

ORIGINAL PAPER

Open Access



Modeling hourly weather-related road traffic variations for different vehicle types in Germany

Nico Becker^{1,2*} , Henning W. Rust^{1,2} and Uwe Ulbrich¹

Abstract

Weather has a substantial influence on people's travel behavior. In this study we analyze if meteorological variables can improve predictions of hourly traffic counts at 1400 stations on federal roads and highways in Germany. Motorbikes, cars, vans and trucks are distinguished. It is evaluated in how far the mean squared error of Poisson regression models for hourly traffic counts is reduced by using precipitation, temperature, cloud cover and wind speed data. It is shown that in particular motorbike counts are strongly weather-dependent. On federal roads the mean squared error is reduced by up to 60% in models with meteorological predictor variables, when compared to models without meteorological variables. A detailed analysis of the models for motorbike counts reveals non-linear relationships between the meteorological variables and motorbike counts. Car counts are shown to be specifically sensitive to weather in touristic regions like seaside resorts and nature parks. The findings allow for several potential applications like improvements of route planning in navigation systems, implementations in traffic management systems, day-ahead planning of visitor numbers in touristic areas or the usage in road crash modelling.

Keywords: Road traffic, Traffic volume, Count data, Weather, Germany, Poisson regression, Prediction, Vehicle type, Motorbikes

1 Introduction

There is strong evidence that weather has a substantial influence on people's travel behavior. However, both strength and direction of the relationship between weather parameters and travel behavior can vary between different locations, depending on characteristics of the local climate or region-specific travel culture [1]. In particular in the mid-latitudes, temperatures can change from adverse to pleasant conditions within the year. Higher temperatures generally lead to an increase of outdoor activities [2–4] and an increasing use of bicycles [1, 5, 6]. However, very high temperatures above 25 to 30°C can be disadvantageous for outdoor activities [7] and cycling [8, 9]. Low temperatures can lead to reduced car traffic, however in case of trucks the impact is less pronounced [10].

Precipitation generally leads to reduced outdoor activities [11–13]. Also car traffic is reduced during rainfall [14, 15], which appears to be particularly the case at weekends [16]. It might play a role that in case of shopping and leisure activities trips are canceled or the mode of transportation and the destination changes due to rainfall [17]. Considerable traffic reductions are reported with snowfall [14, 18–22]. In general, truck traffic is less affected than car traffic [18], because commercial vehicles are less likely to divert trips due to adverse weather [23]. In urban areas precipitation can lead to switching from active (open-air) to motorized (sheltered) transport modes [24], leading to higher levels of transit ridership [25] and public transportation [26].

Compared to precipitation and temperature, wind speed is often overlooked in traffic studies [1]. Some studies document negative impacts of wind speed on cycling [27, 28]. In the case of motorized road traffic, some studies show that wind speed decreases traffic

*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

counts [17], other studies find mostly non-significant impacts of wind speed [29].

Boecker et al. [1] find substantial differences between the outcomes of different studies addressing the impact of weather on travel behaviour. They conclude that the existing literature presents an “incomplete and fragmented picture”, identify gaps and suggest ideas for further research. Based on their findings, we address the following points in an approach to model the impact of weather on hourly traffic counts:

- Traffic and meteorological data need to be matched accurately in time and space to study their relationships. This can be difficult, if traditional weather station data is used, because stations might be located far away from the location of the traffic measurement. This is particularly relevant in case of precipitation, which can vary strongly in time and space and might not be captured well by station data. Therefore, we use reanalysis and radar-based precipitation products to derive meteorological parameters from high-resolution gridded data sets.
- Existing studies make use of a wide variety of multivariate modeling techniques. However, in many studies linear relationships are assumed between weather and different types of travel behavior, although not all effects seem to be linear in all situations [1]. By applying a stepwise predictor selection procedure, we explore non-linear relationships in a controlled setting.
- While most studies focus on weather impacts on bicycles, cars, or trucks, little is known about weather impacts on motorbike usage. By analyzing a comprehensive database of long-term traffic measurements in Germany that includes motorbike counts, we can fill this gap.

This study aims to quantify to which extent meteorological parameters can improve the predictive skill of models for hourly traffic counts of different vehicle types. This is particularly relevant for application purposes, where an accurate estimation of traffic flow is important. Fields of application are, for example, road crash models, where traffic flow is the dominant factor for crash risk [30, 31], travel-demand and mode-change modeling, traffic management, route planning in navigation systems, and air pollution management.

2 Data

2.1 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 and driving speeds of 100 km/h and more, while federal roads usually have one lane per direction and driving speeds of 100 km/h and less. At the traffic counting stations the hourly number of passing vehicles is registered separately for the two directions of travel. Since it was shown that driving direction is not relevant regarding weather effects [17], the sum of the hourly counts of both directions is used for the analyses at each station. The data set provides counts for different vehicle types. The vehicle types and corresponding abbreviations used in this study are motorbikes (mot), cars (car), vans (van), and trucks (trk).

Count data from 2005 to 2018 is considered in this study. However, many of the available measurement stations have been installed after 2005 or show periods with missing data. Therefore, only stations for which at least five years of data are available are used. This ensures that enough data is available for the modeling procedure. Based on these criteria, 696 stations on highways and 704 stations on federal roads are selected for the analyses.

2.2 Reanalysis data

The fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) global atmospheric reanalysis (ERA5) is a synthesis of various heterogeneous observational data and model simulations, which is produced using a physical model together with a data assimilation scheme [32]. ERA5 contains different atmospheric and surface variables on a global grid with a spatial resolution of 30 km and an hourly temporal resolution. The advantage of ERA5 over station-based observations is the spatial and temporal homogeneity. But it should be noted that local station measurements can deviate from the gridded ERA5 values.

For each traffic counting station the corresponding ERA5 grid cell is identified and the hourly time series of temperature at 2 m height, maximum wind gusts, and total cloud cover is extracted. Using the hourly weather parameters directly as a predictor variable is problematic, in particular in case of temperature. Both temperature and traffic volume is high during the day and low at night, but not because of a causal relationship between the two variables, but because both variables depend on the elevation of the sun. To exclude this spurious relationship from the regression models, the daily maximum temperature, daily maximum wind gusts and daily average cloud cover is used for further analyses.

2.3 Radar data

The RADOLAN data set [33] provided by the German Meteorological Service contains hourly precipitation sums on a grid with a spatial resolution of 1×1 km for

the area of Germany. RADOLAN combines radar reflectivities, 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. The RADOLAN data set thus combines the benefits of high spatial resolution of the radar network and the accuracy of gauge-based measurements.

While the other meteorological predictor variables are aggregated in time, precipitation is included in the model in form of hourly values. All RADOLAN grid points within the radius of 10 km around a traffic station are selected and the spatial average of the hourly precipitation sum is calculated. This results in a predictor variable, which is representative for a larger area around a traffic station. This is reasonable, since the travel behavior of drivers passing a traffic station does not depend solely on the precipitation directly at the station.

2.4 Population data

To analyze whether the impact of weather on traffic flow differs between urban and rural areas, population density data from the German census (Zensus 2011) is used [34]. The data set provides the number of inhabitants in Germany on a grid with a resolution of 1 × 1 km. The number of inhabitants per grid cell is provided as a discrete variable with seven classes. Each class corresponds to a certain range of inhabitant numbers (Table 1). For simplicity, we assume that the actual inhabitant number in a grid cell corresponds to the average of the class range. Since class 7 has no upper bound, the lower bound is used. For each traffic station, all grid cells within a radius of 10 km around the station are selected and the average population density is computed.

Table 1 Inhabitants classes of the German census

Class	Inhabitants (class range)	Inhabitants (as used for aggregation)	Frequency
1	0–3	1.5	146,845
2	3–250	126.5	158,317
3	250–500	375	20,342
4	500–2000	1250	26,183
5	2000–4000	3000	6741
6	4000–8000	6000	2543
7	8000–∞	8000	507

Inhabitants classes and class ranges as used in the German census *Zensus 2011*, the corresponding inhabitant numbers used for spatial aggregation in this study, and the number of grid points of a specific class

3 Methods

3.1 Linear regression and breakpoint detection

The standard linear regression model

$$y_i = \alpha + X_i \beta \quad (i = 1, \dots, n), \tag{1}$$

is a well known technique to relate a target variable y_i to a linear combination of l predictor variables $X_i = (X_{i1}, \dots, X_{il})$, where $\beta = (\beta_1, \dots, \beta_l)$ are the corresponding model parameters, α is the intercept and n is the number of available observations. Predictor variables can be continuous or categorical. Interaction terms can be used when the effect of a predictor variable on the target variable changes, depending on the value of other predictor variables [35].

In Eq. 1 β is usually assumed to be constant with respect to i . However, in case of traffic count data, modifications of the road network in the vicinity of a traffic station can lead to abrupt changes of traffic characteristics. Such breakpoints in the time series can be caused for example by construction sites, road closures or the opening of new roads. In this case, the relationship between X_i and y_i may change and the assumption of constant α and β is no longer valid.

