How a Bayesian Hierarchical Model makes our roads safer

Project Report

Authors: Gururaj M. Desai, Eda D. Özdemir, Tim R. Schleicher
Supervisor: Prof. Dr. Burkhardt Funk

Course 82005000 "Probabilistic Modelling"
Institute of Information Systems
Leuphana Universität Lüneburg

July 31, 2019

Abstract

This report is the result of an in-depth project on the paper "A novel Bayesian hierarchical model for road safety hotspot prediction" by Fawcett, Thorpe, Matthews & Kremer (2017). The model introduced by Fawcett et al. intends to enable practitioners to take proactive road safety measures at specific sites by predicting the number of accidents and ranking them according to their safety hotspot potentials. In contrast to previous models, the confounding effects of regression-to-mean and trend are met. Aim of this study project was to analyze the model and to replicate it based on historical data of accident counts from the German city of Halle. As indicated in detail throughout the report, we were able to successfully replicate the model. With this report, we provide the reader with an overview of the model by Fawcett et al., discuss our approach to model replication, compare the results, and give an overall outlook including recommendations to Fawcett et al. It is useful for everyone who intends to study the potentials and implementations of Bayesian models as well as their practical applications.

Contents

List of Figures

1 Introduction

Advancements in data science enable unprecedented opportunities to increase road safety. In Germany alone, more than 3,000 people each year die in road accidents (Destatis 2018). Consequently, improving road safety is high on the political agenda. Although in 2018 57% of traffic deaths occurred on country roads, the recent rise of killed cyclists has put road safety in cities further in the spotlight (ibid.). A lower rate of casualties can be one of the core achievements of successful road safety measures. However, especially due to budget and time constraints, many local authorities struggle to implement these. A data-driven approach that predicts road safety hotspots without confounding issues could therefore support authorities to effectively and efficiently decrease road accidents.

Fawcett, Thorpe, Matthews & Kremer (2017) have developed a Bayesian Hierarchical Model that they proclaim to be well-suited to predict road safety hotspots and to be meeting the demands of both effectiveness and efficiency. Their model based on a past accident data is special as it accounts for "regression-to-mean (RTM) and trend" (ibid. 262). Other models, e.g. by Sacchi et al. (2015) who apply a multivariate Poisson-lognormal regression model to predict road safety hotspots in the city of Vancouver, do not handle these unwanted effects well. Also, the model by Fawcett et al. is said to have the "ability for more recent accident counts to inform model-based estimates of safety with greater precision" (ibid. 263). Some research (e.g. by Cheng & Washington 2008) already indicates that modeling road safety hotspots with a Bayesian approach seems to be a superior method compared to other approaches like accident rate rankings.

The report at hand is the result of an in-depth project on the paper by Fawcett et al. The authors of this report have not only closely analyzed the paper but have replicated the model based on accident counts in the Germany city of Halle, which also Fawcett et al. use to illustrate their model. The report is structured as follows: After introducing the overall challenges around road safety hotspot prediction in the next subsection, we introduce the data set and provide some descriptive analysis (Section 2). Afterwards (Section 3), we give an overview of the model developed by Fawcett et al. before then closely outlining our approach in replicating their model (Section 4). Later, we compare and evaluate the results (Section 5), provide an outlook (Section 6), and conclude (Section 7).

1.1 Challenges in Road Safety Hotspot Prediction

Road safety practitioners want to take measures at sites that are most likely to cause accidents. By using common accident prediction models, they encounter two central challenges: RTM and trend. The first refers to the problem that a recent peak in accidents at a site, e.g. in the previous year can either be due to a severe safety lack or due to a random increase. A random increase would lead to a decrease in accidents (regression to the mean) in the following year anyway—no matter whether measures were implemented or not. To invest the available budget efficiently, a practitioner would not want to treat sites with random increases in accidents because they are "inherently 'safe'" (ibid. 262) but would want to focus only on the sites with a true safety risk. A similar challenge is imposed on road safety practitioners by the problem of trend. Also here, if a measure has been implemented, accidents might decrease due to either the measure itself—or due to site-specific or global trends, i.e. that people drive slower in general.

Another challenge in road safety hotspot prediction is the problem that if historical data is available, commonly the different years in the past are given the same weight. However, also this can be misleading since especially if time series data is available, practitioners would prefer recent years to stronger inform decisions—a lot could have happened in the meantime. The model discussed in this report is said to counter the challenges introduced in this subsection.

2 The Data and Descriptive Analysis

2.1 The Data

To demonstrate the proposed Bayesian Hierarchical Model, an open-source road network of the German city of Halle is selected by Fawcett et al. In our endeavour to replicate their model, we also adopt the same dataset. The available data consists of accident counts at 734 sites in Halle for the years 2004–2012 inclusive, which are the so-called observations. We have historical points in time which are represented by $\{y_j(t); t = -7, ..., -1, 0; j = 1, ..., 734\}$ for the years 2004, ..., 2011 and a future time period $\{y_j(t); t = +1; j = 1, ..., 734\}$ for the year 2012. We keep the year 2012 for validating the model's predictions. For each site $j = 1$, ..., 734, we have covariates which include amongst others traffic volume, speed limit, or whether a respective site is located in an urban area.

