

BIVARIATE POISSON CREDIBILITY MODEL AND BONUS-MALUS SCALE FOR CLAIM AND NEAR-CLAIM EVENTS

Pierre-Alexandre Simon, Julien Trufin,
Michel Denuit

REPRINT | 2025 / 07



ISBA

Voie du Roman Pays 20 - L1.04.01

B-1348 Louvain-la-Neuve

Email : lidam-library@uclouvain.be

<https://uclouvain.be/en/research-institutes/lidam/isba/publication>



Bivariate Poisson Credibility Model and Bonus–Malus Scale for Claim and Near-Claim Events

Pierre-Alexandre Simon,¹ Julien Trufin,¹  and Michel Denuit²

¹*Department of Mathematics, Université Libre de Bruxelles (ULB), Brussels, Belgium*

²*Institute of Statistics, Biostatistics and Actuarial Science - ISBA, Louvain Institute of Data Analysis and Modeling - LIDAM, UCLouvain, Louvain-la-Neuve, Belgium*

With the emergence of telematics, huge amounts of data have become available to insurers, raising the question of their integration into motor insurance pricing. Near-claim events are dangerous events that could have triggered a claim. Typically, they correspond to harsh deceleration, acceleration, or cornering. Based on the motor third-party liability insurance portfolio of an insurance company operating in Belgium, we highlight that deceleration events are the best candidates for characterizing near-claim events. Then, we propose to integrate these events in the pricing structure in addition to a policyholder's claim history with the help of a bivariate Poisson-Lognormal credibility model. Finally, we design an original bonus–malus scale based on both claim and near-claim events.

1. INTRODUCTION

Premium calculation in motor insurance routinely uses explanatory variables including driver's profile (age, place of residence, etc.) or vehicle characteristics (power, make and model, etc.). In addition to these classical features, insurance companies have started collecting telematics data, offering new opportunities for risk evaluation. These data are typically recorded by telematics boxes installed in each vehicle or policyholder's smartphone. They typically include distance traveled, speed, or acceleration measured in real time.

Because telematics are not yet predominantly deployed, a standard pricing model is still missing. Several proposals have been made in the literature. A first research stream concentrates on the construction of new explanatory variables based on the processing of telematics signals, to supplement traditional risk factors. Gao, Meng, and Wüthrich (2019) constructed driving scores from speed–acceleration heat maps and included them as additional explanatory variables for claim frequency prediction. Gao, Wang, and Wüthrich (2022) applied a combination of feedforward and convolutional neural networks to process these heat maps. So, Boucher, and Valdez (2021) proposed a three-way simulation procedure to generate a telematics portfolio using the synthetic minority oversampling technique in conjunction with feedforward neural networks. The analysis conducted by Gao, Meng, and Wüthrich (2019) and Duval, Boucher, and Pigeon (2022) suggested that only 3 months or 4000 kms of driving experience is enough to characterize the driving style of a given policyholder. This illustrates the richness of telematics data. We refer the reader to Gao, Meng, and Wüthrich (2022) for further information about the preprocessing techniques of driving data.

Another approach consists of modeling telematics signals as supplementary response variables along claim counts. As part of this approach, Guillen et al. (2020) and Sun et al. (2021) defined different types of near-miss events (a term borrowed to the aviation safety) such as excess speed, harsh acceleration, harsh deceleration, or cornering events. They highlighted that significant risk factors (traditional and telematics) depend on the type of near-miss events under consideration. Both studies concluded that considering all near-miss events leads to misleading interpretation because all near-misses are not necessarily associated with a higher claim frequency.

A last stream of research assumes that the prediction of claim frequency based on traditional rating factors has to be adjusted based on the policyholder's history, including past claims and telematics signals. This echoes multiple-event

Address correspondence to Julien Trufin, Department of Mathematics, Université Libre de Bruxelles, Brussels 1050, Belgium. E-mail: julien.trufin@ulb.be

experience rating systems initially proposed to combine several products or to integrate claim severities into premium updates, in a simplified way. See, for example, Dionne and Vanasse (1992), Pinquet (1998), Gomez-Deniz (2016), Park, Kim, and Ahn (2018), Cheung et al. (2021), Denuit and Lu (2021), Tzougas and di Cerchiara (2021), and Verschuren (2022), as well as the references therein. Recognizing that telematics data are recorded while policyholders are driving, similar to claim experience, Denuit, Guillen, and Trufin (2019) and Corradin et al. (2022) proposed using multivariate credibility models (also called mixed models, in statistics) to capture the joint dynamics of telematics data and claim experience. Past telematics signals are not included in the score like ordinary risk factors but used for premium correction in experience rating. This is the approach followed in this article.

Even if credibility models are powerful tools to integrate driving experience in future premiums, motor insurance generally favors their commercial simplification in the form of bonus–malus scales. These scales comprise a number of levels, and policyholders’ movements depend on the number of claims they reported to the insurer. The relativities associated to each level correspond to the relative corrections applied to the individual basis premium. Bonus–malus scales are effectively modeled as Markov chains and relativities are generally computed from the corresponding stationary distribution. In this article, we propose a novel scale, integrating both claim and near-claim events in a simple and transparent way, and we adapt the approaches proposed by Pitrebois, Denuit, and Walhin (2003) and Tan et al. (2015) to derive the relativities.

The analysis conducted in this article is based on the motor third-party liability (MTPL) insurance portfolio of an insurance company operating in Belgium, observed over the period 2017 to 2019. The remainder of the text is organized as follows. Sections 2 and 3 describe the MTPL dataset. We first consider several candidates to define appropriate near-claim events. Then, we correct frequencies of claims and near-claim events for standard features used in motor insurance pricing using gradient boosting machine (GBM; Section 3). In Section 4, we design a bivariate Poisson-Lognormal credibility model combining claim and near-claim events. The bivariate model is compared to the univariate credibility model only based on reported claims. Section 5 proposes a bonus–malus scale integrating claim and near-claim events. Section 6 briefly discusses the results obtained in this article as well as possible extensions.

2. NEAR-CLAIM EVENTS

2.1. Telematic Signals Analysis

Following Guillen, Nielsen, and Pérez-Marín (2021), we propose to overcome the relatively low claim frequency in motor insurance by identifying relevant near-claim events. This term is borrowed from aviation safety and denotes an event in which an aircraft comes dangerously close to a collision but avoids it by a narrow margin. Similar to studying the signals recorded by the black box in an airplane after a crash, we investigate the telematic signals the day of claim occurrence to highlight possible dangerous events, that could have resulted in a claim. Because we consider MTPL insurance, we focus our study only on claims at fault. The telematic database considered in this article gathered logs of vehicle events recorded during policyholders’ trips over the period 2017 to 2019. These events are typically harsh deceleration (DE), acceleration (AE), quick lateral movement (QLM), cornering (CE), and speed excess (SE) events. They are labeled as such by the manufacturer for all events but SE events when the vehicle’s instantaneous acceleration exceeds an absolute threshold of $0.2g$ and when the vehicle’s instantaneous speed exceeds the road speed limit for SE events. For each event, associated telematic signals as GPS quality, distance traveled, instantaneous speed, or acceleration are recorded. Among all observations, the majority of records relate to CE events (130,665,548), followed by DE events (80,653,333), AE events (48,291,158), QLM events (40,065,254), and finally SE events (9,371,138). For each type of event, we propose analyzing the associated telematic signals to highlight a possible association between these events and the occurrence of a claim.