The foundation for estimating single breakpoints in linear regression models was given by Bai [36] and was subsequently extended to multiple breaks [37–39]. To identify breakpoints in the traffic count time series, we use the R package `strucchange` [40, 41], which implements the algorithm described in Bai and Perron [42] for simultaneous estimation of multiple breakpoints. Eq. 1 is extended to

$$y_i = \alpha_j + X_i \beta_j \quad (i = i_{j-1} + 1, \dots, i_j, j = 1, \dots, m + 1)$$

where j is the segment index, $\mathcal{J}_{m,n} = i_1, \dots, i_m$ denotes

the set of the m breakpoints, and by convention $i_0 = 0$ and $i_{m+1} = n$. For a given a set of breakpoints i_1, \dots, i_m the least-squares estimates for the β_j can be obtained. The resulting minimal residual sum of squares is given by

$$RSS(i_1, \dots, i_m) = \sum_{j=1}^{m+1} r_{ss}(i_{j-1} + 1, i_j)$$

$r_{ss}(i_{j-1} + 1, i_j)$ is the minimal residual sum of squares in

the j th segment. The R package `strucchange` applies an efficient algorithm to find the breakpoints $\hat{i}_1, \dots, \hat{i}_m$ that minimize the objective function

$$(\hat{i}_1, \dots, \hat{i}_m) = \underset{(i_1, \dots, i_m)}{\operatorname{argmin}} RSS(i_1, \dots, i_m) \tag{4}$$

over all partitions (i_1, \dots, i_m) with $i_j - i_{j-1} \geq n_h$, where n_h is the minimum length of a segment, which is specified by the user.

3.2 Poisson regression

If y is a count variable, the Poisson regression model

$$y_i = \exp(\alpha + \mathbf{X}_i \beta) \quad (i = 1, \dots, n), \tag{5}$$

can be applied, which belongs to the family of generalized linear models and uses the exponential function as the inverse link function to assure that $y_i \geq 0$ [43]. β is estimated using the iteratively reweighted least squares method [44].

3.3 Assessing model performance

The mean squared error

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (f_i - o_i)^2,$$

is a common metric to evaluate model performance by (6)

comparing the modeled values f_i to the observed values o_i . The squared difference leads to a strong penalization of predictions with larger errors.

A skill score is a relative measure of how a model performs compared to a reference model. The mean squared error skill score

$$\text{MSESS} = 1 - \frac{\text{MSE}_f}{\text{MSE}_r}, \tag{7}$$

where MSE_f is the score of the model under evaluation and MSE_r is the score of the reference model. Positive values of the MSESS indicate an improvement compared to the reference model.

Cross-validation is applied by estimating model coefficients using a training data set and computing scores on an independent testing data set. Here, we split the data randomly into 10 sets. Parameters are estimated on 9 sets and the score is calculated on the remaining set. This is repeated 10 times such that for each set the resulting score is computed. These are then averaged and used for model comparison.

3.4 Model selection procedure

It is infeasible to manually inspect the functional relationships between traffic counts and various meteorological and non-meteorological predictor variables at all 1400 traffic stations. Therefore we apply an automatic

procedure that selects relevant predictor variables based on objective criteria and allows the evaluation of the benefit of including meteorological variables compared to a model without meteorological variables. The following three steps are applied successively for each traffic station and for each of the four vehicle types.

3.4.1 Step 1: Breakpoint detection

Breakpoints are detected in the traffic count time series as described above to identify systematic changes of traffic characteristics, e. g. due to modifications of the road network in the vicinity of a station. Although an efficient algorithm is used for estimating the locations of breakpoints, a considerable computational effort is required for long time series like the hourly traffic counts used in this study. Furthermore, since the breakpoint detection is based on linear regression, the method assumes that residual errors are normally distributed. However, this is not the case due to the nature of the count data. Both issues are solved by applying the breakpoint detection to daily instead of hourly sums of traffic counts. Firstly, the amount of data is reduced significantly. Secondly, by aggregating the data the distribution of the residual errors becomes approximately normal, which we tested using the Shapiro-Wilk test and the Anderson-Darling test [45]. The month of the year and the day of the week are included as categorical predictor variables in Eq. 2 to account for an annual and weekly cycle of traffic counts. The minimum length of a segment n_h is set to 300 days to avoid too many and too short segmentations. The number of breakpoints m is selected by iteratively increasing it from 0 to 4. If an increase of m does not improve the RSS by more than 1%, the iteration is stopped and m is selected. Finally, a categorical variable with hourly resolution is generated, in which each segment corresponds to one category, based on the identified breakpoints. This variable is included in the model selection process described below. Note that the daily traffic data is only used to determine the dates of the breakpoints and that the following modeling steps are carried out with hourly data.

3.4.2 Step 2: Model without meteorological variables

After the identification of breakpoints based on daily aggregated traffic counts, Poisson regression models for hourly traffic counts are estimated. The BAST uses daily, weekly and annual cycles to classify the characteristics of individual traffic stations and distinguishes between periods with and without holidays [46]. We adopt this approach to develop a model NO_MET using only non-meteorological predictor variables and relevant interaction terms (see Table 2). NO_MET is used

Table 2 Description of predictor variables and interaction terms included in models without (NO_MET) and with (MET) meteorological variables

Variable name	Variable type	Variable description
<i>Potential predictor variables for NO_MET</i>		
hour	Categorical (24)	Hour of the day
dow	Categorical (7)	Day of the week (public holidays are treated as Sundays)
mon	Categorical (12)	Month of the year
holiday	Categorical (2)	School holiday in the federal state of the traffic station
trend	Continuous	Linear trend in time
hour:dow	Interaction	Different diurnal cycles on different days of the week
break	Categorical (n)	$n = 2, 3, 4$ segments determined by breakpoint detection (if $n = 1$ this term and its interactions are excluded)
break:trend	Interaction	Different trends in different segments
break:hour	Interaction	Different diurnal cycle in different segments
<i>Potential predictor variables for MET</i>		
temp ^{k}	Continuous	Daily maximum temperature at 2 m height at the ERA5 grid cell closest to the traffic station
cloud ^{k}	Continuous	Daily mean total cloud cover at the ERA5 grid cell closest to the traffic station
wind ^{k}	Continuous	Daily maximum wind gust at 10 m height at the ERA5 grid cell closest to the traffic station
precip ^{$1/k$}	Continuous	Average hourly precipitation sum of all RADOLAN grid cells within a radius of 10 km around traffic station
weekend	Categorical (3)	Distinguish between working day, Saturday, and Sunday (only included in interaction terms below, but not as single variable)
weekend:temp ^{k}	Interaction	Different relationships between temperature and traffic
weekend:precip ^{$1/k$}	Interaction	Different relationships between precipitation and traffic
weekend:clt ^{k}	Interaction	Different relationships between cloud cover and traffic
weekend:wind ^{k}	Interaction	Different relationships between wind speed and traffic

For categorical variables the number of categories is shown in brackets. Meteorological variables are introduced to the model selection process with different exponents $k = \{1, 2, 3, 4\}$ to allow non-linear functional relationships

as a benchmark to quantify the improvement achieved by including meteorological variables later. Predictor variables are added to NO_MET in a step-wise procedure. Starting with an intercept-only model, all remaining non-meteorological variables and interactions are added to the model individually and the MSESS is computed using 10-fold cross-validation with random samples. The variable that leads to the largest improvement with respect to the MSESS is added to the model, if the MSESS is larger than 0.01, indicating the reduction of the MSE of more than 1%. The iteration is repeated with all remaining variables. If the MSESS is smaller than or equal to 0.01 the iteration is stopped.