2.2 Descriptive Analysis

To gain a deeper understanding of the available data from Halle, we deploy some descriptive statistics in the following. Figure 1 shows the distributions of yearly accident totals from 2004 to 2011. The below histograms show the skewed distributions to the left. Furthermore, the distributions of yearly accident totals already indicate a negative binomial distribution which we want to learn with the Bayesian model. The mean of yearly accident totals varies between 3.1 accidents per site in the year 2011 and 3.7 accidents per site in 2004, 2005, and 2007. The yearly maximum count of accidents at an individual site was measured highest with 52 accidents in 2007, whereas the years 2004 and 2008 indicate the lowest values for yearly maximum accidents at an individual site. The standard deviations for yearly accident totals range from 19.3 in 2010 to 25.7 in 2006.

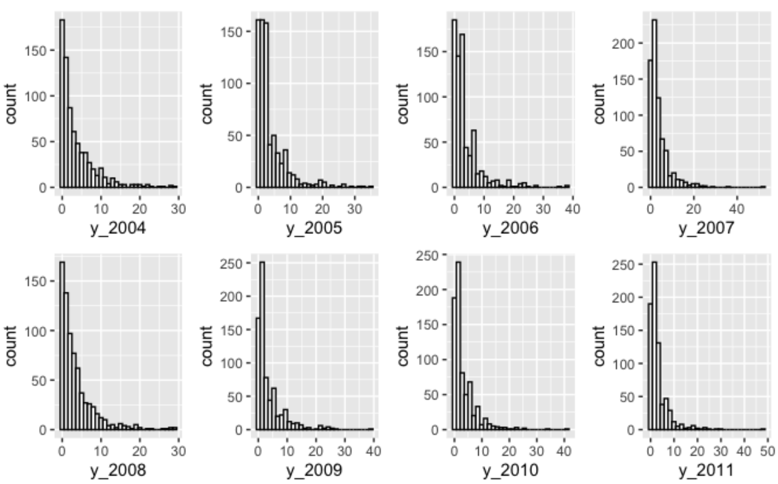

Figure 1: Distributions of yearly accident totals *y(t)*

Figure 2 gives an overview of all available covariates and their distributions. Understanding these covariates is important for our model in terms of their impact on accident rates. The mean of average daily vehicles passing through a site per year is 7,432, the maximum value is 65,484 vehicles. More than one-third of all sites have a speed limit of 30 km/h, only 7% of all sites have the highest possible speed limit of 80 km/h. Almost all sites are located in an urban area and have an intersection. Less than one third of all sites are signalised, but almost two-third is located at a major road. Every fifth site is at a major intersection and every fourth site is a so-called four-legged junction.

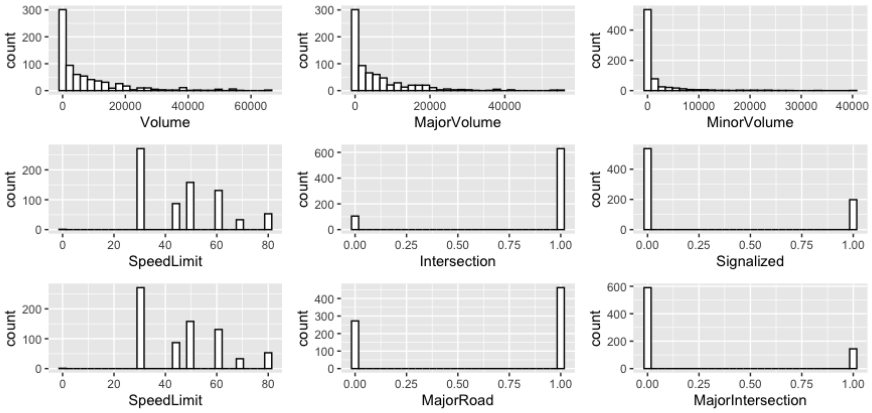

Figure 2: Distributions of the covariates

Figure 3 gives the time series plots of accidents at four selected sites. The four sites (163, 309, 677, 706) have been chosen by Fawcett et al. to indicate the effects of RTM and trend we aim to tackle in our modeling.