By inspecting marginal distributions of telematics signals separately for policyholders with and without a claim at fault in Figures 1 to 3, we notice that a clear divergence is only observable for deceleration events. This suggests that harsh deceleration events seem to be correlated with claims at fault occurrence. These events basically translate the late reaction of the policyholder in the event of danger. It is also worth mentioning that more than half of the claim days exhibit a harsh deceleration. Speed excess and acceleration events are related more to a specific driving attitude rather than directly to claim at fault occurrence. Moreover, this information is partially included in deceleration signals because harsh decelerations are more likely to happen if the policyholder drives at high speed and in an aggressive way. Therefore, we propose considering exclusively deceleration events in our definition of near-claim events. Then, we investigate further raw telematics signals of policyholders the day of claim occurrence. Figures 4 and 5 display the observed telematics data available for two policyholders the day of claim occurrence. In both figures, horizontal dotted lines represent the speed limit in force in Belgium with an additional 10% margin for the speed signal (middle), the acceleration/deceleration thresholds used by Guillen, Nielsen, and Pérez-Marín (2021) for acceleration signal (right), and an arbitrary 1-minute threshold for event duration signal (left). The vertical dotted red line

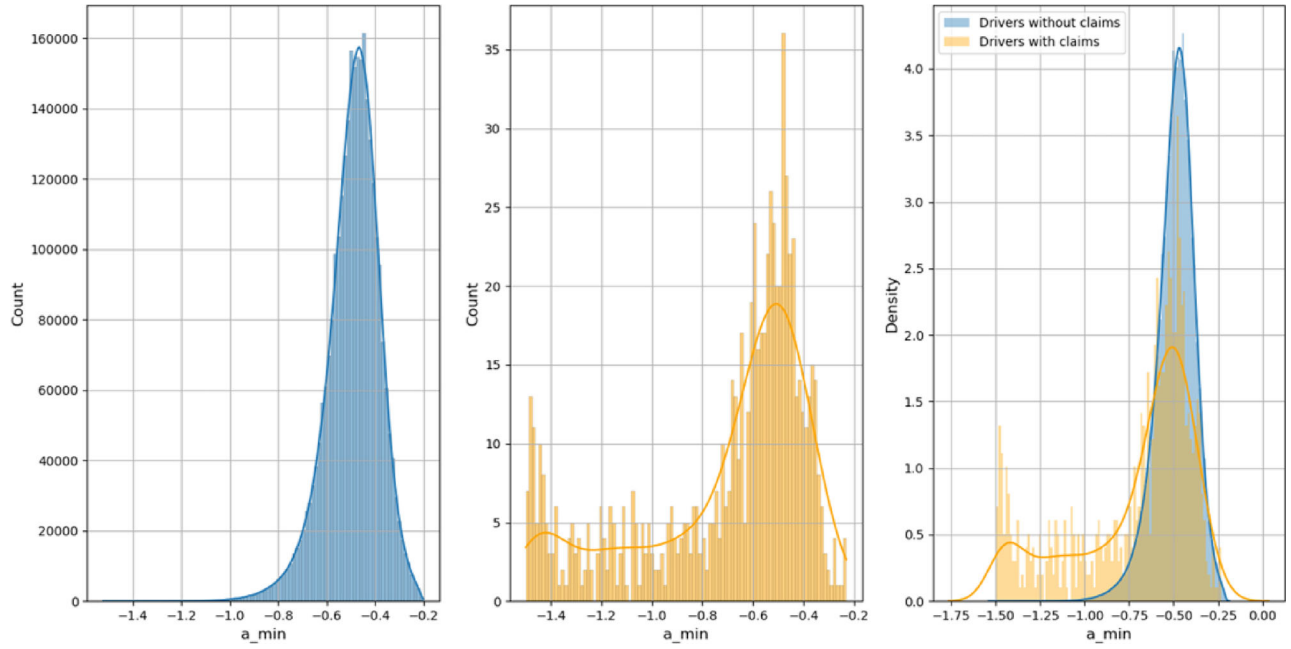


FIGURE 1. Deceleration Distribution for Drivers without (Blue) and with Claims (Orange).

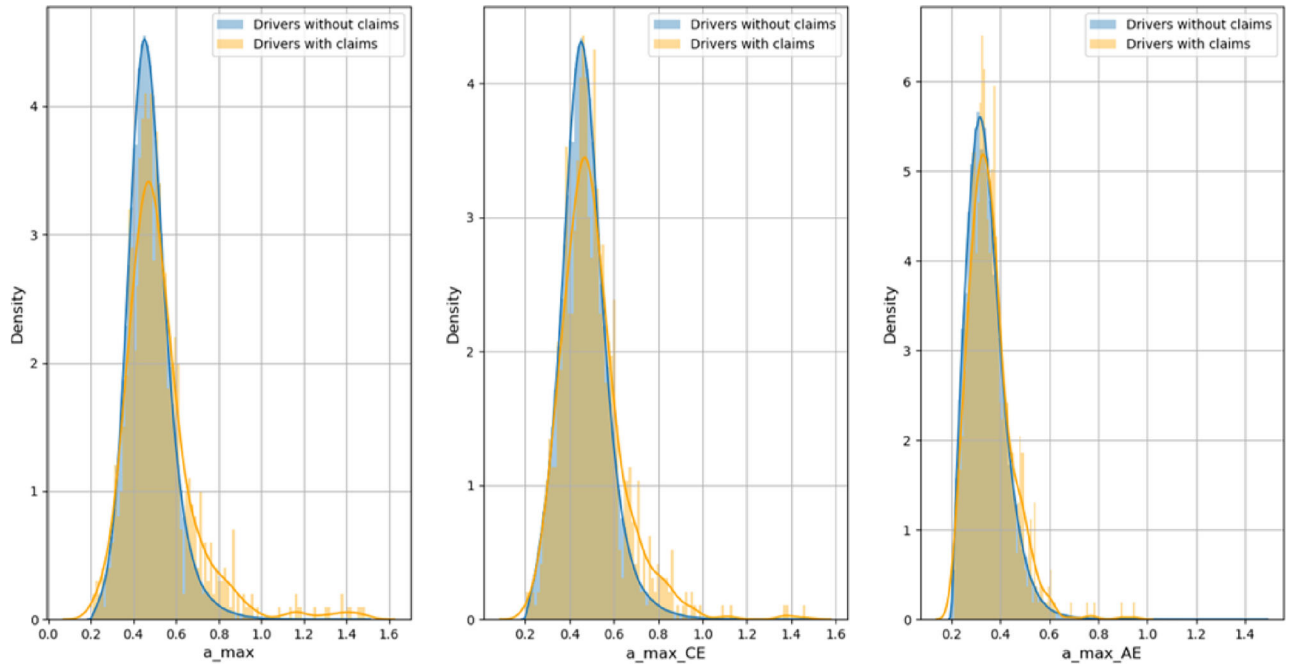


FIGURE 2. Maximum Acceleration Distribution for Drivers without (Blue) and with Claims (Orange) for (a) All Events, (b) CE Events, and (c) AE Events.

represents the declared claim occurrence time by the policyholder. The vertical dotted black line represents the time of observed harsh deceleration and the next event occurrence following this harsh deceleration. Finally, we represent by a vertical green line the next vehicle startup event when the policyholder effectively starts a new trip following this event.