3.4.3 Step 3: Model with meteorological variables

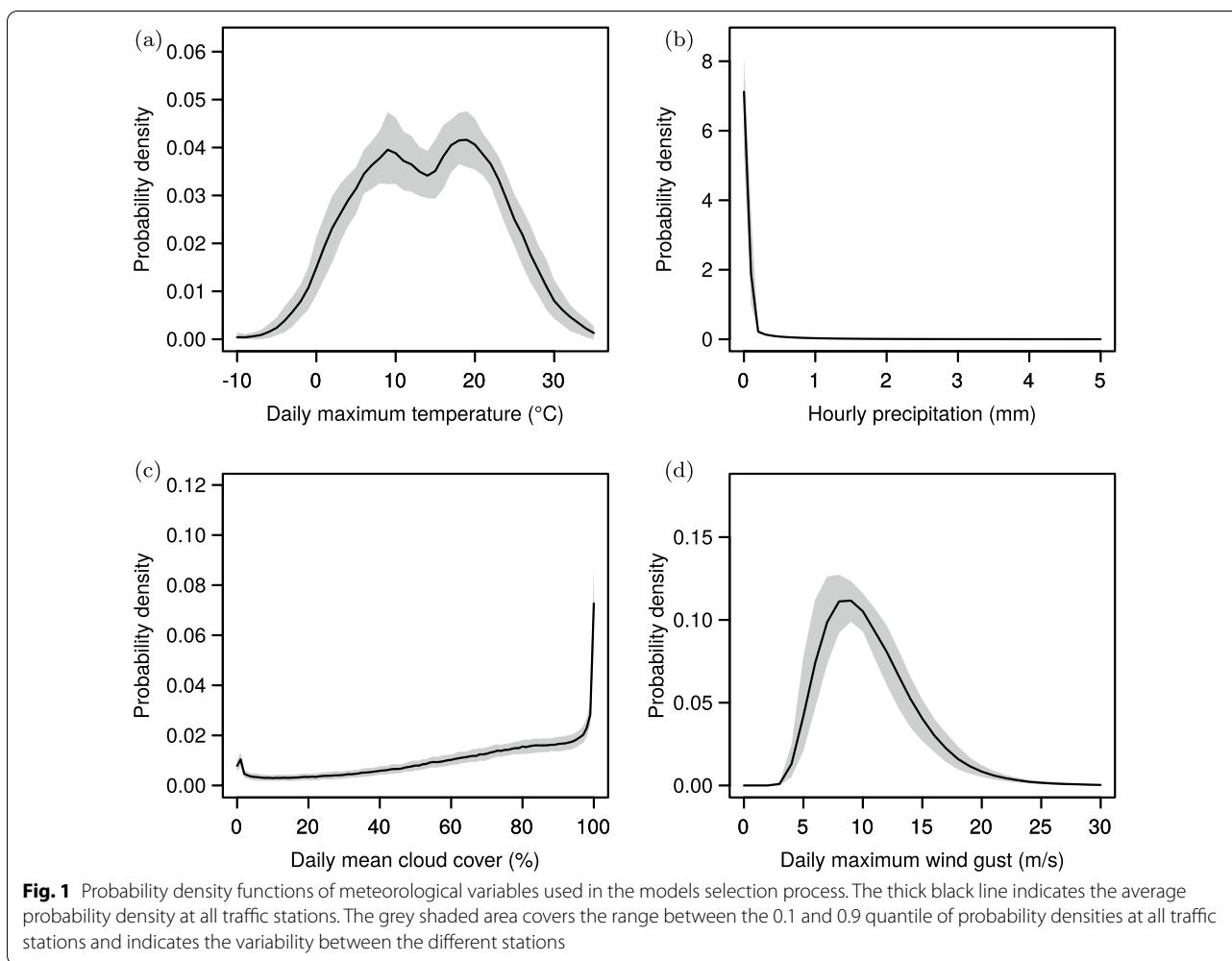
The iterative model selection procedure is repeated as described in step 2, but this time starting with the model NO_MET and iteratively adding meteorological predictor variables (Table 2). This model (MET) is used to quantify the improvement of traffic count predictions by including meteorological variables compared to NO_MET using the MSESS. To allow non-linear functional relationships between meteorological predictors and

traffic counts, the meteorological variables are considered in the selection procedure with different exponents. Temperature, cloud cover and wind are considered with exponents k and precipitation with the exponents $1/k$, with $k = \{1, 2, 3, 4\}$. In case of precipitation, the fraction allows for a sudden increase or decrease of crash counts with onsetting precipitation. This has already been successfully applied in a previous study for modeling the relationship between precipitation and road crash probabilities [47]. Additionally, each meteorological variables is included in the selection procedure as an interaction term with the categorical variable *weekend*, which has the three categories *working day* (Monday to Friday), *Saturday* and *Sunday*. This allows, for example, that precipitation can have a different effect on traffic on a Sunday, compared to a working day.

4 Results

4.1 Statistics of meteorological variables

Before studying the effect of meteorological parameters on traffic volume, the occurrence frequencies and correlations of the meteorological parameters is analyzed. For each traffic station the probability density function



of each meteorological variables is computed. The probability density function of daily maximum temperature, averaged over all stations, shows that temperature varies mainly between 0 and 30°C (Fig. 1a). In case hourly precipitation the distribution is strongly skewed towards low values (Fig. 1b). On average 71% of all hourly time steps show a precipitation of 0 mm, 19% show a precipitation between 0 and 0.1 mm and the remaining 10% correspond to precipitation amounts above 0.1 mm. The probability density for mean daily cloud cover is highest at cloud covers of 100%. Days with lower cloud covers are less frequent (Fig. 1c). Daily maximum wind gusts occur most frequently within the range between 5 and 20 m/s. The probability density function of temperature and wind gusts vary considerably between the different stations, while in case of precipitation and cloud cover the variability between the stations is much smaller.

To determine the strength and direction of potentially non-linear and monotonous relationships between the different meteorological variables, Spearman’s rank-order

correlations [48] are computed for each combination of the four variables at each traffic station (Table 3). This step is important to be aware of potential multicollinearity when estimating regression models. The strongest correlation of -0.37 is found between daily maximum temperature and daily mean cloud cover, indicating that low cloud cover correlates with high temperatures. Furthermore, positive correlations around 0.2 are found between cloud cover and precipitation, as well as between daily maximum wind gusts and precipitation and cloud cover. These correlations are reasonable and physically meaningful, however, they are small enough to justify the use of all three variables in the model selection process.

4.2 Selection of predictor variables

The model selection procedure described above is executed to develop models for hourly counts of different vehicle types at each traffic station by identifying those variables and interaction terms that improve the predictive skill of the model. Table 4 shows how frequently the

variables defined in Table 2 are selected at the different stations. In case of all vehicle types, *hour*, *dow* (day of the week) and *month* are selected at 100% of the stations. *hour* and *dow* are always selected as an interaction term, indicating that the diurnal cycle of traffic counts changes on different days of the week. The variable *break*, which indicates breakpoints, is selected as an interaction with a linear trend or *hour* in most cases.

In case of motorbikes and cars, temperature is selected at almost all stations, both on highways and federal roads. In case of vans, temperature is selected twice as often on highways than on federal roads. In most cases, temperature is selected as an interaction term with weekend, indicating that the effect of temperature on traffic counts is different on working days, Saturdays and Sundays. Mostly temperature is selected as a linear term without an exponent. However, in case of motorbikes also higher order terms are selected, indicating a more complex functional relationship.

Cloud cover seems to have an important effect on motorbike counts, in particular on federal roads. But also in case of cars on federal roads, cloud cover is selected at 45% of all stations. Wind speed and precipitation are selected at the majority of federal road station in case of motorbikes, but not in case of cars. In case of trucks, meteorological variables are rarely selected.

4.3 Skill scores

For each vehicle type at each traffic station the cross-validated MESS of MET is computed, with NO_MET as the reference. The MESS quantifies the improvement of the model predictions that results from including meteorological predictor variables in the regression models. It should be noted that due to the setup of the model selection procedure no negative MESS values occur, because predictors are only added to the model, if they improve the MESS. The largest improvements due to meteorological variables occur in models for motorbike counts on federal roads with a median MESS of 0.35, which corresponds to a reduction of the MSE by 35% compared to a model without meteorological variables (Fig. 2). At 25% of all federal road stations the MESS for motorbike counts is larger than 0.42, which constitutes a considerable reduction of the model error due to the inclusion of meteorological information in the model. The median MESS, and thus the improvement against the model without meteorological variables, is about 3 times larger on federal roads than on highways. The median MESS of car counts is 0.04, which is considerably smaller than the MESS of motorbikes. However, at individual stations MESS of cars values reach more than 0.3. For vans on federal highways the MESS are almost as large as for cars. For vans on federal roads and for truck in general

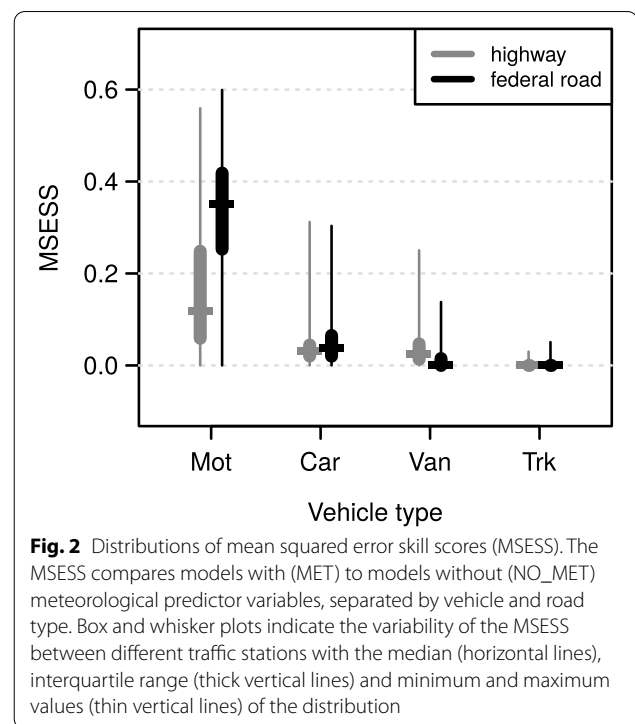


Fig. 2 Distributions of mean squared error skill scores (MSESS). The MESS compares models with (MET) to models without (NO_MET) meteorological predictor variables, separated by vehicle and road type. Box and whisker plots indicate the variability of the MESS between different traffic stations with the median (horizontal lines), interquartile range (thick vertical lines) and minimum and maximum values (thin vertical lines) of the distribution

the improvements due to weather predictors is zero or negligibly small, which is consistent with the previous observation that in most cases no meteorological predictor variables were added to these models.