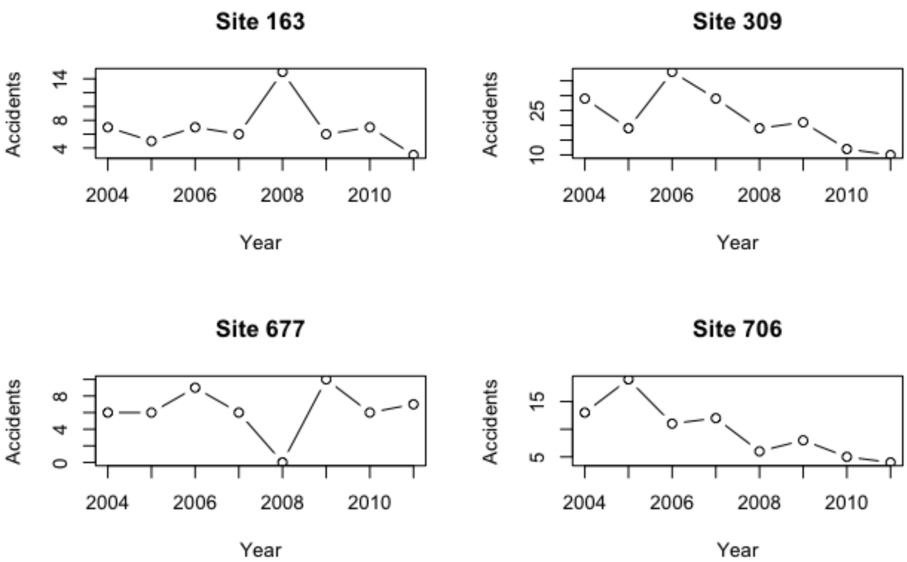

Figure 3: Time series plots of accidents at selected sites

While the yearly accident count was in between zero and eight for the site 163 over the eight-year recording period (2004–2011 inclusive), more than fifteen accidents were observed for the same site in 2008. Similarly, yearly accident counts in site 677 ranged from six to eleven but fell to zero in 2008. Both sites 163 and 677 of Figure 3 show that after a random and temporary change in the year 2008 accident counts reverted back to an average level. This indicates that an RTM effect may have occurred after 2008: the number of accidents naturally may have reverted back to an average level in 2009.

Furthermore, sites 309 and 706 suggest evidence of a temporal trend. In both plots of site 309 and 706 in Figure 3, there is an indication of a trend that the yearly accident counts decrease. Figure 3 proves that trend can vary at a local level since downward-sloping is observed only for site 309 and 706 (and not the other two sites displayed). Therefore, it is crucial that our model should include both the effects of the local and global trend for road safety hotspot prediction.

3 The Model by Fawcett et al.

The model that has been studied and implemented is basically an extension to the classical empirical Bayesian method for analysis, also popularly known as Before-and-After studies. The main idea is to include data from multiple time periods for training and analysis. Common Bayesian approaches only look at a single time period, the so-called before period. As the data has accidents for different times, t is considered as the discrete-time indicator. For the model, $t = 0$ represents current time, $t < 0$ represents the before period and $t > 0$ is reserved for the after period. By using the time indicator, accidents are divided into the respective time periods. Like in the case of most accident prediction models (global APMs), the assumption is that current and future accidents counts are Poisson distributed with variance λ_j. Fawcett et al. model historical counts with variance $\lambda_j(t)$ *c(t). The main reason behind choosing λ_j is to adjust for RTM and trend. The reason behind including *c(t) is to give more weight to recent accidents.

For the before time period $(t < 0)$, we are assuming a Negative Binomial distribution (NegBin). For the after time period, accident counts $(t \geq 0)$ are assumed to be Poisson distributed. Both distributions are similar but when it comes to their distributional properties, the Negative Binomial distribution allows for different mean and variance, whereas for the Poisson distribution mean and variance are assumed to be the same. This means the Negative Binomial distribution allows us to have extra robustness. As a side note, the distributions make sense because we have discrete data and the events occurring are independent of the time since the last event. The modeling process is illustrated in Figure 4, and the overall model is discussed in the next section. Before continuing, it shall be noted that the model is specifically useful since we want to know the individual parameters for each site instead of assuming just the parameters across sites, grasp the relation between the model parameters and have prior knowledge.

3.1 Model Components

As described in the previous section, the model has two components (or submodels). These are Negative Binomial-distributed for the before time period and Poisson-distributed for the after time period. The respective model equations are given below:

$$y_j(t)|\lambda_j(t)) = \begin{cases} Poisson(\lambda_j(t)), & t \geq 0 \\ NegBin\left(r = \frac{\lambda_j(t)}{c(t)-1} \ , \ p = \frac{1}{c(t)} \right), & t < 0 \end{cases} \tag{3}$$

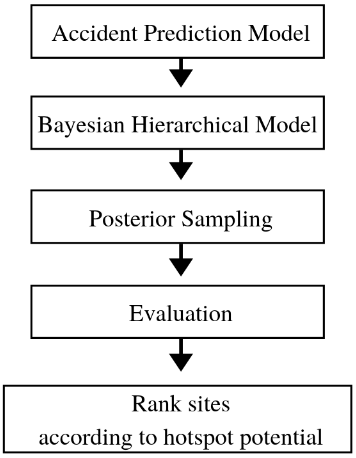

Figure 4: Modeling steps according to Fawcett et al.