In Figure 4, we observe that the next event occurrence a couple minutes after the harsh deceleration is a vehicle stopping event and that no events are available afterwards until the declared claim occurrence time, suggesting that the harsh deceleration event is directly associated with the claim occurrence. In Figure 5, we observe that after the harsh deceleration event, the vehicle stops and starts again about 10 minutes later. This delay could be associated with the duration of vehicle stops for the

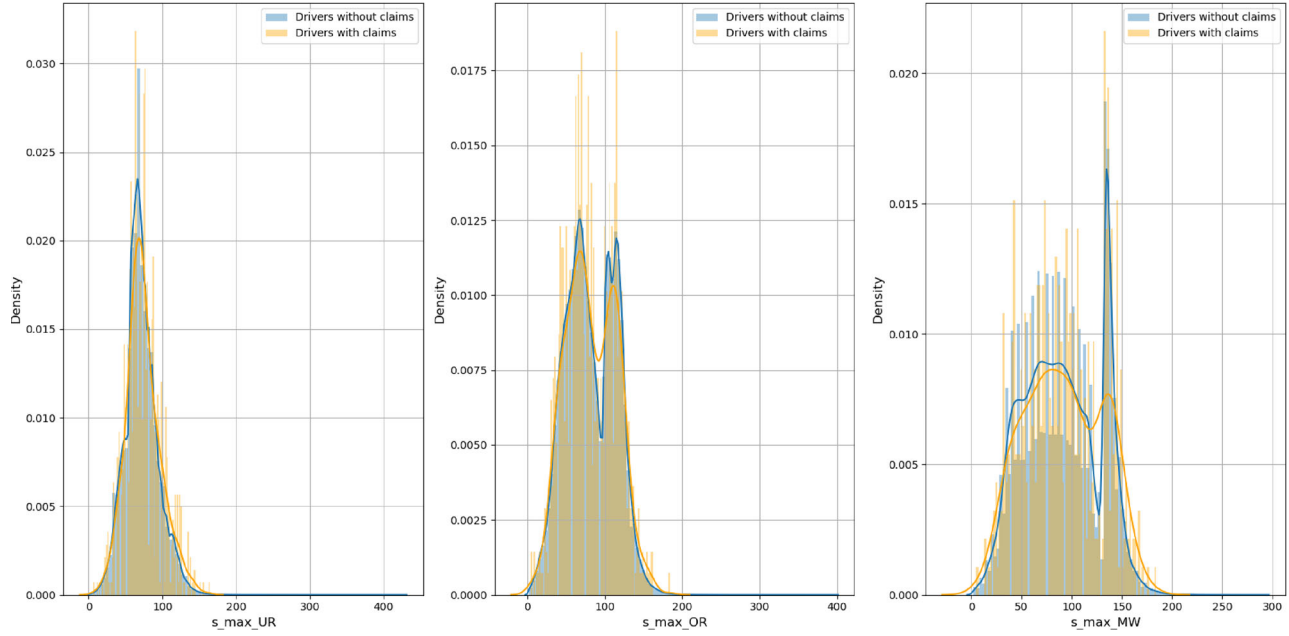


FIGURE 3. Maximum Speed Distribution for Drivers without (Blue) and with Claims (Orange) on (a) Urban Roads, (b) County Roads, and (c) Motorways.

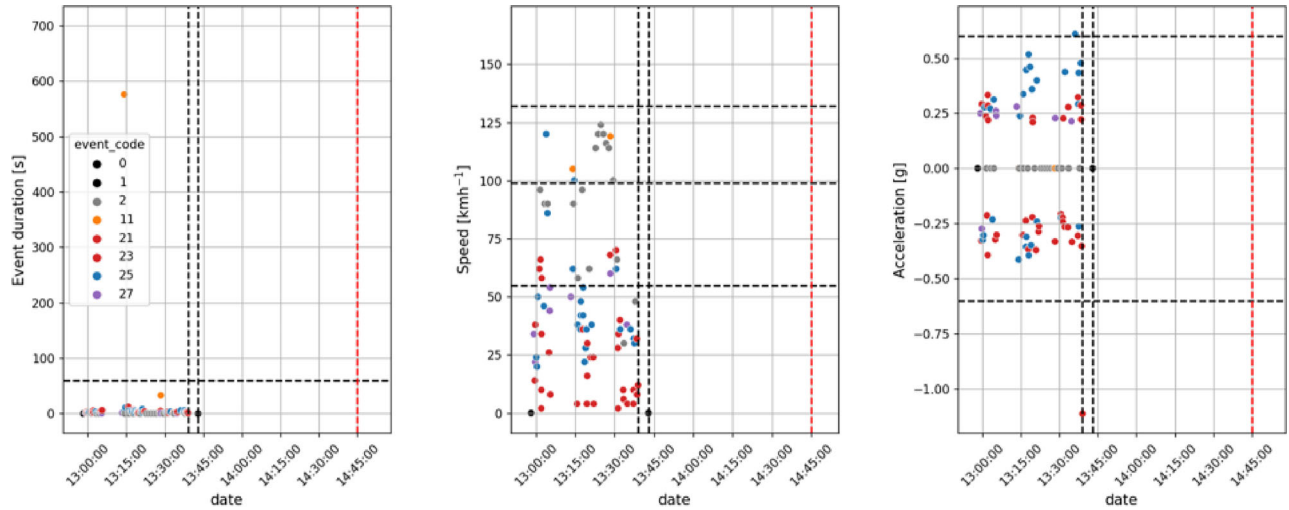


FIGURE 4. Telematics Signals of a Policyholder Exhibiting Harsh Deceleration the Day of Claim Occurrence.

completion of an accident report. However, the declared claim occurrence time is much later and located in the neighborhood of other events, preventing any conclusion and implying an indirect association. One can note that the policyholder in [Figure 5](#) also has exhibiting excess speed events and harsh acceleration events, whereas the second one does not exhibit such events. As in both figures, the declared claim occurrence time seems approximate, and we cannot state anything more than a correlation between harsh deceleration event and claim occurrence.