The spatial distribution of the MESS values of motorbike counts shows that in case of stations on federal highways the largest MESS occur in areas with high population density, like Berlin, Munich or the Ruhr area, Cologne and Bonn (Fig. 3a). On federal roads the spatial distribution is more homogeneous (Fig. 3b). In case of cars most stations show a relatively low MESS, but some stations with considerably larger MESS values stand out, which are closely linked to touristic regions. For example, MESS values of more than 0.2 are found on routes from cities like Hamburg and Bremen towards seaside resorts at the North Sea and Baltic Sea (Fig. 3c, d). The largest MESS for cars of about 0.3 is found on the highway from Munich towards touristic areas in the Bavarian Alps (Fig. 3c). In case of cars on federal roads, stations with large MESS values are located at roads leading to recreation areas and nature parks like Sauerland, Eifel, Swabian Alb and Franconian Switzerland (Fig. 3d).

The visual inspection of the spatial distribution of MESS values of models for motorbike counts indicated a larger relevance of weather in densely populated areas. To quantify this relationship, the Spearman correlations between the MESS values and the population density within a radius of 10 km around the specific

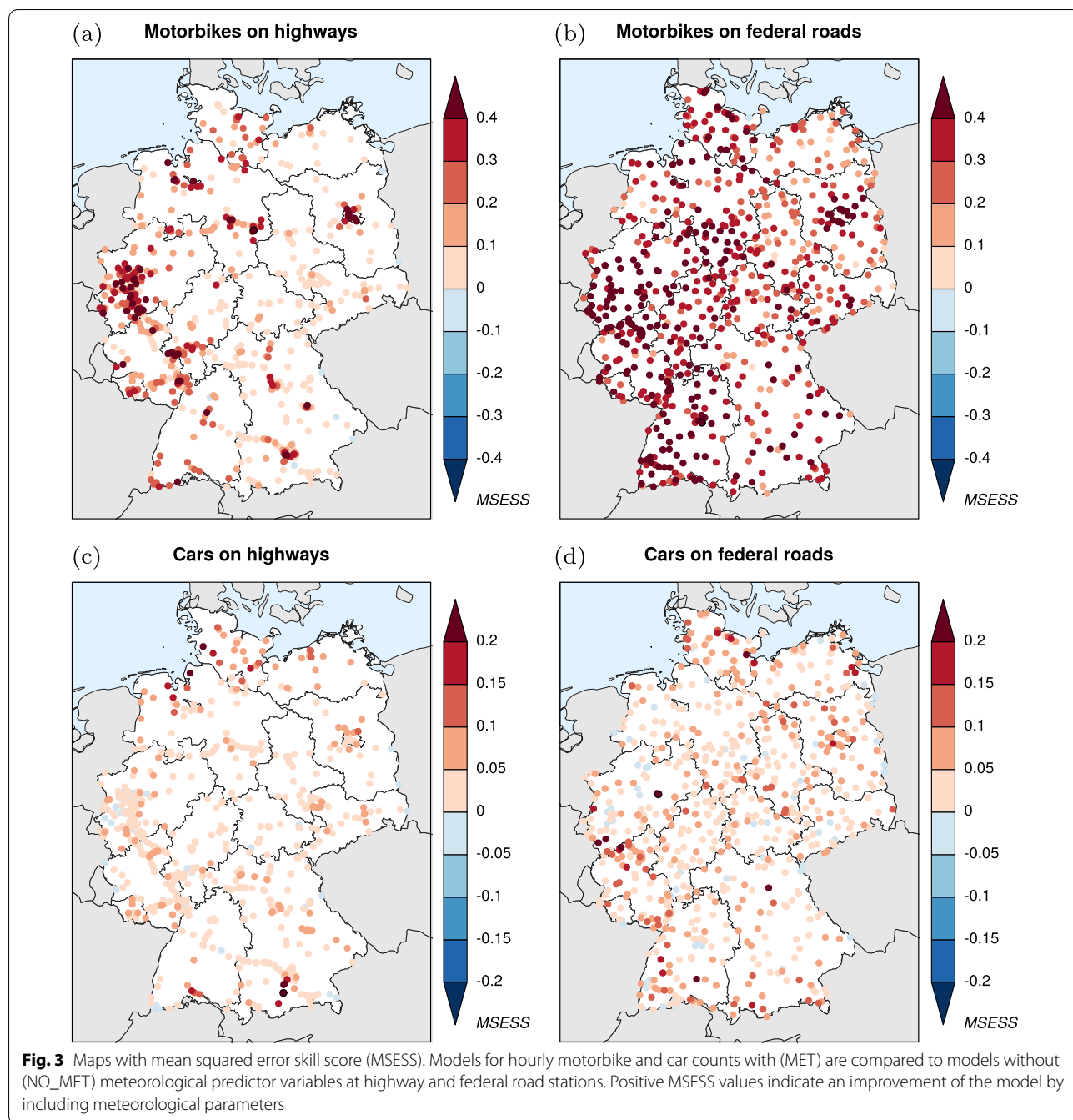


Table 3 Spearman correlations between different meteorological parameters

	Temp	Precip	Cloud	Wind
Temp	1.00			
Precip	-0.03	1.00		
Cloud	-0.37	0.22	1.00	
Wind	-0.07	0.21	0.19	1.00

The Spearman correlations between different meteorological parameters are calculated for each traffic station. The table shows the average correlations of all traffic stations

traffic stations is computed (Table 5). The largest correlation of 0.53 is found in case of motorbikes on highways, indicating that in regions with high population densities meteorological predictor variables lead to the largest improvement of models for motorbike counts. In case of cars the correlations are smaller in magnitude and negative, indicating that in regions with low population densities meteorological predictor variables improve the prediction of car counts.

Table 4 Fraction of traffic stations (in percent) for which a predictor variable or interaction term is selected by the model selection procedure

Predictor variable	Highway				Federal road			
	Mot	Car	Van	Trk	Mot	Car	Van	Trk
hour	0	0	0	0	0	0	0	0
dow	0	0	0	0	0	0	0	0
mon	100	100	100	100	100	100	100	100
hour:dow	100	100	100	100	100	100	100	100
holiday	6	54	74	99	1	78	79	91
trend	13	38	12	58	19	23	9	27
break	2	4	11	11	1	6	8	4
break:trend	36	32	79	27	10	44	70	38
break:hour	52	33	70	23	14	46	74	48
temp	95	89	76	8	93	86	38	16
temp	33	0	0	0	55	0	0	1
temp ²	3	4	1	0	7	0	0	0
temp ³	7	0	0	0	17	0	0	0
temp ⁴	15	0	0	0	14	1	0	0
weekend:temp	59	82	54	5	39	82	29	13
weekend:temp ²	10	18	30	2	18	11	9	3
weekend:temp ³	7	3	1	1	23	8	1	0
weekend:temp ⁴	23	4	1	0	17	18	1	0
cloud	68	10	1	0	99	45	4	0
cloud	3	0	0	0	3	0	0	0
cloud ²	9	1	0	0	25	0	0	0
cloud ³	1	0	0	0	3	0	0	0
cloud ⁴	0	0	0	0	1	0	0	0
weekend:cloud	18	5	1	0	9	13	2	0
weekend:cloud ²	31	3	0	0	50	15	1	0
weekend:cloud ³	5	1	0	0	10	10	1	0
weekend:cloud ⁴	1	0	0	0	2	7	0	0
wind	52	7	4	4	89	15	3	1
wind	1	0	0	0	2	0	0	0
wind ²	3	0	0	0	10	0	0	0
wind ³	1	0	0	0	6	0	0	0
wind ⁴	0	0	0	0	0	0	0	0
weekend:wind	24	6	4	4	7	9	3	0
weekend:wind ²	14	1	1	0	27	4	0	0
weekend:wind ³	9	0	0	0	30	2	0	0
weekend:wind ⁴	1	0	0	0	7	0	0	0
precip	40	5	25	0	90	1	2	0
precip	0	0	0	0	0	0	0	0
precip ^(1/2)	1	0	0	0	1	0	0	0
precip ^(1/3)	5	0	1	0	3	0	0	0
precip ^(1/4)	3	0	0	0	9	0	0	0
weekend:precip	0	0	0	0	0	0	0	0
weekend:precip ^(1/2)	8	0	8	0	11	0	1	0
weekend:precip ^(1/3)	11	1	14	0	23	0	1	0
weekend:precip ^(1/4)	9	3	1	0	43	1	0	0