In equation (3), the mean accident rate ($\lambda_j(t)$) is dependent on mean accidents represented by $\mu_j(t)$ (also called global APM), a_j which accounts for random effects (RTM) and b_j which tackles the confounding effects of local trend. The mean accidents are built as a log-linear model with coefficients and site details as covariates—it is the standard way to find the MLEs. The variable $c(t)$ allows us to give more weight to recent accidents than to those accidents further in time. Equations for $c(t)$ and $\mu_j(t)$ can be found below:

$$c(t) = exp\{-t\tau\}, t < 0, \tau > 0 \tag{4}$$

$$\mu_j = exp\left\{\beta_0 + \beta_t t + \textstyle\sum_{p=1}^{n_p} \beta_p x_{p,j}\right\} \tag{5}$$

As mentioned, to account for local trend at given sites and also to account for random effects at each site, the two variables b_j and a_j are used as parameters of the distribution. Afterwards, when the model is trained, each site will have coefficients for random effects and local trend. Overall, the mean accident rate parameter looks as follows:

$$\lambda_j(t) = a_j \mu_j(t) exp(b_j t), a_j > 0; -\infty < b_j < \infty; t < 0 \tag{6}$$

3.2 Prior Distributions

The model has in total four prior beliefs and three different distributions. To allow time dependent inflation for the Negative Binomial distribution's variance, τ has been used with a Gamma distribution as a prior. For each site, random effects are accounted for with the a_j parameter which is used to represent this in combination with a Gamma distribution as an another prior. a_j helps us to capture causes we have not captured in our observations. As discussed earlier, the local trend has been accounted for the model and represented with b_j, which is the product of b_n and b_z. In b_j, b_z acts as the zero-inflation component that penalizes the global trend detected by the global APM. Each of these prior distribution equations is given below:

τ - Gamma(2,20); a_j - Gamma(γ, γ); b_N - N(0,0.1); b_Z - Bernoulli(0.5)

3.3 Model Parameters

In total, the model has 5 different parameters (θ_j):

$$\theta_j^{(i)} = \{a_j, b_j, \tau_j, \lambda_j(t)\}^{(i)}, j = 1, ..., 734; t = -7, ..., 0 \tag{7}$$

Inference on these parameters has been made using MCMC sampling technique to simulate approximate draws from the marginal posteriors. Specifically, Metropolis-within-Gibbs MCMC has been used to update the model parameters. Fawcett et al. have trained the model with 10,000 iterations to get the optimal parameters.

Fawcett et al. are doing all this with the aim to identify future road safety hotspots and to rank them according to their potential to cause accidents. Practitioners will want to know which sites they should treat – and which would be inefficient to treat. With the data from Halle, Fawcett et al. learn the posterior predictive distributions for 2012 that allows afterwards to formulate predictions and to quantify uncertainty.

4 Our Approach to Model Replication

This section describes step-by-step the endeavour to replicate the model by Fawcett et al. and its application to the open-source data from the city of Halle. In the project, we followed the same process steps as outlined in the paper. Due to the fact that for some steps only limited descriptions were available, we took the freedom to handle these parts as it seemed most reasonable to us.

The major goal was to train a model to learn the so-called mean accident rate $y_j(t)$, as discussed in the previous section. The mean accident rate is dependent on the outcome of the global APM and the parameters that account for RTM and trend to increase the predictive power. As outlined in the model overview in the previous section, the modeling starts with estimating the global APM—and then continuing with the Bayesian modeling. In the following subsections, we discuss the three central steps involved in the process.

4.1 Global Accident Prediction Model (APM)

To start modeling the global APM to get the μ_j values for all sites and every single year, we first had to rearrange the data towards a horizontal representation. The original data has site-wise accident counts for each year in different columns, which made it impossible to get one time coefficient to estimate the global APM without rearrangement. In the horizontal way, the respective years t_1, t_0, ..., t_{-7} are then indicated in one column.

With this transformed data, we can fit the Negative Binomial Model as discussed in Section 3.1. By doing so, we get coefficients for all covariates as well as time and intercept. Once these parameters are obtained for all the variables, we can calculate the respective mean accidents μ_j for each year and site. As we discussed extensively already, the mean accidents, however, do not account for the confounding effects of RTM and trend. In fact, the global APM is, as Fawcett et al. state, a simple and commonly used model to predict accidents. Our trained Global APM model can be seen in Figure 5:

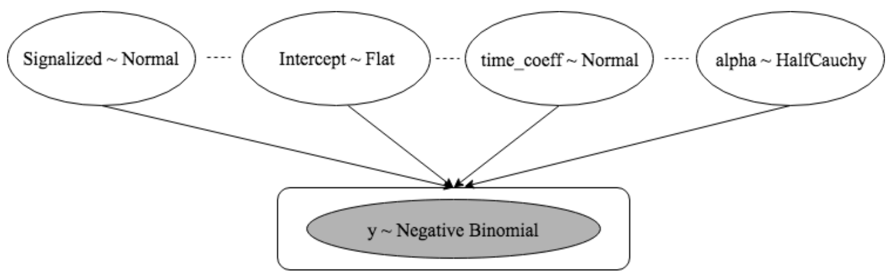

Figure 5: Global APM

4.2 Before Time and After Time Model Training

After having obtained the mean accidents μ_j for each site and point in time with the global APM, the mean accident rate $y_t(t)$ is learned by taking the mean accidents as a prior. Also, we include the priors that account for RTM and global as well as local trend, as specified in Section 3.2. All other priors/hyperpriors are initialized as has been specified by Fawcett et al.

