
no injuries	4.61 (0.73)	2.47 (0.77)	15.02 (0.74)	15.66 (0.77)
	any	multi-car	single-car	single-truck

Vehicles involved in crash

AUCSS in %   −20 −10 0 10 20

Fig. 7 Area Under Reviewer Operating Characteristics Curve Skill Score (AUCSS), with positive values indicating an improvement of the predictive power if meteorological predictor variables are included in models for hourly probabilities of different crash types. The Area Under Reviewer Operating Characteristics (AUC) of the models with meteorological predictor variables is given in brackets

Becker *et al. European Transport Research Review*   (2022) 14:37

Page 16 of 18

are higher in case of those crash types, which also showed a strong relationship to one or more of the meteorological variables. Highest AUCSS values occur in case of single-vehicle crashes, at higher speed limits, at locations with curves, descents or ascents and in case of crashes without or with minor injuries. Lower AUCSS values occur in case of multi-car crashes at lower speed limits, in case of crashes with other vehicles (except head-on crashes), at crossroads and in case of crashes with severe injuries and fatalities.

5 Discussion

In general, the results of our analysis are in line with previous studies. For example, an increase in crash probabilities due to precipitation has been found quite consistently in the literature [3]. However, the comprehensive crash data set in combination with the chosen modeling approach allows more precise and quantitative statements about the functional relationships between the meteorological parameters and probabilities of different crash types.

We have compared results for a large number of different crash types, which comes at the expense of detail regarding evaluation of the individual models. An in depth diagnostic of the fitting procedure of each generalized additive model was not possible within the frame of this article. For example, we have found that cubic regression splines generally lead to reasonable functional relationships in the generalizes additive models, but a more detailed analysis could reveal that for specific parameters or crash types other smoothing functions might be more appropriate. Furthermore, we have assumed that the standard setting for the number of basis dimensions of the splines is appropriate.

The radar data used in this study provides highly resolved precipitation estimates. However, it does not distinguish between rain and snowfall. Instead we analyzed the combined effect of precipitation and surface temperature. In future research, novel data products could be used, which combine radar-data and atmospheric models, to provide additional information about the precipitation type.

Furthermore, it should be noted that only the weather conditions within a specific hourly time interval are considered for predicting the hourly crash probability; possible time-lagged effects are neglected. For example, if precipitation occurred before the hour of a crash, but the road surface could still wet. Also the accumulation of snow cover on roads over several hours is not considered. Both effects could lead to an underestimation of crash probabilities during hours without precipitation. A potential effect of winter road maintenance is also not included, because appropriate data is missing.

Also the analysis of the combined effects of cloud cover and sun elevation revealed missing meteorological factors in the model, which are related to time-lagged effects of precipitation and evaporation processes. Future research could focus on including such effects, for example by using information from physical road surface energy balance models taking into account evaporation processes [21].

We have shown that sun glare particularly increases probabilities of rear-end crashes, which is in line with findings for Tucson, Arizona [14]. However, we have also identified a stronger effect of sun glare on crash probabilities at higher speed limits and in case of increasing crash severity, which has not been found in previous studies [15].

In previous research weather station data is frequently used to study the impact of weather on crashes, while we have used post-processed gridded meteorological data sets. Using weather station data assumes that the point measurement is representative for the location of the crash, which might be a certain distance away. Using gridded data, which is spatially aggregated, assumes that the spatially aggregated weather information is representative for a certain location within the aggregation area and that the variability within the area is sufficiently small so that this assumption is valid. We think that using gridded data is most appropriate for our study design. One might consider to compare both approaches in future analyses.

While for Germany count data of motorized road traffic is available for a large number of stations on federal roads and highways, there is no comprehensive data set for smaller roads. Here, we assumed that federal road stations are representative for smaller roads as well. The validity of this assumption could be tested in districts, where additional traffic data is available. Furthermore, there is little measurement data of bicycle and pedestrians volume. The numbers of such non-motorized road users themselves depend on the weather conditions. This is problematic when analyzing crashes involving such road users without an estimate of their actual share in the total traffic volume. For example, an increase in crash probabilities during hours with sun glare could be due to the effects of reduced visibility, but also due to an increased number of bicycles or pedestrians during fair weather conditions. This should be kept in mind, when interpreting the corresponding numbers.