Table 4 (continued)

Note that for one station a meteorological variable can be selected multiple times, e.g. with different exponents. Rows with italic font show the fraction of stations for which a specific meteorological variable was selected at least once with any exponent or in an interaction term

Table 5 Spearman correlation between MSESS values of models with meteorological predictor variables and the population density within a radius of 10 km around the specific traffic stations

Vehicle type	Highways	Federal roads
Mot	0.53	0.37
Pkw	- 0.12	- 0.17
Lfw	- 0.05	- 0.16
Lkw	- 0.10	0.02

For a more detailed analysis of weather impacts on model performance, the cross-validated MSESS is computed separately for the hours of the day, the days of the week and the months of the year. In case of motorbike counts on federal roads the largest MSESS values occur during daytime, in particular in the afternoon hours, where median MSESS values of almost 0.4 are reached (Fig. 4). Between 0 and 5 AM the MSESS show almost no improvement, at some stations even negative MSESS. Furthermore, on Saturdays and Sundays the MSESS values are generally higher than on workdays. The largest improvements in the course of the year is found during the transitional seasons, in particular in March and October, where the median MSESS reach almost 0.5. In contrast, in the winter months December and January the median MSESS values are almost zero. This is likely due to the effect, that in winter the conditions for motorbiking are generally bad due to low temperatures and therefore the addition of weather predictors to the model brings no benefit compared to simply providing the the information of climatological low temperatures by using the month of the year. However, in the transitional months the weather can change frequently between fair and adverse conditions and climatology given by the month of the year is not a good predictor. Thus, the availability of weather predictors in the models is beneficial to differentiate between these situations. In case of motorbike counts on federal highways the patterns are similar, but the MSESS values are smaller compared to federal roads.

In case of models for car counts, the MSESS values are again smaller than for motorbikes. However, at some stations a considerable improvement is evident during weekends and in the afternoon, with maximum MSESS values of more than 0.4 (Fig. 5). An interesting difference compared to motorbikes are the relatively high MSESS values in January and low values in April. It could play

Table 6 Setting of predictor variables used for predictions of hourly vehicle counts displayed in Fig. 6

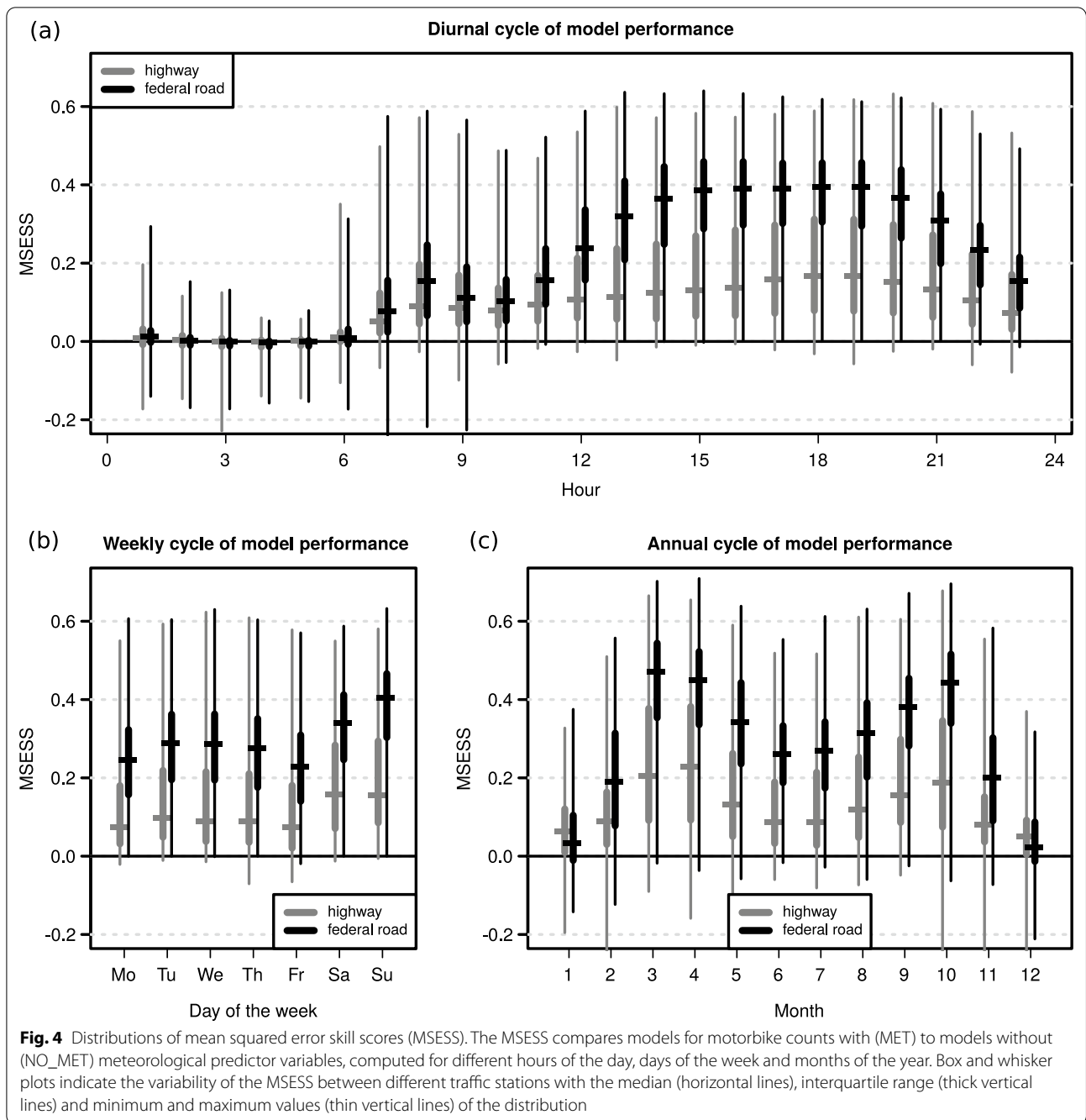
Variable	Value
Hour	12 PM
Month	June
Holiday	no
Trend	First day of time series
Breakpoint	First segment
Precip	0 mm/h
Temp	25°C
Cloud	50%
Wind	10 m/s

a role here that a car, as a sheltered mode of transport, can be easily used at low temperatures and otherwise fair weather conditions. Motorbike rides at low temperatures, however, might be unpleasant, or seasonal licenses, which are common in Germany, might prohibit the use of motorbikes in winter.

4.4 Functional relationships

The iterative predictor selection procedure chooses from a set of relevant meteorological parameters with different exponents. This allows non-linear functional relationships between the meteorological predictor variables and traffic flow. To study these functional relationships, one specific meteorological predictor variable is varied, while all other variables are held constant (see Table 6 for details). The variables are chosen to represent weather situations typical for the summer season. This is done separately for Mondays, Saturdays and Sundays to assess the differences between the functional relationships on working days and weekends. Tuesday to Friday are comparable to Mondays and are therefore not shown here. To compare the model predictions of traffic counts at the different stations, the traffic counts are rescaled, so that 0 and 1 correspond to the average daily minimum and maximum hourly traffic flow at the specific station. For visualization of the functional relationships the modeled rescaled traffic counts of all stations are averaged (thick colored lines in Fig. 6). Additionally the 0.1 and 0.9 quantiles are computed to show the variability between the different stations (shaded areas in Fig. 6).