The Poisson Model is used for the after time period and the Negative Binomial Model for before time period. The mean accident rate λ_j represents the main parameter for this process. MCMC sampling technique has then been used to obtain the model parameters and our current model has been trained with 1,000 iterations. Figure 5 gives an overview of the model in plate notation:

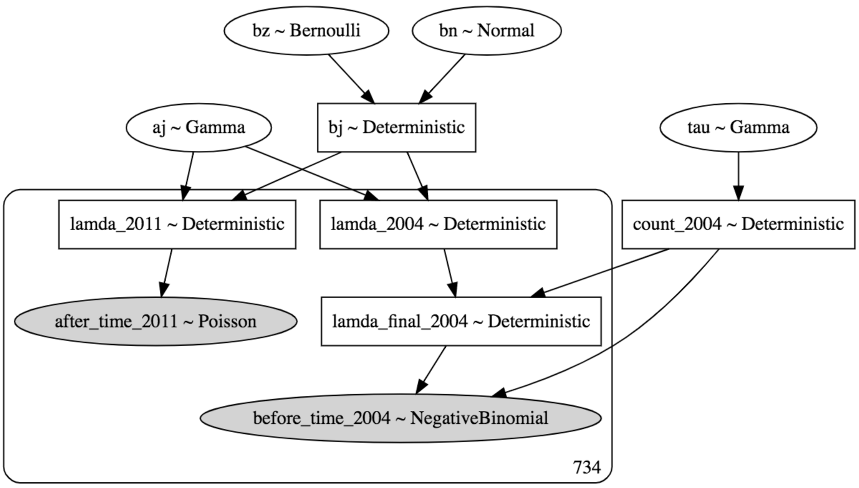

Figure 6: Bayesian Hierarchical Model for hotspot prediction

4.3 Predicting Accidents

For accident predictions, we need to use the parameters that counter RTM, local trend and global trend along with the mean accident rate. To do so, we wrote a final function which mainly takes 3 functional variables: site, year and time. The prediction function first finds the random effect and trend parameters for the given site and then fetches mean accidents (global APM) for the same site. Final predictions are made using these derived values and combining them using equation (6) as described in Section 3.1.

4.4 Language and Libraries Used

All the computations, model training, posterior sampling and result evaluation have been done using the Python language with the help of the pymc3 library. We have experienced the library as a very useful tool to build the even an advanced Bayesian Hierarchical Model in a clearly laid out and fast way. Our code is available on GitHub: https://github.com/topmodels/accident_prediction_bayesian.

5 Comparison and Evaluation

The main focus of the research discussed in this report is to predict accident counts for each site and to rank them accordingly. For evaluation of the model's performance, the 2012 year's accidents are used. As discussed in the previous section, the replicated model has been trained with the observed values from the years 2004-2011. In the aftermath, posterior sampling has been generated for the year 2012 to compare the results with the actual accidents that were measured. Mean Squared Error (MSE) has been used to evaluate the model's accuracy for the year 2012. Currently, the MSE for the replicated model is around 20. As expected, as data (years) are increased in the replicated model, the better the model's predictions turn out.

5.1 Posterior Comparison with Fawcett et al.

Posterior samples have been drawn for four sites that are being discussed throughout the report: 163, 309, 677 and 706. In total, 100 samples have been drawn from the trained model to generate the posterior distributions for the above mentioned sites. The posterior parameters can be seen in the below displayed table. The table shows that a_j and b_j have been learned for each site. The below table shows the observed data, the parameter values we have generated after training and also predictions for the year 2012.

	a_j	b_j	2008			2011			2012
Site			Observed	$\mu_j(t=-4)$ (APM)	$\lambda(t=-4)$	Observed	$\mu_j(t=0)$	$\lambda(t=0)$	Prediction (t = 1)
162	0.83	-0.050	15	3.57	3.79	3.69	3	3.07	3
309	1.60	-0.004	19	4.76	8.24	4.91	10	7.87	8
677	1.92	0.020	0	2.50	4.84	2.58	7	4.97	5
706	0.39	0.002	0	4.54	1.89	4.69	4	1.84	2

Figure 7: Posterior summary of the replicated model

We can see from the Figure 7 that our model has learned the trend element correctly and is fitting well. When we compare the prediction fit of the replicated model with the one of Fawcett et al. it becomes clear that our model has predicted values which are not only close to the observed data, but also close to the results of Fawcett et al. Figure 8 and Figure 9 illustrate the predictions of Fawcett et al. and of the replicated model which we implemented. In these graphs, the blue line indicates the model fitting of the mean accident rates and the red line indicates the model fitting of the Global APM.