6 Conclusions

While previous studies on weather effects on road crash have often focused on specific weather conditions or certain crash types, we have applied a modeling approach that is able to capture the combined effects

of meteorological parameters on a large number of different crash types. By using additive logistic regression models, we could capture and analyze non-linear functional relationships between meteorological parameters and crash probability, which would have been difficult using other methods like traditional logistic regression.

We have shown that including meteorological variables can substantially improve predictions of crash probabilities. This is particularly true for single-vehicle crashes on road sections with high speed limits, where the largest improvement of verification scores was observed. Our findings can help authorities to identify crash types and road characteristics, where weather-dependent driving restrictions like variable speed limits or situation-specific warnings could be beneficial for road safety. Such warnings could be communicated via on-board computers and navigation systems depending on vehicle type, speed limit and characteristics of the road characteristics. This would be an important step towards moving from traditional weather forecasts towards impact-based warnings, which is heavily promoted by the World Meteorological Organization and national weather services [22, 23].

Abbreviations

AUC: Area under receiver operating characteristic curve; AUCSS: Area under receiver operating characteristic curve skill score; ERA5: European centre for medium-range weather forecasts reanalysis v5; RADOLAN: Radar-based precipitation data (*RAdar-ONLine-ANeichung*, radar online adjustment); RR: Relative risk; RRI: Relative risk increase.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12544-022-00561-2.

Additional file 1. Additional model verification results.

Availability of data and materials
The crash data for Germany was obtained from the Research Data Centre of the Federal Statistical Office and Statistical Offices of the Länder [24]. The hourly traffic count data for highways and federal roads in Germany is available via the Bundesanstalt für Straßenwesen [25]. The RADOLAN data set is available via the German Weather Service [26]. The ERA5 reanalysis data is available at the Climate Data Store [27]. The R [28] package "mgcv" [19] was used for the development of the generalized additive models.

Declarations

Ethics approval and consent to participate
Not applicable

Consent for publication
Not applicable

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Institute of Meteorology, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6-10, 12165 Berlin, Germany. [2]Hans-Ertel-Centre for Weather Research, Berlin, Germany.

References
1. Peden, M., & Sminkey, L. (2004). World health organization dedicates world health day to road safety. *Injury Prevention, 10*(2), 67–67. https://doi.org/10.1136/ip.2004.005405.
2. BASt: Verkehrs- und unfalldaten - kurzzusammenstellung der entwicklung in deutschland. Technical report, Bundesanstalt für Straßenwesen, Bergisch Gladbach, Germany (2020). https://www.bast.de/DE/Publikationen/Medien/VU-Daten/VU-Daten.pdf. Accessed 7 Jan 2022.
3. Theofilatos, A., & Yannis, G. (2014). A review of the effect of traffic and weather characteristics on road safety. *Accident Analysis and Prevention, 72,* 244–256. https://doi.org/10.1016/j.aap.2014.06.017.
4. Ziakopoulos, A., & Yannis, G. (2020). A review of spatial approaches in road safety. *Accident Analysis and Prevention, 135,* 105323. https://doi.org/10.1016/j.aap.2019.105323.
5. Becker, N., Rust, H. W., & Ulbrich, U. (2020). Predictive modeling of hourly probabilities for weather-related road accidents. *Natural Hazards and Earth System Sciences, 20*(10), 2857–2871. https://doi.org/10.5194/nhess-20-2857-2020.
6. El-Basyouny, K., Barua, S., & Islam, M. T. (2014). Investigation of time and weather effects on crash types using full Bayesian multivariate Poisson lognormal models. *Accident Analysis and Prevention, 73,* 91–99. https://doi.org/10.1016/j.aap.2014.08.014.
7. Wang, J., & Huang, H. (2016). Road network safety evaluation using Bayesian hierarchical joint model. *Accident Analysis and Prevention, 90,* 152–158. https://doi.org/10.1016/j.aap.2016.02.018.
8. Qiu, L., & Nixon, W. A. (2008). Effects of adverse weather on traffic crashes: Systematic review and meta-analysis. *Transportation Research Record, 2055*(1), 139–146. https://doi.org/10.3141/2055-16.
9. Edwards, J. B. (1998). The relationship between road accident severity and recorded weather. *Journal of Safety Research, 29*(4), 249–262. https://doi.org/10.1016/S0022-4375(98)00051-6.
10. Malin, F., Norros, I., & Innamaa, S. (2019). Accident risk of road and weather conditions on different road types. *Accident Analysis and Prevention, 122,* 181–188. https://doi.org/10.1016/j.aap.2018.10.014.
11. Liu, C., & Subramanian, R.(2009). Factors related to fatal single-vehicle run-off-road crashes. Technical report, NHTSA's National Center for Statistics and Analysis, Washington, DC. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811232. Accessed 7 Jan 2022.
12. Baker, C., & Reynolds, S. (1992). Wind-induced accidents of road vehicles. *Accident Analysis and Prevention, 24*(6), 559–575. https://doi.org/10.1016/0001-4575(92)90009-8.
13. Naik, B., Tung, L.-W., Zhao, S., & Khattak, A. J. (2016). Weather impacts on single-vehicle truck crash injury severity. *Journal of Safety Research, 58,* 57–65. https://doi.org/10.1016/j.jsr.2016.06.005.