2.2. Near-Claim Events Characterization

The difficulty lies in determining the appropriate threshold to qualify such events as near-claim events. Guillen, Nielsen, and Pérez-Marín (2021) suggested that some DE events (with intensities 1 and 3) are associated with higher claim frequency and some (with intensity 2) with lower claim frequency where DE events characterization is carried out by the telematics manufacturer. Guillen et al. (2020) considered the same absolute threshold of 0.6g to characterize DE and AE events as near-

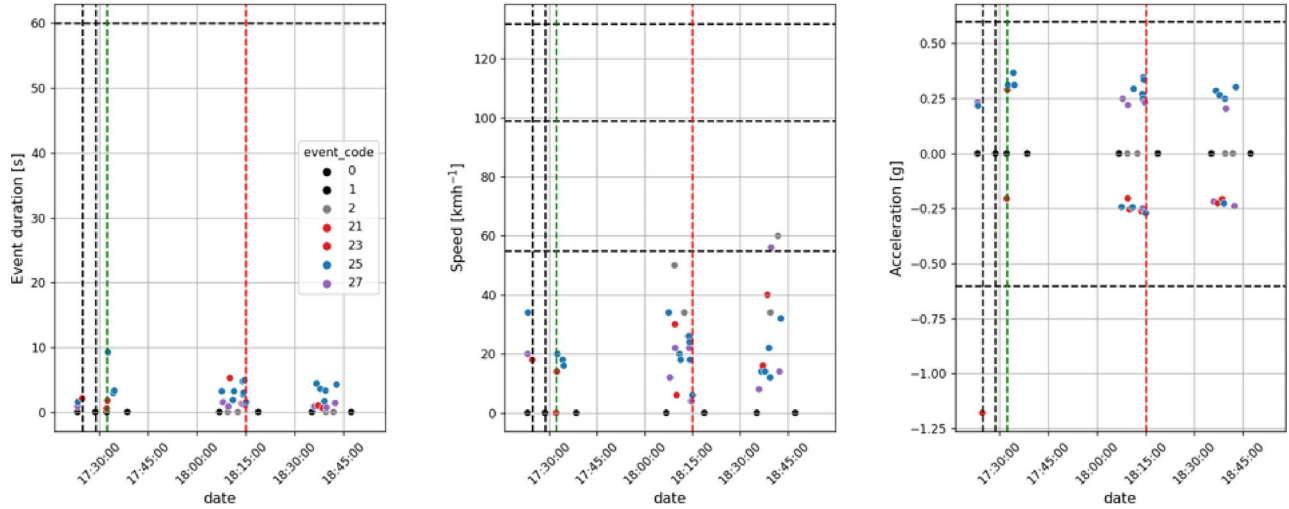


FIGURE 5. Telematics Signals of a Policyholder Exhibiting Harsh Deceleration the Day of Claim Occurrence.

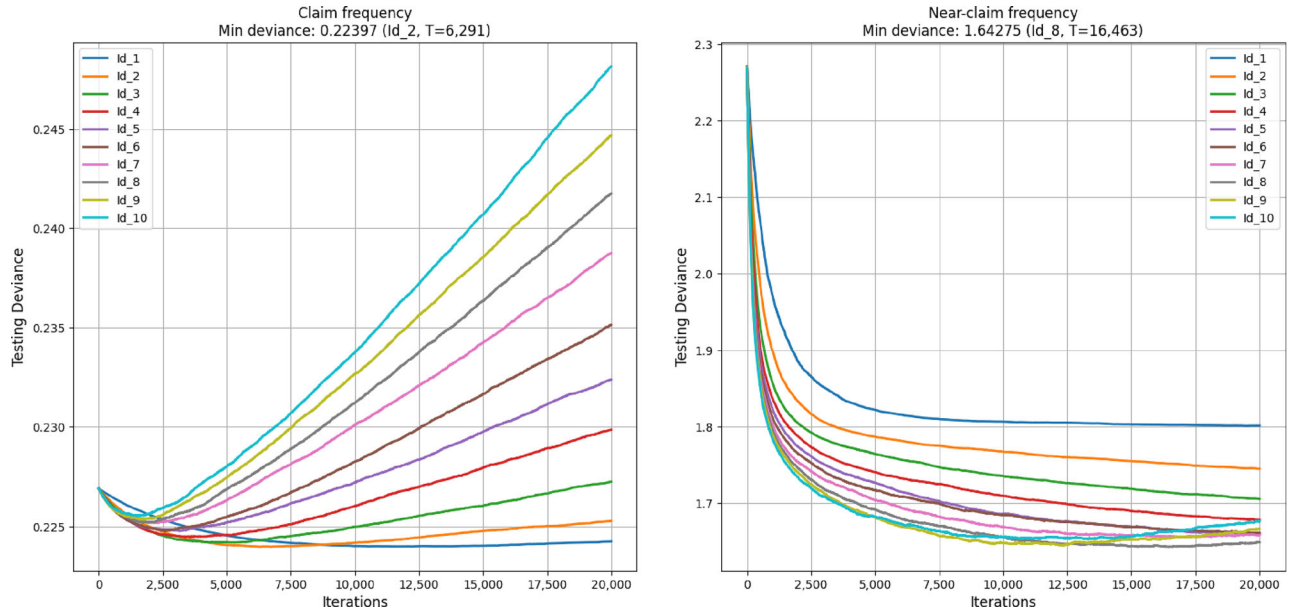


FIGURE 6. Mean Poisson Deviance Observed for Different Interaction Depths for (a) Claim Frequency Models and (b) Near-Claim Frequency Models.

miss events. However, based on our preliminary study, we see that empirical distributions exhibit some divergences, suggesting that a proper threshold for DE events should be considered. From accident analysis and prevention literature, Chevalier et al. (2016) demonstrated the association between contrast sensitivity decline for older drivers and rapid deceleration events. These events are characterized as such for a deceleration rate lower than $-0.75g$, corresponding to a 26.5 km/h speed reduction in 1 second, and used as a surrogate safety measure. Fu and Sayed (2021) and Zheng, Sayed, and Essa (2018) studied the deceleration rate to avoid a crash. They point out that the maximum available deceleration rate above which the crash is unavoidable depends on the vehicle weight, tires, braking system, and road pavement condition. They recommended using the mean of the truncated Normal distribution (with mean 8.45 and standard deviation 1.402) as a threshold (in ms^{-2}) for crash occurrence. De Ceunynck (2017) recommended $-0.8g$ as a conservative value for the maximum deceleration rate that all automobiles can achieve in case of a braking emergency. In our study, because the telematics portfolio contains heterogeneous vehicles, we decided to use this threshold to characterize a particular DE event as a near-claim event, corresponding to a braking emergency situation. The dataset was divided into two parts: a training set containing data for 2017 and 2018 and a testing set containing 2019 half-year data. Table 1 describes the number of near-claim events and claims in the two respective datasets.

3. A PRIORI RISK CLASSIFICATION

We denote by $N_{i,t}$ (respectively $Z_{i,t}$) the number of claims at fault (respectively the number of near-claim events) recorded for policyholder i , $i = 1, 2, \dots, n$, during period t , $t = 1, 2, \dots, T_i$, and by $d_{i,t}$ the corresponding exposure to risk, which is the distance driven in kilometers or time spent behind the wheel in the context of telematics data. Here, $d_{i,t}$ is the distance driven in kilometers. The proper inclusion of $d_{i,t}$ has been studied in the literature. For instance, Boucher, Côté, and Guillen (2017) and Ayuso, Guillen, and Nielsen (2019) documented the relationship between distance driven and expected claim frequency, which may be due to learning effects associated to long distances travelled or the driving environment, which may differ according to annual mileage. In this article, we consider the number of events recorded by kilometers driven. This means that we work with claim (respectively near-claim) frequencies per kilometer obtained by dividing $N_{i,t}$ (respectively $Z_{i,t}$) by $d_{i,t}$ and weighting the observed frequencies by the corresponding distances $d_{i,t}$. In the Poisson regression setting, using response $N_{i,t}/d_{i,t}$ weighted by $d_{i,t}$ is equivalent to analyzing $N_{i,t}$ by specifying $\ln d_{i,t}$ as offset; that is, by multiplying the expected number of claims per kilometer by $d_{i,t}$. The same comments apply to near-claim events. This explains why the distance driven in kilometers is included as exposure to risk in the model adopted in this article.