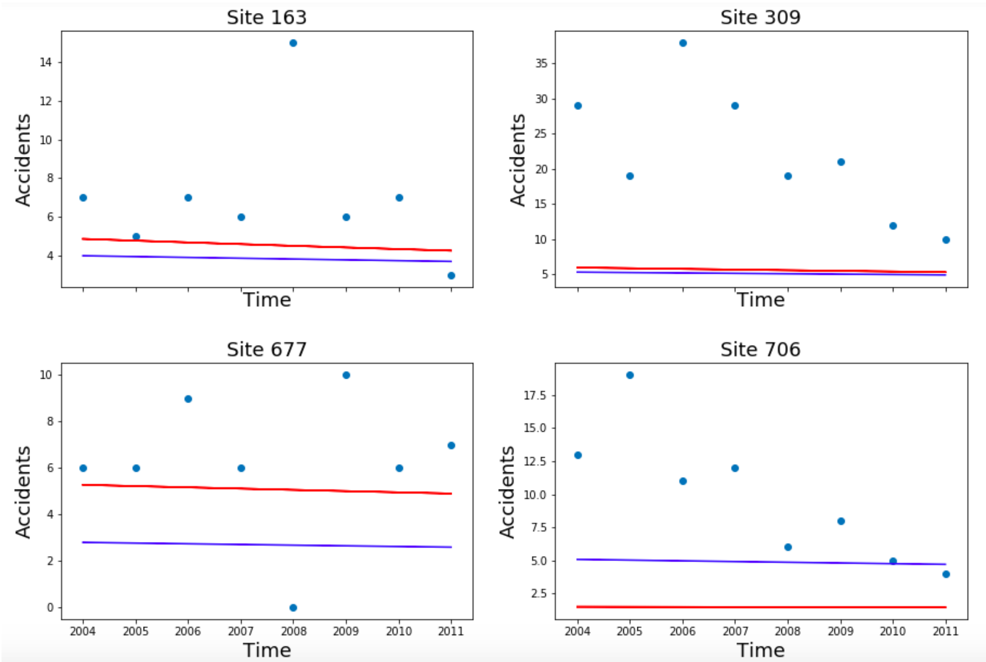

Figure 8: Predictions of the replicated model

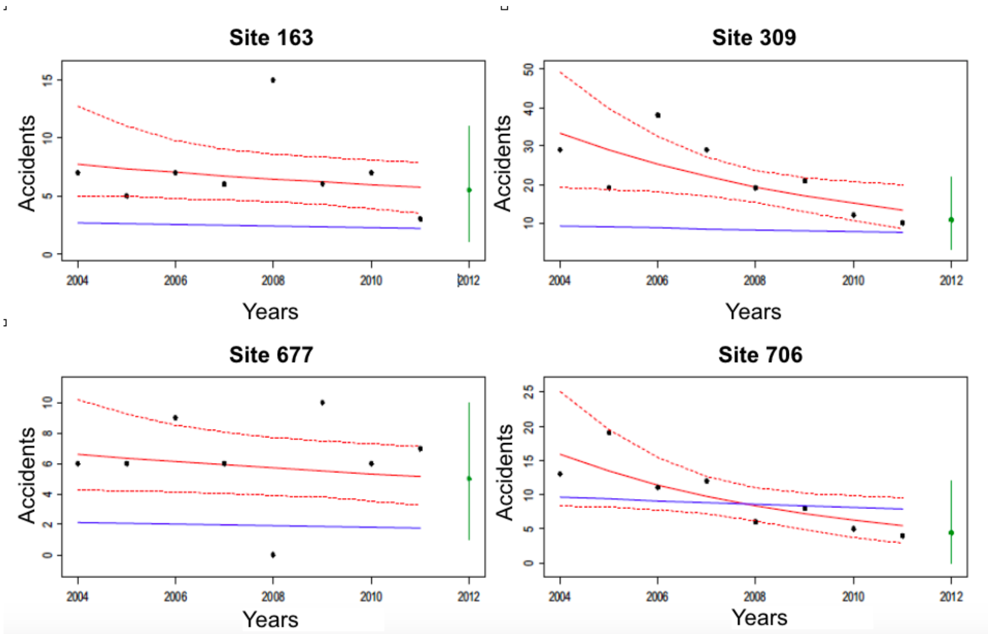

Figure 9: Predictions of Fawcett et al.

5.2 Interpreting the Results

One of the main aims of the model is to rank the sites according to the accident rates predicted from the model. The exceedance probability graph (Figure 10) helps us to visualize the accidents and their respective probabilities. This graph helps in identifying thresholds and to discuss the risk potential of different sites. As the graph indicates, the probabilities for accidents varies between sites: We can see that site 677 (blue) has a higher probability than the others to have more accidents. Accordingly, practitioners would presumably choose to treat this site first. Just by looking at the graph, it might make sense to place a threshold for instance at 10 accidents. In this case, sites 163 and 706 do not exceed this threshold and would correspondingly not be treated—since they have a much lower risk to become a road safety hotspot. The probability that these two sites will have more than 10 accidents is zero. Only the outer two sites would now be treated, because they are likely to exceed the threshold. Although there a slight differences to the exceedance probability graph of Fawcett et al., the graph we produced is still very similar.