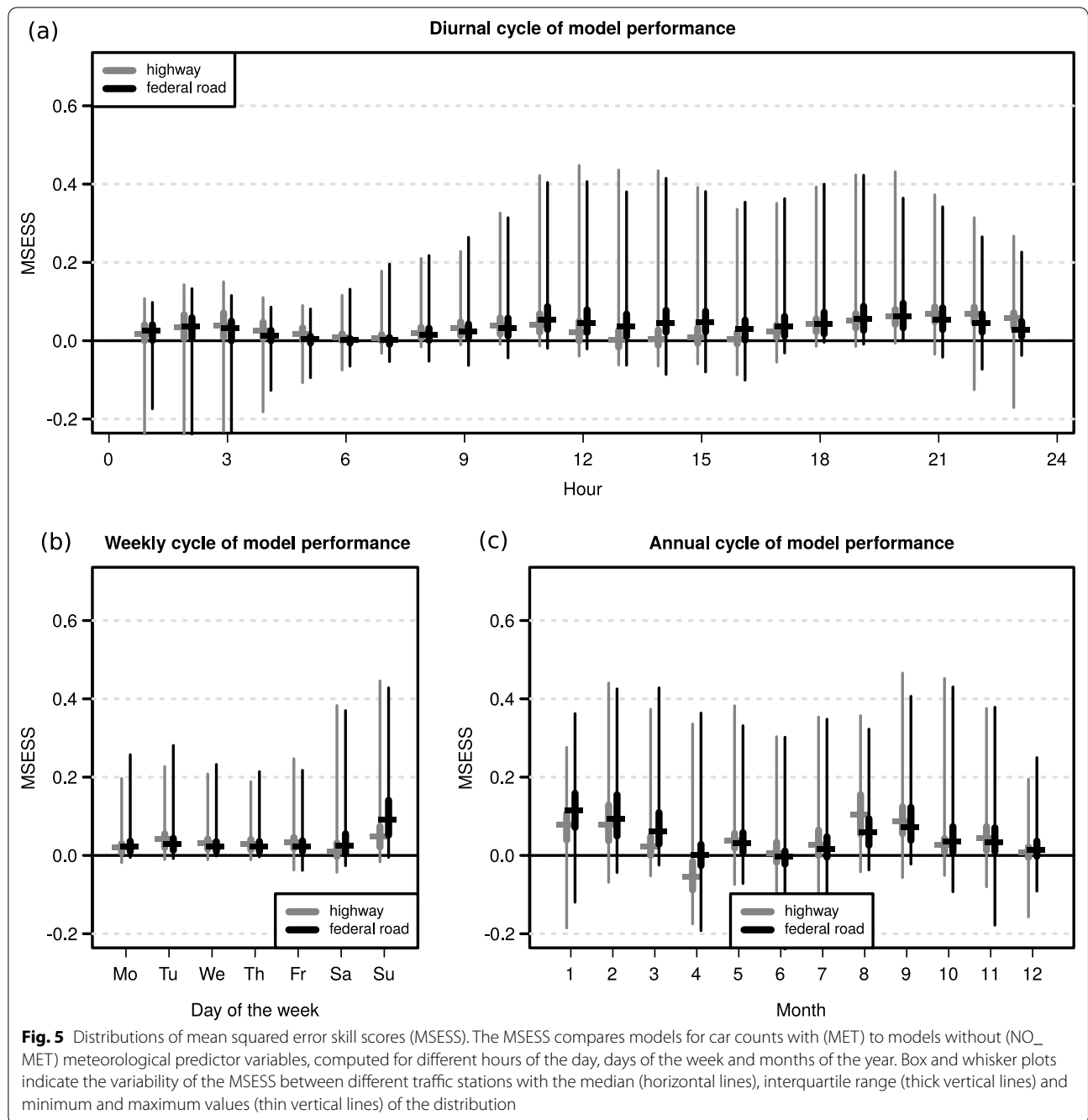
In case of motorbike counts on federal roads, the station-average traffic counts are highest on Sundays,



followed by Saturdays and Mondays (Fig. 6a–d). This indicates that motorbikes are often used for leisure activities. The traffic flow as a function of daily maximum temperature shows that motorbike counts decrease strongly at lower temperatures (Fig. 6a). The maximum is reached at about 25°C. Higher temperatures lead to a reduction of motorbike counts.

Motorbike counts as a function of hourly precipitation show highest values when there is no precipitation

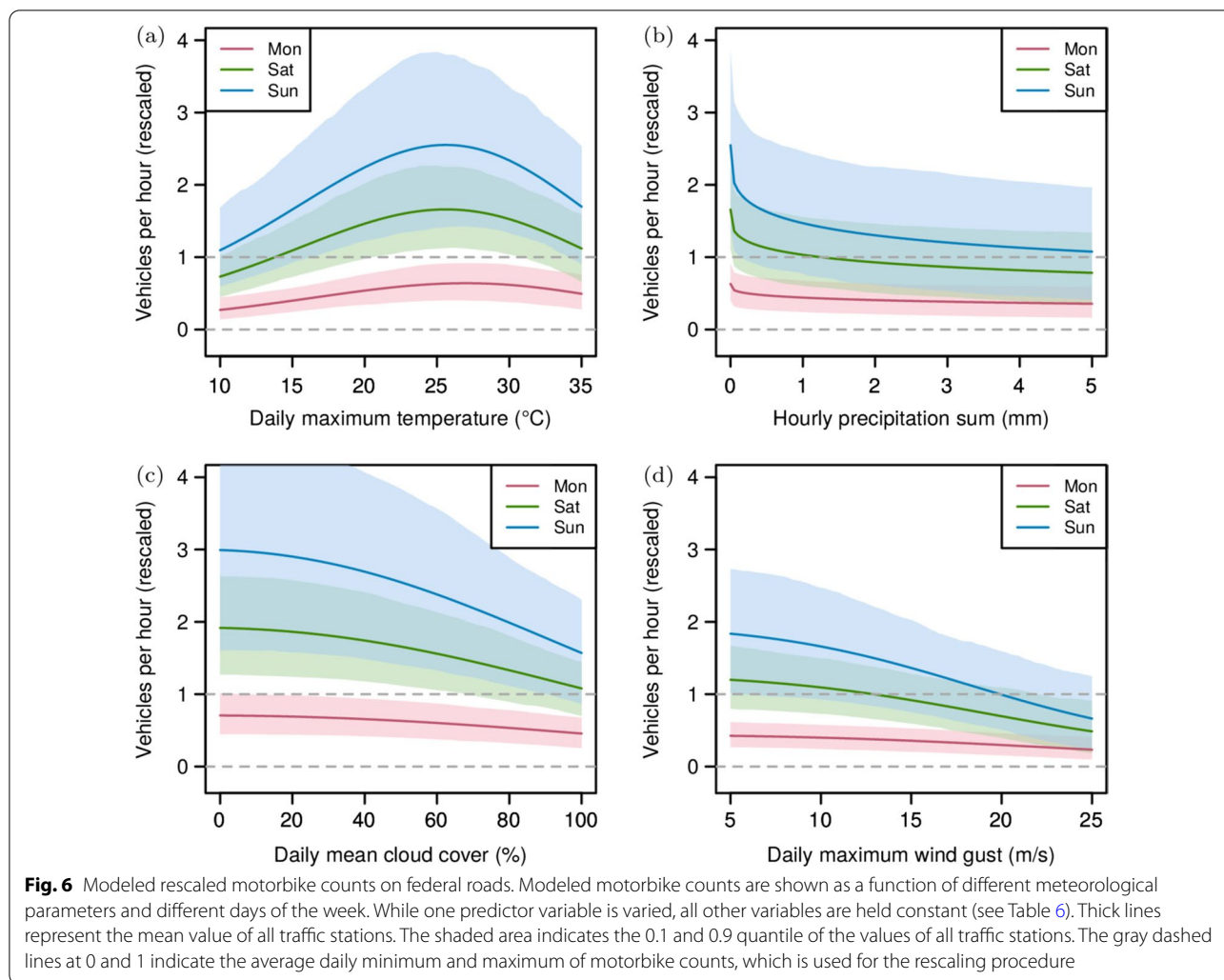
(Fig. 6b). During hours without precipitation motorbike counts are almost 5 times larger on Sundays compared to Mondays. An increase of hourly precipitation leads to a sudden drop in motorbike counts and an almost asymptotic flattening of the curve where precipitation exceeds 2 mm/h. Precipitation of 2 mm/h leads to reduction of motorbike counts by 50% compared to hours without precipitation. This reasonable non-linear functional relationship between precipitation and motorbike counts is



established by using $1/k$ as an exponent for precipitation, with $k = \{1, 2, 3, 4\}$ (see Table 2). One could argue that a sharp break between “no precipitation” and “precipitation”, which could be introduced by using a binary variable, would be more appropriate. However, the smooth transition better reflects the uncertainties related to the precipitation data and the model formulation. For example, due to the lack of an unambiguous relationship

between radar echo and the actual precipitation amount, RADLOAN data may show precipitation, although there was no precipitation on the ground. Also a potential time-lagged effect of onsetting precipitation on motorbike counts is not included in the model.

The relationship between motorbike counts and daily average cloud cover reveals particularly large motorbike counts in cloud-free situations and a reduction of motorbike counts with increasing cloudiness (Fig. 6c).



On cloud-free Sundays the motorbike counts are almost twice as large as on cloudy Sundays. However, one should be aware of the correlation between cloud cover and precipitation and between cloud cover and temperature, which could affect the results. Furthermore, the variability between the different traffic stations is relatively large at low cloud covers.

Motorbikes are especially vulnerable to strong wind speeds and one can expect that motorcyclists avoid trips under windy conditions. This is also reflected by the models. Increasing daily maximum wind gusts lead to a strong reduction of motorbike counts (Fig. 6d). Extreme wind gusts of more than 25 m/s lead to the lowest motorbike counts, also when compared to the effects of the other meteorological parameters. Such wind speeds occur, for example, in summer during thunder storms or in winter in conjunction with extratropical cyclones.