The distance traveled could also have been included among explanatory variables for both numbers of claims and near-claim events. Some care is nevertheless required because this information is generally not available a priori (that is, before the coverage period starts) but only recorded a posteriori, as is the case for $N_{i,t}$ and $Z_{i,t}$. Past values of $d_{i,t}$ could nevertheless be used to distinguish several drivers' profiles based on the intensity of the use of the vehicle. Here, we do not follow this route and we adopt $d_{i,t}$ as a measure of exposure.

At the beginning of each insurance period, the actuary has at his or her disposal some information about each policyholder summarized into p features $x_{i,t,j}$, $j = 1, \dots, p$. The a priori risk classification is based on the a priori information $\mathbf{x}_{i,t} = (x_{i,t,1}, \dots, x_{i,t,p})$ that is integrated into scores $\eta_{i,t} = \eta(\mathbf{x}_{i,t})$ for the number of claims at fault and $\nu_{i,t} = \nu(\mathbf{x}_{i,t})$ for the number of near-claim events for policyholder i in period t (taken to be a calendar year), so that the expected number of claims at fault and near-claim events are

$$E[N_{i,t} | X_{i,t} = x_{i,t}] = d_{i,t} \exp(\eta_{i,t})$$

and

$$E[Z_{i,t} | X_{i,t} = x_{i,t}] = d_{i,t} \exp(\nu_{i,t}),$$

respectively. Scores $\eta_{i,t}$ and $\nu_{i,t}$ are thus functions of the features and are estimated with the GBM algorithm developed by Friedman (2008). To this end, the dataset was aggregated by policyholder on a yearly basis. The calendar year is thus the unit of time in the rest of the article. We acknowledge that it is a limitation of our dataset because we only have 2 years available in the training set. The explanatory variables used for scoring are described in the Appendix. Notice that exposure has been restrained in the range [100km, 40, 000km] to avoid too high volatility in the predicted expected frequencies.

3.1. Model Calibration

To assess the robustness of each model, we perform five fold stratified cross-validation. Each fold is designed so that the empirical probability mass function of the number of claims at fault (respectively near-claim events) remains similar. Table 2 displays the exposures (number of kilometers driven) on each slice as well as the claim (respectively near-claim) numbers per 1000 km driven.

To tune the GBM on each fold, we considered a bag fraction equal to 0.5. The number of trees is closely tied with the shrinkage parameter retained. By considering a standard shrinkage parameter $\nu = 0.01$, a preliminary study indicated that the minimum of the Poisson deviance on a out-of-sample set containing 20% of the training set observations was reached for a number of trees lower than 20, 000. Thus, we retained this value for the maximum number of trees M . We also observed that the minimum of the mean Poisson deviance was obtained for a lower number of trees for claim frequency model (< 7500) than for near-claim frequency model ($< 15,000$). As a consequence, we considered $\nu = 0.01$ for the near-claim model and $\nu = 0.001$ for the claim model, letting the maximum number of trees be sufficiently large to obtain the minimum of the mean Poisson deviance on the interval $[0; 20,000]$. Because the minimum number of observations per leaf constrains the interaction depth of the trees, we decided to set it to one while letting the interaction depth vary in the range $[1, 10]$. Once the GBM tuning was performed on each fold of the stratified dataset for claim and near-claim counts, the mean Poisson deviance for each interaction depth was computed over the different folds as depicted in Figure 6. This leads to the best model parameters selection for each model, which are summarized in Table 3.

TABLE 1
Response Variables Statistics

Dataset	Training set	Testing set
No. of near-claim events	29,167	6,765
No. of claims	2,336	435
No. of claims at fault	1,111	234
No. of claims not at fault	1,225	201
Exposure (km)	258,585,950	57,335,704

TABLE 2
Stratified Cross-Validation

Slice	S1	S2	S3	S4	S5
Nb of rows	7,707	7,706	7,708	7,708	7,707
Exposure (km)	51,413,880	52,122,710	52,345,880	52,580,810	50,867,180
Mean observed claim frequency (per 1000 km)	0.0051	0.0050	0.0050	0.0050	0.0051
Mean observed near-claim frequency (per 1000 km)	0.1099	0.1085	0.1079	0.1075	0.1096

TABLE 3
Optimal GBM Parameters

Parameters	Claim GBM	Near-claim GBM
Shrinkage	0.001	0.01
Bag fraction	0.5	0.5
Minimum observations per leaf	1	1
Interaction depth	2	8
Number of trees	6291	16,483

3.2. Model Interpretation

3.2.1. Importance of Explanatory Variables

For each best claim (respectively near-claim) frequency model trained on each fold, the respective importance of explanatory variables is represented in [Figure 7](#). For both frequency models, the two most important variables responsible for more than 10% of the Poisson deviance reduction are the vehicle make followed by the main driver's age. For the other variables responsible for more than 5% of the Poisson deviance reduction, we retrieved common explanatory variables for both models such as the geographical coordinates of the place of residence, vehicle value, geomosaic code, and vehicle power and weight. The main differences between both models were age variable and vehicle weight, whose relative importance order was interchanged between both models.

3.2.2. Partial Dependencies

From the importance of the explanatory variables to each response variable identified in the previous section, we represent the effect of the most important variables on the predicted claim frequency (respectively near-claim frequency) in [Figure 8](#). We chose to use the accumulated local effect (ALE) plot of Apley and Zhu (2016) to isolate the interaction effect of the other

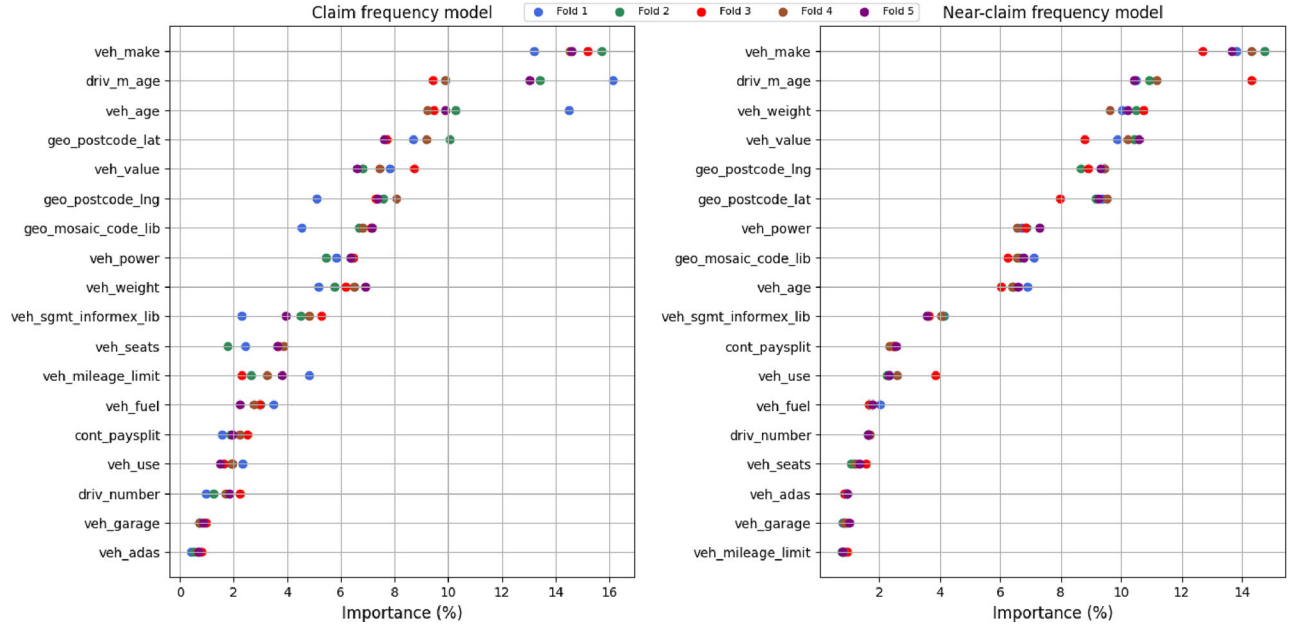


FIGURE 7. Explanatory Variables Importance for (a) Claim Frequency Model and (b) Near-Claim Frequency Model on Each of the Five Folds.