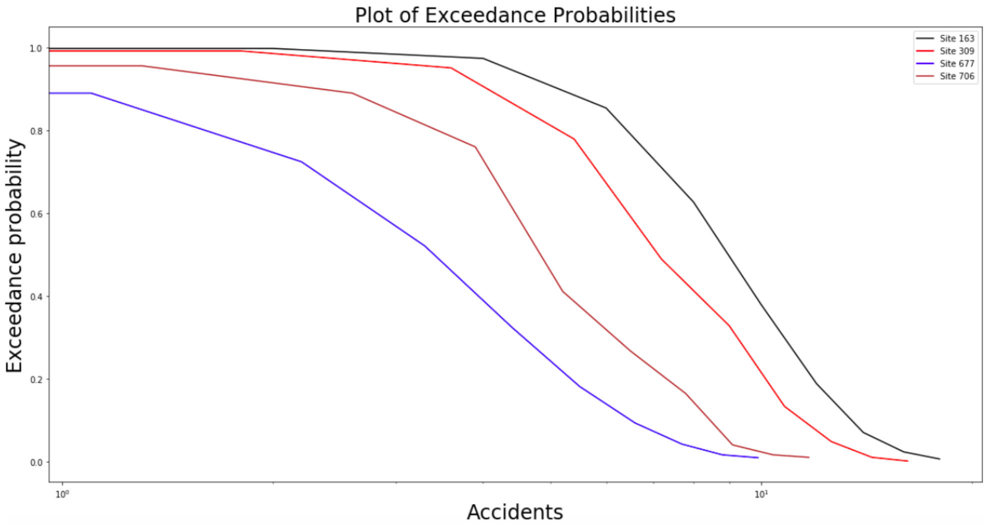

Figure 10: Exceedance probability graph

6 Outlook

After introducing the Bayesian Hierarchical Model, outlining its implementation, and discussing the results, this section aims at identifying possible improvements. At first, we describe how the replicated model's predictive performance could be increased and how harmonization of the results with those of Fawcett et al. could be achieved. Afterwards, we provide ideas on how the original model by Fawcett et al. could be refined and identify further promising approaches to take into account when predicting road safety hotspots.

6.1 Improvements for Replicated Model

The replicated model has been trained with 1,000 iterations on a Macbook machine with 8GB RAM. This was done mainly due to time and space constraints. Fawcett et al., however, trained their model with 10,000 iterations. Correspondingly, increasing the predictive power of the replicated model would require amongst others to increase the number of iterations to 10,000. One important benefit of training the model with more iterations is that the model will learn the parameters better. Since the posterior sampling depends on these parameters, the comparison with the implementation of Fawcett et al. would be much more promising and the replicated model's accuracy is expected to increase.

It could also be beneficial to train the replicated model with another library or language. Fawcett et al. trained their modes using the RJAGS library in the R language. The replicated model was trained with Python using the pymc3 library. For the sake of better comparison and a potential increase in model performance, it could be considered to implement the replicated model in the same environment, namely using the same language and the same library as Fawcett et al.

6.2 Suggestions to Fawcett et al.

Despite the fact that some previous research (e.g. by Cheng & Washington 2008) indicates that a Bayesian approach to road safety hotspot prediction seems most promising, other approaches to road safety hotspot prediction should still be considered. In fact, recent research (e.g. by Yu & Li 2018) around aircraft accident predictions shows that discriminative machine learning methods—in this case Support Vector Machines—predict accidents well.

Motivated by recent advancements in discriminative machine learning, we briefly implemented a relatively simple Neural Network Model on the Halle dataset. The network has been trained with the 2004 to 2011 data to predict 2012 year's accidents. The training does not have any intermediate steps of calculating mean accidents, random effects, etc. It is a straight forward regression model. We have trained the model in Python using the scikit-learn library. We used 3 layers and the MLPRegressor built-in model of scikit-learn with default parameters. In specific, for the activation function of the hidden layer, the identity function has been used. In total, 100 nodes have been deployed in the hidden layer. The weights were optimized with Adam's solver and we conducted 1,000 iterations with the learning rate being 0.001. Although the replicated Bayesian Hierarchical Model has outperformed the Neural Network Model, the accuracy was only slightly better: the MSE for the Bayesian Hierarchical Model is slightly less in comparison with the Neural Network.

Despite discriminative machine learning being another option for predicting road safety hotspots, classical advanced time series modeling implementing ARIMA and ARIMAX techniques (e.g. by Ihueze & Onwurah 2018) still lead to predictions with high accuracy, as in the case of road safety hotspot prediction in Nigeria.