14. Mitra, S. (2014). Sun glare and road safety: An empirical investigation of intersection crashes. *Safety Science, 70,* 246–254. https://doi.org/10.1016/j.ssci.2014.06.005.

15. Hagita, K., & Mori, K. (2014). The effect of sun glare on traffic accidents in Chiba prefecture, Japan. *Asian Transport Studies, 3*(2), 205–219. https://doi.org/10.11175/eastsats.3.205.

16. Becker, N., Rust, H. W., & Ulbrich, U. (2022). Modeling hourly weather-related road traffic variations for different vehicle types in Germany. *European Transport Research Review,14*(16), https://doi.org/10.1186/s12544-022-00539-0.

17. Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, A., Kohler, O., & Gerlach, N., et al.:(2004) Projekt RADOLAN–Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodennied-erschlagsstationen (Ombrometer). Deutscher Wetterdienst, Germany. https://www.dwd.de/DE/leistungen/radolan/radolan_info/abschlussbericht_pdf. Accessed 7 Jan 2022.

18. Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society, 146*(730), 1999–2049. https://doi.org/10.1002/qj.3803.

19. Wood, S. N. (2017). *Generalized additive models: an introduction with R.* Chapman and Hall/CRC.

20. Hosmer, D. W., Jr., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons.

21. Jacobs, W., & Raatz, W. E. (1996). Forecasting road-surface temperatures for different site characteristics. *Meteorological Applications, 3*(3), 243–256. https://doi.org/10.1002/met.5060030306.

22. WMO: WMO guidelines on multi-hazard impact-based forecast and warning services - Part II: Putting Multi-hazard IBFWS into Practice. World Meteorological Organization, Geneva, Switzerland (2021). https://library.wmo.int/doc_num.php?explnum_id=10936. Accessed 7 Jan 2022.

23. RCC, IFRC, MetOffice, UKAid: The future of forecasts: impact-based forecasting for early action. Climate Centre, International Federation for Red Cross and Red Crescent Societies, UK Met Office, UK Aid (2020). https://www.forecast-based-financing.org/wp-content/uploads/2020/09/Impact-based-forecasting-guide-2020.pdf. Accessed 7 Jan 2022.

24. Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder: Statistik der Straßenverkehrsunfälle. https://www.forschungsdatenzentrum.de/de/sonstige-wirtschaftsstatistiken/strassenverkehrsunfaelle. Accessed 7 Jan 2022.

25. Bundesanstalt für Straßenwesen: Automatische Zählstellen auf Autobahnen und Bundesstraßen. https://www.bast.de/BASt_2017/DE/Verkehrstechnik/Fachthemen/v2-verkehrszaehlung/Stundenwerte.html. Accessed 7 Jan 2022.

26. Deutscher Wetterdienst: RADOLAN (Radar-Online-Aneichung): Analysen der Niederschlagshöhen aus radar- und stationsbasierten Messungen im Echtzeitbetrieb. https://www.dwd.de/DE/leistungen/radolan/radolan.html. Accessed 7 Jan 2022.

27. Copernicus: Climate Data Store. https://cds.climate.copernicus.eu/. Accessed 7 Jan 2022.

28. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2022). R Foundation for Statistical Computing. https://www.R-project.org. Accessed 17 May 2022.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.