explanatory variables on the dependence between the studied explanatory variable and the response. Notice that the ALE plot is centered at zero, expressing the difference in expected claim (respectively near-claim) frequency with the respective mean expected frequency on the entire portfolio. For the vehicle make, partial dependence with expected claim frequency exhibited high volatility, mainly produced by outliers. This is due to the large number of vehicle makes related to the number of claims at fault present in the aggregated database. This prevented us from drawing any robust conclusion about an eventual similarity in the dependence of some vehicle makes on claim frequency and those with expected near-claim frequency. However, we observed that dependence of the numerical explanatory variables on each response variable showed the same behavior. Guillen et al. (2020) emphasized that driver age was associated with lower near-miss frequency. Interestingly, we observed that vehicle age and weight were associated in both cases with higher (respectively lower) expected frequency. We propose in the following section to incorporate a posteriori individual claim and near-claim history.

4. A POSTERIORI RISK CLASSIFICATION

Random effects Δ_i and Γ_i are added to the scores $\eta_{i,t}$ and $\nu_{i,t}$, respectively, to recognize the residual heterogeneity of the portfolio. Precisely, we assume that

$$E[N_{i,t}|X_{i,t} = x_{i,t}, \Delta_i] = \frac{d_{i,t} \exp(\eta_{i,t} + \Delta_i)}{E[\exp(\Delta_i)]}$$

and

$$E[Z_{i,t}|X_{i,t} = x_{i,t}, \Gamma_i] = \frac{d_{i,t} \exp(\nu_{i,t} + \Gamma_i)}{E[\exp(\Gamma_i)]}.$$

Let us write

$$\lambda_{i,t} = \frac{d_{i,t} \exp(\eta_{i,t})}{E[\exp(\Delta_i)]}$$

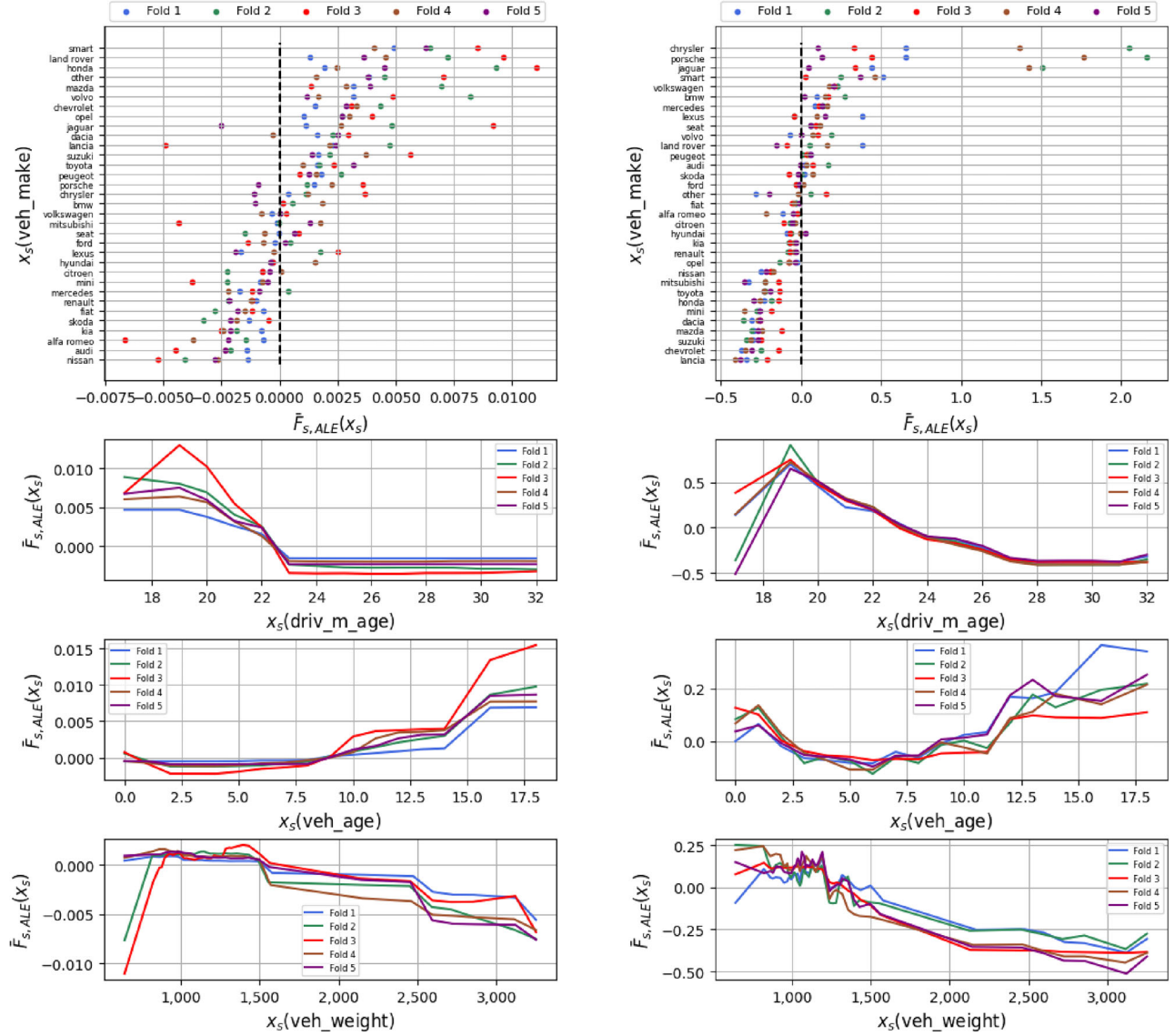


FIGURE 8. Most Important Explanatory Variables Partial Dependence (ALE Plot) for (a) Claim Frequency Model and (b) Near-Claim Frequency Model on Each of the Five Folds.