5 Discussion

While previous studies have addressed weather impacts bicycle, car or truck traffic, there was little research on the direct effect of weather on motorcycle traffic. This study presents evidence that motorcycles, as a non-sheltered mode of transport, is strongly depending on weather conditions. The findings that motorcycle flow increases with temperature and decreases with precipitation is in line with a number of studies addressing bicycle travel behavior [5, 6] and outdoor activities in general [2, 3]. Cloudiness and wind speed are mostly not considered in studies of traffic and outdoor activity. We showed that low cloud cover and low wind speeds coincide with a higher motorbike traffic flow, which is in line with the general findings that fair weather increases open-air activity [1]. We could also show that high temperatures above 25°C lead to a decline in motorbike counts, which is similar to bicycle usage [8, 9] and outdoor activities in general [7].

Results of previous studies, which have analyzed individual traffic stations, suggested that traffic flow in recreational areas is more dependent on weather compared to urban areas [17]. By analyzing a large number of stations and specific vehicle types, we can confirm that this is particularly true for car traffic. At the majority of traffic stations we found that the improvement of prediction of hourly car counts by including meteorological variables is small. However, traffic stations along routes towards seaside resorts and nature parks showed a substantial improvement and thus a pronounced dependence on meteorological variables.

As suggested by previous research [1], we established non-linear relationships between meteorological predictor variables and traffic flow by choosing from meteorological predictor variables with different exponents in an automatic selection procedure. Another methodology specifically designed to describe non-linear relationships are generalized additive models (GAMs), which have been applied for example to predict crash frequencies [49, 50]. GAMs use smooth function like cubic splines to find the optimum functional relationships between predictor and target variable. As a test we have also applied GAMs to our data and found that it leads to unrealistic behavior at the extreme ends of the distributions. Also the strong drop of traffic flow with onsetting precipitation lead to considerable overshooting behavior of the splines. It appears to be unsuitable to apply GAMs in an automatic procedure to a large number of stations, where a detailed evaluation of each individual model is infeasible. However, it may be suitable to apply GAMs to individual traffic stations in a detailed study, where fine-tuning of the model is possible.

Böcker et al. [1] suggested to consider interactions between different meteorological variables. For example, the impact of wind speed on motorbike counts may be different on days with precipitation compared to days without. Under rainy conditions motorcyclist already refrain from making trips, so that additional strong wind speeds make no difference. We have included the interaction of precipitation as a categorical variable with the other meteorological predictor variables. However, in general no major improvement of the model was found. The changes of the MSE less than 1% in most cases. Therefore the results were not included in this paper. Due to the increasing complexity of the models when using interactions, future research in this direction could focus more on individual stations, which have been shown do be strongly affected by weather.

The Poisson regression model assumes an equality of mean and variance of the count data. In our case this

assumption does not hold due to an overdispersion of the data. We have tested if the use of a negative binomial regression model would lead to an improvement of the predictive skill, but that was not the case. Instead, the predictive skill decreased, in particular at hours with high traffic volume. Therefore, we decided to use the Poisson model, which is acceptable, because the overdispersion mainly the estimation of standard errors, which were not the focus of this study.

6 Conclusions

We have shown that the use of meteorological predictor variables can substantially increase the predictive skill of models for hourly traffic flow, although the magnitude of the improvement depends strongly on vehicle type and location of the traffic station. A particular result was that motorbike counts are strongly weather-dependent and showed a highly non-linear relationship to the meteorological variables. Mean squared errors of motorbike counts could be reduced by up to 60% by including meteorological variables in the models. This is reasonable, since motorbikes are a non-sheltered transportation mode, frequently used for leisure activities and less frequently for commercial purposes. In case of cars the analysis showed mixed results. As a sheltered mode of transportation, which is used for commuting, leisure activities as well as commercial purposes, car counts showed the tendency to be less sensitive to weather in urban areas, but strongly weather dependent in touristic regions like seaside resorts and nature parks. Lastly, counts of delivery vans and trucks, which are mainly used for commercial purposes, showed only low weather dependence.

These findings open up several potential applications of such models. First, analyses of weather impacts on crash probabilities can be improved by including weather-related variation of traffic flow as a predictor variable. Second, taking into account weather effects in traffic flow predictions could improve route planning in navigation systems and could assist in traffic management systems to compensate or redistribute high traffic volumes, in particular in touristic regions. Furthermore, prediction of traffic volumes taking into account weather forecasts would allow day-ahead planning of visitor numbers in touristic areas.

Abbreviations

ECMWF: European Centre for Medium-Range Weather Forecasts; ERA5: European Centre for Medium-Range Weather Forecasts Reanalysis v5; GAM: Generalized additive model; MET: Model with meteorological predictor variables; MSE: Mean squared error; MSESS: Mean squared error skill score; NO_MET: Model without meteorological predictor variables; RSS: Residual sum of squares.

Acknowledgements

This research was carried out within the framework of the Hans-Ertel-Centre for Weather Research [51]. This research network of universities, research institutes and the Deutscher Wetterdienst is funded by the Bundesministerium für Verkehr und Digitale Infrastruktur. We would like to thank the HPC service of ZEDAT at Freie Universität Berlin for the computing time and assistance provided.

Author contributions

Data analysis and visualization was done by Nico Becker; all authors contributed to writing the manuscript. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This research has been supported by the Bundesministerium für Verkehr und Digitale Infrastruktur (grant no. 4818DWDP3A).

Availability of data and materials

The hourly traffic count data for highways and federal roads in Germany is available via the Bundesanstalt für Straßenwesen [52]. The RADOLAN data set is available via the German Weather Service [53]. The ERA5 reanalysis data is available at the Climate Data Store [54]. The gridded population density data for Germany is available at the website of the Zensus 2011 [55].

Declarations

Ethics approval and consent to participate

Not applicable.