It seems that despite the fact that modeling road safety hotspots is beneficial with a Bayesian approach, further research comparing different approaches on different real-world as well as simulated datasets could be beneficial to advancing research around road safety hotspot prediction. This could also provide road safety practitioners with advice on selecting the best model in a certain context.

In addition, the rise of open-source projects might indicate a useful opportunity for Fawcett et al. On the one hand, an open-source approach to their model might lead to further improvements of the model by making use of the highly-skilled supportive communities, the so-called crowd. On the other hand, an open-source approach could increase the relevance of the model by leading to further practical applications. As discussed by Fawcett et al., the overall goal is to make our roads safer. The interactive website featuring the paper (available at `https://mas-shiny.ncl.ac.uk/hotspot-demo/`) is already a valuable step into this direction.

7 Conclusion

This report outlined an in-depth project on the paper by Fawcett et al. to predict road safety hotspots. Since the core goal of the project was to replicate the model, it can be stated that this has been successfully achieved. In this section, an overall conclusion shall be drawn:

As has been discussed throughout the report, the Bayesian Hierarchical Model developed by Fawcett et al. is a relatively new and useful model to inform road safety practitioners and to thereby increase road safety. It successfully meets current challenges in road safety hotspot prediction, namely the confounding effects of RTM and trend. The model predicts which sites in a certain area will—based on historical data—pass a pre-defined threshold. Based on this site-specific information, practitioners can then implement measures at the most relevant sites and are assured that, with a high probability, the predictions are not flawed by random peaks in the past or global as well as site-specific trend. With this model, investments into measures at sites are well justified.

Fawcett et al. used a Bayesian modeling approach since prior knowledge is available. Instead of having the same parameters across all sites, the model allows for site-specific information. The fact that the model puts more weight on recent accidents increases predictive power additionally. Also, the comparison with the neural net that we implemented provides further support that the Bayesian approach to road safety hotspot prediction enables better predictions. Nevertheless, as discussed in the outlook, other approaches to road safety hotspot prediction are still worth considering to advance research in this field.

As outlined, replicating the model by Fawcett et al. was done using the same process steps as suggested in their paper: We started off with calculating the accident prediction model, then conducted the Bayesian modeling, and closed with posterior sampling and model evaluation. As suggested throughout the report, the way Fawcett et al. have designed there model has proven useful. The results we calculated are close to the results of Fawcett et al., taking into account that at certain points in time, we kept the model slightly leaner. Overall, our predictions, as can be seen especially in terms of the exceedance probabilities, match with the results of Fawcett et al.

In academia today, only a very limited amount of research aims at replicating models and respective results. This project aimed at doing exactly that. The findings of this project clearly show that the approach and the detailed modeling of Fawcett et al. not only seem reasonable, but are now also "cross-validated."

References

Cheng, W., Washington, S. (2008). *New criteria for evaluating methods of identifying hot spots.* Transportation Research Record, 2083(1), 76-85.

Destatis (2019). *Verkehrsunfälle - Statistisches Bundesamt.* Retrieved from https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Verkehrsunfaelle/_inhalt.html

Fawcett, L., Thorpe, N., Matthews, J., Kremer, K. (2017). *A novel Bayesian hierarchical model for road safety hotspot prediction.* Accident Analysis Prevention, 99, 262-271.

Ihueze, C. C., Onwurah, U. O. (2018). *Road traffic accidents prediction modelling: An analysis of Anambra State, Nigeria.* Accident Analysis Prevention, 112, 21-29.

Ketter, W., Peters, M., Collins, J., Gupta, A. (2016). *A multiagent competitive gaming platform to address societal challenges.* Mis Quarterly, 40(2), 447-460.

Sacchi, E., Sayed, T., El-Basyouny, K. (2015). *Multivariate full bayesian hot spot identification and ranking: New technique.* Transportation research record, 2515(1), 1-9.

Yu, H, Li, X. (2019). *On the chaos analysis and prediction of aircraft accidents based on multi-timescales.* Physica A: Statistical Mechanics and its Applications, 120828.

AUTHORS' CONTRIBUTIONS: All authors contributed extensively to the project. The success of the project is due to equally large contributions by all three authors in terms of time and task complexity. In joint discussions, the authors developed the main conceptual ideas on how to tackle the overall model replication, how to frame the two presentations, and how to structure the report. At the same time, certain areas of focus were defined: EÖ focused especially on the bigger picture, the impact of the research, communication with the original authors and the data. With the help of EÖ and TS, GD concentrated on coding the model, looking for opportunities on how to improve it, and further technical details. TS focused predominantly on the theory behind the model and its analysis, the model comparisons, creating a structure for the presentations and the report. GD and EÖ put special attention on interpreting the analysis, while TS helped on how to convey the content to the audience/reader. All authors helped in verifying the numerical results by an independent implementation of certain steps in the R language. In addition, all authors discussed the results, commented extensively, and developed an outlook.