and

$$\mu_{i,t} = \frac{d_{i,t} \exp(\nu_{i,t})}{\mathbb{E}[\exp(\Gamma_i)]}.$$

Given $\Delta_i = \delta$, the random variables $N_{i,t}$, $t = 1, 2, \dots$, are assumed to be independent and Poisson distributed with means $\lambda_{i,t} \exp(\delta)$. At the portfolio level, the sequences $(\Delta_i, N_{i,1}, N_{i,2}, \dots)$, $i = 1, 2, \dots, n$, are assumed to be mutually independent. Similarly, given $\Gamma_i = \gamma$, the random variables $Z_{i,t}$, $t = 1, 2, \dots$, are supposed to be independent and Poisson distributed with means $\mu_{i,t} \exp(\gamma)$, and the sequences $(\Gamma_i, Z_{i,1}, Z_{i,2}, \dots)$, $i = 1, 2, \dots, n$, are assumed to be mutually independent. Moreover, given Δ_i , claim counts $N_{i,1}, N_{i,2}, \dots$ are assumed to be independent of $\Gamma_i, Z_{i,1}, Z_{i,2}, \dots$, and given Γ_i , near-claim counts $Z_{i,1}, Z_{i,2}, \dots$ are assumed to be independent of $\Delta_i, N_{i,1}, N_{i,2}, \dots$. The random vectors (Δ_i, Γ_i) , $i = 1, 2, \dots, n$, are supposed to be mutually independent and to obey the bivariate Normal distribution with zero mean, variances σ_Δ^2 and σ_Γ^2 , and correlation $\rho_{\Delta,\Gamma}$, so that $\mathbb{E}[\exp(\Delta_i)] = \exp(\sigma_\Delta^2/2)$ and $\mathbb{E}[\exp(\Gamma_i)] = \exp(\sigma_\Gamma^2/2)$. This is henceforth denoted as $\mathcal{N}(\mathbf{0}, \Sigma)$ with the variance-covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_{\Delta}^2 & \rho_{\Delta,\Gamma}\sigma_{\Delta}\sigma_{\Gamma} \\ \rho_{\Delta,\Gamma}\sigma_{\Delta}\sigma_{\Gamma} & \sigma_{\Gamma}^2 \end{pmatrix}$$

and corresponding probability density function $f_{\Sigma}(\cdot, \cdot)$. So the vectors (Δ_i, Γ_i) , $i = 1, 2, \dots, n$, are mutually independent, and for any i , Δ_i and Γ_i are dependent unless $\rho_{\Delta,\Gamma} = 0$. When $\rho_{\Delta,\Gamma} \neq 0$, the number of near-claim events will bring information about the number of claims.

4.1. Estimation of the Covariance Matrix

The estimation approach adopted in this article can be decomposed into several steps as follows. First, we perform a marginal analysis to account for individual risk profiles in the number of claims and in the number of near-claim events. Based on separate Poisson GBM regression analyses, we estimate the scores $\eta_{i,t}$ and $\nu_{i,t}$ entering the expected number of claims and near-claim events using the information available about the policyholder in the database. In a second stage, we allow for heterogeneity and we estimate the variance of the random effects separately for claims at fault and near-claim events. This is done by maximizing the likelihood associated to numbers $N_{i,t}$ of claims at fault with respect to σ_{Δ}^2 , integrating out the random effects Δ_i , on the one hand, and by maximizing the likelihood associated to numbers $Z_{i,t}$ of near-claim events with respect to σ_{Γ}^2 , integrating out the random effects Γ_i , on the other hand. Separate optimizations of these likelihoods are equivalent to setting $\rho_{\Delta,\Gamma} = 0$ because 0 correlation makes Γ_i and Δ_i independent and the joint likelihood factors into separate likelihood functions for the numbers of claims and near-claim events. The second stage thus uses 0 as an initial value for $\rho_{\Delta,\Gamma}$. Then, $\rho_{\Delta,\Gamma}$ is estimated by maximizing the joint likelihood for claims and near-claim events, integrating out the random effects (Δ_i, Γ_i) but freezing the variances σ_{Δ}^2 and σ_{Γ}^2 to the estimated values obtained before. In the third and final stage, we simultaneously estimate all elements of the variance–covariance matrix Σ of the bivariate Poisson mixture model for the number of claims and the number of near-claim events. Joint likelihood for both types of events is maximized with respect to σ_{Δ}^2 , σ_{Γ}^2 , and $\rho_{\Delta,\Gamma}$ using the estimations obtained before as starting values. This likelihood is computed by integrating out the random effects (Δ_i, Γ_i) entering the mixed Poisson model for the sequence of observed pairs $(N_{i,t}, Z_{i,t})$, by two-dimensional numerical integration. However, when estimating the elements of the variance–covariance matrix Σ of the bivariate Poisson mixture model, the scores $\eta_{i,t}$ and $\nu_{i,t}$ entering the expected number of claims and of near-claim events are fixed to their values obtained from the GBM analysis.

This estimation method is in line with the approach developed in Pechon, Trufin, and Denuit (2018). In the present study, we do not perform global likelihood optimization simultaneously over the scores $\eta_{i,t}$ and $\nu_{i,t}$ and over the parameters σ_{Δ}^2 , σ_{Γ}^2 , and $\rho_{\Delta,\Gamma}$ entering the variance–covariance matrix Σ , in contrast to Denuit, Guillen, and Trufin (2019) and Corradin et al. (2022), for instance. The reason is that Generalized Linear Model (GLM)/Generalized Additive Model (GAM) analysis of the scores given random effects in these papers is replaced with GBM in the present study. In principle, it is possible to use GBM in each cycle of the expected conditional maximization algorithm proposed by Corradin et al. (2022), but this would require considerable computation time. For this reason, we adopt the pragmatic approach of Pechon, Trufin, and Denuit (2018) to estimate the parameters entering the bivariate credibility model joining the numbers of claims at fault and near-claim events.

The likelihood associated to the observations (k_{it}, l_{it}) can be written as

$$\begin{aligned} \mathcal{L} &= \prod_{i=1}^n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(\prod_{t=1}^{T_i} e^{-\lambda_{it}e^{\delta}} \frac{(\lambda_{it}e^{\delta})^{k_{it}}}{k_{it}!} e^{-\mu_{it}e^{\gamma}} \frac{(\mu_{it}e^{\gamma})^{l_{it}}}{l_{it}!} \right) f_{\Sigma}(\delta, \gamma) d\delta d\gamma \\ &= \prod_{i=1}^n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(\prod_{t=1}^{T_i} \frac{\lambda_{it}^{k_{it}} \mu_{it}^{l_{it}}}{k_{it}! l_{it}!} \right) \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} e^{\delta \sum_{t=1}^{T_i} k_{it} + \gamma \sum_{t=1}^{T_i} l_{it} - \sum_{t=1}^{T_i} \lambda_{it}e^{\delta} - \sum_{t=1}^{T_i} \mu_{it}e^{\gamma} - \frac{1}{2}(\delta, \gamma)\Sigma^{-1}(\delta, \gamma)^{\top}} d\delta d\gamma \\ &= \prod_{i=1}^n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(\prod_{t=1}^{T_i} \frac{\lambda_{it}^{k_{it}} \mu_{it}^{l_{it}}}{k_{it}! l_{it}!} \right) \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} e^{\delta k_{i\bullet} + \gamma l_{i\bullet} - \lambda_{i\bullet}e^{\delta} - \mu_{i\bullet}e^{\gamma} - \frac{1}{2}(\delta, \gamma)\Sigma^{-1}(\delta, \gamma)^{\top}} d\delta d\gamma, \end{aligned}$$