Informed consent

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

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

Received: 16 September 2021 Accepted: 11 April 2022

Published online: 22 April 2022

References

- Böcker, L., Dijst, M., & Prillwitz, J. (2013). Impact of everyday weather on individual daily travel behaviours in perspective: A literature review. *Transport Reviews*, 33(1), 71–91. <https://doi.org/10.1080/01441647.2012.747114>.
- Dwyer, J. F. (1988). Predicting daily use of urban forest recreation sites. *Landscape and Urban Planning*, 15(1–2), 127–138. [https://doi.org/10.1016/0169-2046\(88\)90021-7](https://doi.org/10.1016/0169-2046(88)90021-7).
- Thorsson, S., Lindqvist, M., & Lindqvist, S. (2004). Thermal bioclimatic conditions and patterns of behaviour in an urban park in Göteborg, Sweden. *International Journal of Biometeorology*, 48(3), 149–156. <https://doi.org/10.1007/s00484-003-0189-8>.
- Aultman-Hall, L., Lane, D., & Lambert, R. R. (2009). Assessing impact of weather and season on pedestrian traffic volumes. *Transportation Research Record*, 2140(1), 35–43. <https://doi.org/10.3141/2140-04>.
- Bergström, A., & Magnusson, R. (2003). Potential of transferring car trips to bicycle during winter. *Transportation Research Part A: Policy and Practice*, 37(8), 649–666. [https://doi.org/10.1016/S0965-8564\(03\)00012-0](https://doi.org/10.1016/S0965-8564(03)00012-0).
- Nankervis, M. (1999). The effect of weather and climate on bicycle commuting. *Transportation Research Part A: Policy and Practice*, 33(6), 417–431. [https://doi.org/10.1016/S0965-8564\(98\)00022-6](https://doi.org/10.1016/S0965-8564(98)00022-6).
- Sharifi, E., & Boland, J. (2018). Limits of thermal adaptation in cities: Outdoor heat-activity dynamics in Sydney, Melbourne and Adelaide. *Architectural Science Review*, 61(4), 191–201. <https://doi.org/10.1080/00038628.2018.1482824>.
- Ahmed, F., Rose, G., & Jacob, C. (2010). Impact of weather on commuter cyclist behaviour and implications for climate change adaptation. In *Australasian transport research forum*. https://www.australasiantransportresearchforum.org.au/sites/default/files/2010_Ahmed_Rose_Jacob.pdf.
- Phung, J., & Rose, G. (2007). Temporal variations in usage of Melbourne's bike paths. In *Proceedings of 30th Australasian transport research forum, Melbourne*. https://www.australasiantransportresearchforum.org.au/sites/default/files/2007_Phung.pdf.
- Roh, H.-J. (2020). Assessing the effect of snowfall and cold temperature on a commuter highway traffic volume using several layers of statistical methods. *Transportation Engineering*, 2, 100022. <https://doi.org/10.1016/j.treng.2020.100022>.
- Tucker, P., & Gilliland, J. (2007). The effect of season and weather on physical activity: A systematic review. *Public Health*, 121(12), 909–922. <https://doi.org/10.1016/j.puhe.2007.04.009>.
- Chan, C. B., & Ryan, D. A. (2009). Assessing the effects of weather conditions on physical activity participation using objective measures. *International Journal of Environmental Research and Public Health*, 6(10), 2639–2654. <https://doi.org/10.3390/ijerph6102639>.
- Spinney, J. E., & Millward, H. (2011). Weather impacts on leisure activities in Halifax, Nova Scotia. *International Journal of Biometeorology*, 55(2), 133–145. <https://doi.org/10.1007/s00484-010-0319-z>.
- Al Hassan, Y., & Barker, D. J. (1999). The impact of unseasonable or extreme weather on traffic activity within Lothian region, Scotland. *Journal of Transport Geography*, 7(3), 209–213. [https://doi.org/10.1016/S0966-6923\(98\)00047-7](https://doi.org/10.1016/S0966-6923(98)00047-7).
- Changnon, S. A. (1996). Effects of summer precipitation on urban transportation. *Climatic Change*, 32(4), 481–494.
- Keay, K., & Simmonds, I. (2005). The association of rainfall and other weather variables with road traffic volume in Melbourne, Australia. *Accident Analysis & Prevention*, 37(1), 109–124. <https://doi.org/10.1016/j.aap.2004.07.005>.
- Cools, M., Moons, E., & Wets, G. (2010). Assessing the impact of weather on traffic intensity. *Weather, Climate, and Society*, 2(1), 60–68. <https://doi.org/10.1175/2009WCAS1014.1>.
- Call, D. A. (2011). The effect of snow on traffic counts in western New York state. *Weather, Climate, and Society*, 3(2), 71–75. <https://doi.org/10.1175/WCAS-D-10-05008.1>.
- Hanbali, R. M., & Kuemmel, D. A. (1993). Traffic volume reductions due to winter storm conditions. *Transportation Research Record* (1387).
- Knapp, K. K., & Smithson, L. D. (2000). Winter storm event volume impact analysis using multiple-source archived monitoring data. *Transportation Research Record*, 1700(1), 10–16. <https://doi.org/10.3141/1700-03>.
- Maze, T. H., Agarwal, M., & Burchett, G. (2006). Whether weather matters to traffic demand, traffic safety, and traffic operations and flow. *Transportation Research Record*, 1948(1), 170–176. <https://doi.org/10.1177/0361198106194800119>.
- Datla, S., & Sharma, S. (2010). Variation of impact of cold temperature and snowfall and their interaction on traffic volume. *Transportation Research Record*, 2169(1), 107–115. <https://doi.org/10.3141/2169-12>.
- Maze, T. H., Crum, M. R., & Burchett, G. (2005). An investigation of user costs and benefits of winter road closures. Technical Report 21, Iowa State University. https://lib.dr.iastate.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1019&context=intrans_reports.
- Sabir, M. (2011). Weather and travel behaviour. Ph.D. thesis, Vrije Universiteit Amsterdam. <https://research.vu.nl/en/publications/weather-and-travel-behaviour>.
- Khattak, A. J., & De Palma, A. (1997). The impact of adverse weather conditions on the propensity to change travel decisions: A survey of brussels commuters. *Transportation Research Part A: Policy and Practice*, 31(3), 181–203. [https://doi.org/10.1016/S0965-8564\(96\)00025-0](https://doi.org/10.1016/S0965-8564(96)00025-0).
- Nissen, K., Becker, N., Dähne, O., Rabe, M., Scheffler, J., Solle, M., & Ulbrich, U. (2020). How does weather affect the use of public transport in Berlin? *Environmental Research Letters*. <https://doi.org/10.1088/1748-9326/ab8ec3>.
- Aaheim, H. A., & Hauge, K. E. (2005). Impacts of climate change on travel habits: A national assessment based on individual choices. CICERO Report.

28. Flynn, B. S., Dana, G. S., Sears, J., & Aultman-Hall, L. (2012). Weather factor impacts on commuting to work by bicycle. *Preventive Medicine*, 54(2), 122–124. <https://doi.org/10.1016/j.ypmed.2011.11.002>.
29. Sathiaraj, D., Pankasem, T.-O., Wang, F., & Seedah, D. P. (2018). Data-driven analysis on the effects of extreme weather elements on traffic volume in Atlanta, GA, USA. *Computers, Environment and Urban Systems*, 72, 212–220. <https://doi.org/10.1016/j.compenurbysys.2018.06.012>.
30. Van den Bossche, F., Wets, G., & Brijs, T. (2005). Role of exposure in analysis of road accidents: a Belgian case study. *Transportation Research Record*, 1908(1), 96–103. <https://doi.org/10.3141/1908-12>.
31. Fridstrøm, L., Ifver, J., Ingebrigtsen, S., Kulmala, R., & Thomsen, L. K. (1995). Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. *Accident Analysis & Prevention*, 27(1), 1–20. [https://doi.org/10.1016/0001-4575\(94\)E0023-E](https://doi.org/10.1016/0001-4575(94)E0023-E).
32. Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., & Schepers, D., 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>.
33. 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 bodenniederschlagsstationen (ombrometer). Deutscher Wetterdienst, Hydrometeorologie 5.
34. Ämter des Bundes und der Länder, S. (2015). Zensus 2011, Datenangebot zum Zensusatlas - Klassifizierte Ergebnisse. Retrieved August 26, 2021, from www.zensus2011.de.
35. Wood, S.N. (2017). Generalized additive models: An introduction with R. Chapman and Hall/CRC, Boca Raton. <https://doi.org/10.1201/9781315370279>.
36. Bai, J. (1994). Least squares estimation of a shift in linear processes. *Journal of Time Series Analysis*, 15(5), 453–472.
37. Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory*. <https://doi.org/10.1017/S0266466600005831>.
38. Bai, J. (1997). Estimation of a change point in multiple regression models. *Review of Economics and Statistics*, 79(4), 551–563. <https://doi.org/10.1162/003465397557132>.
39. Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*. <https://doi.org/10.2307/2998540>.
40. Zeileis, A., Leisch, F., Hornik, K., & Kleiber, C. (2002). strucchange: An r package for structural change in linear regression models. *Journal of Statistical Software*, 7(2), 1–38. <https://doi.org/10.18637/jss.v007.i02>.
41. Zeileis, A., Kleiber, C., Krämer, W., & Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis*, 44(1–2), 109–123. [https://doi.org/10.1016/S0167-9473\(03\)00030-6](https://doi.org/10.1016/S0167-9473(03)00030-6).
42. Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), 1–22. <https://doi.org/10.1002/jae.659>.
43. Cox, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *Journal of Personality Assessment*, 91(2), 121–136. <https://doi.org/10.1080/00223890802634175>.
44. Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, 135(3), 370–384.
45. Razali, N. M., Wah, Y. B., et al. (2011). Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21–33.
46. Fitschen, A., & Nordmann, H. (2019). Verkehrsentwicklung auf bundesfernstraßen 2016, in: Berichte der bundesanstalt für straßenwesen. Technical Report 323, Bundesanstalt für Straßenwesen, Bergisch Gladbach, Germany. https://bast.opus.hbz-nrw.de/opus45-bast/frontdoor/deliver/index/docId/2311/file/V323_Internet+PDF.pdf.
47. 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>.
48. Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101.
49. Zhang, Y., Xie, Y., & Li, L. (2012). Crash frequency analysis of different types of urban roadway segments using generalized additive model. *Journal of Safety Research*, 43(2), 107–114. <https://doi.org/10.1016/j.jsr.2012.01.003>.
50. Khoda Bakhshi, A., & Ahmed, M. M. (2021). Real-time crash prediction for a long low-traffic volume corridor using corrected-impurity importance and semi-parametric generalized additive model. *Journal of Transportation Safety & Security*. <https://doi.org/10.1080/19439962.2021.1898069>.
51. Hans-Ertel-Zentrum für Wetterforschung: Hans-Ertel-Zentrum für Wetterforschung. Retrieved September, 08, 2021 from <https://www.hans-ertel-zentrum.de/en/index.html>.
52. Bundesanstalt für Straßenwesen: Automatische Zählstellen auf Autobahnen und Bundesstraßen. Retrieved September 08, 2021, from https://www.bast.de/BASSt_2017/DE/Verkehrstechnik/Fachthemen/v2-verkehrszählung/Stundenwerte.html.
53. Deutscher Wetterdienst: RADOLAN (Radar-Online-Aneichung): Analysen der Niederschlagshöhen aus radar- und stationsbasierten Messungen im Echtzeitbetrieb. Retrieved September 08, 2021, from <https://www.dwd.de/DE/leistungen/radolan/radolan.html>.
54. Copernicus: Climate Data Store. Retrieved September 08, 2021, from <https://cds.climate.copernicus.eu/>.
55. Statistische Ämter des Bundes und der Länder: ZENSUS2011 - Bevölkerung- und Wohnungszählung 2011. Retrieved September 08, 2021, from <https://www.zensus2011.de/>.

Publisher's Note

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

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)