where “ $\bullet$ ” indicates a summation over time, like $k_{i\bullet} = \sum_{t=1}^{T_i} k_{it}$, for instance. Moreover, we have $|\Sigma| = \sigma_{\Delta}^2 \sigma_{\Gamma}^2 - \rho_{\Delta,\Gamma}^2 \sigma_{\Delta}^2 \sigma_{\Gamma}^2$. The variance–covariance matrix Σ can be inverted only if $\rho_{\Delta,\Gamma} \neq \pm 1$. If the latter condition holds, then

$$\Sigma^{-1} = \frac{1}{|\Sigma|} \begin{pmatrix} \sigma_{\Gamma}^2 & -\rho_{\Delta,\Gamma}\sigma_{\Delta}\sigma_{\Gamma} \\ -\rho_{\Delta,\Gamma}\sigma_{\Delta}\sigma_{\Gamma} & \sigma_{\Delta}^2 \end{pmatrix}$$

and the exponential term can be expressed as

$$\begin{aligned}
 m &= (\delta \ \gamma) \Sigma^{-1} \begin{pmatrix} \delta \\ \gamma \end{pmatrix} \\
 &= \frac{1}{|\Sigma|} (\delta \ \gamma) \begin{pmatrix} \sigma_\Gamma^2 & -\rho_{\Delta,\Gamma} \sigma_\Delta \sigma_\Gamma \\ -\rho_{\Delta,\Gamma} \sigma_\Delta \sigma_\Gamma & \sigma_\Delta^2 \end{pmatrix} \begin{pmatrix} \delta \\ \gamma \end{pmatrix} \\
 &= \frac{1}{|\Sigma|} (\sigma_\Gamma^2 \delta^2 - 2\rho_{\Delta,\Gamma} \sigma_\Delta \sigma_\Gamma \delta \gamma + \sigma_\Delta^2 \gamma^2) \\
 &= \frac{\sigma_\Gamma^2 \delta^2 - 2\rho_{\Delta,\Gamma} \sigma_\Delta \sigma_\Gamma \delta \gamma + \sigma_\Delta^2 \gamma^2}{\sigma_\Delta^2 \sigma_\Gamma^2 - \rho_{\Delta,\Gamma}^2 \sigma_\Delta^2 \sigma_\Gamma^2}.
 \end{aligned}$$

The log-likelihood is written as

$$\log \mathcal{L} = -\frac{n}{2} \log(|\Sigma|) + \sum_{i=1}^n \log \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{\delta k_{i\bullet} + \gamma l_{i\bullet} - \lambda_{i\bullet} e^\delta - \mu_{i\bullet} e^\gamma - \frac{m}{2}} d\delta d\gamma \right) + \text{constant}.$$

The maximization of the log-likelihood is performed using the Nelder-Mead optimization method where the double integral is computed at each step using the function `dblquad` of the `scipy` library. Because of the computational time for the optimization process on the training set 2017 to 2018, we proceed in two steps. The first step consists of first estimating the random effect standard deviations σ_Δ and σ_Γ with $\rho_{\Delta,\Gamma} = 0$. We get

$$\begin{cases} \hat{\sigma}_\Delta^{(0)} = 0.61889801 \\ \hat{\sigma}_\Gamma^{(0)} = 0.73366228. \end{cases}$$

Then, we estimate the correlation coefficient $\rho_{\Delta,\Gamma}^{(0)}$ with the estimated values $(\hat{\sigma}_\Delta^{(0)}, \hat{\sigma}_\Gamma^{(0)})$ for the random effect standard deviations σ_Δ and σ_Γ , setting the initial value for $\rho_{\Delta,\Gamma}^{(0)}$ to 0.1. The objective function near the optimal solution is displayed in [Figure 9](#), which leads to $\hat{\rho}_{\Delta,\Gamma}^{(0)} = 0.57071$.

In a second step, we again perform the optimization estimating all parameters considering as initial values $(\hat{\sigma}_\Delta^{(0)}, \hat{\sigma}_\Gamma^{(0)}, \hat{\rho}_{\Delta,\Gamma}^{(0)})$. The joint estimation leads to the following estimators:

$$\begin{cases} \hat{\sigma}_\Delta^{(1)} = 0.71943756 \\ \hat{\sigma}_\Gamma^{(1)} = 0.73501123 \\ \hat{\rho}_{\Delta,\Gamma}^{(1)} = 0.54464738. \end{cases}$$

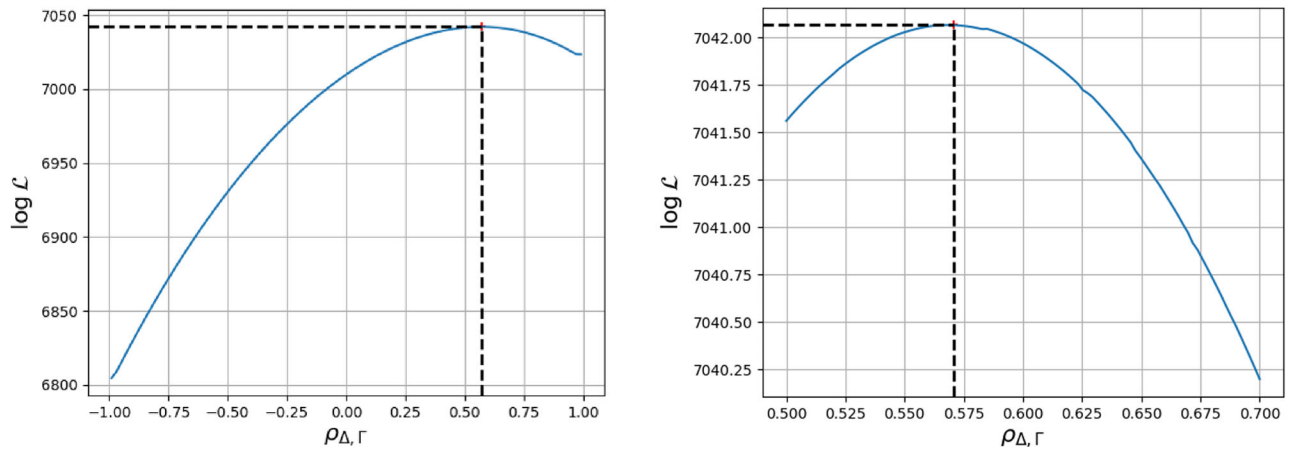


FIGURE 9. Maximum Likelihood Estimator $\hat{\rho}_{\Delta,\Gamma}^{(0)}$ in the (a) Entire Solution Space and (b) Nearby Optimal Solution.

The pretty large estimated correlation coefficient indicates that near-claim events bring relevant information about claims at fault.

4.2. Corrections Based on Experience

4.2.1. Univariate Credibility Model

In the univariate case, the expectation of the number of claims at fault for the year $T_i + 1$ given the claim history $\mathcal{H}_{T_i} = \{k_{i,1}, \dots, k_{i,T_i}\}$ is given by

$$E[N_{i,T_i+1} | \mathcal{H}_{T_i}] = \lambda_{i,T_i+1} E[\exp(\Delta_i) | \mathcal{H}_{T_i}],$$

where

$$E[\exp(\Delta_i) | \mathcal{H}_{T_i}] = \frac{\int_{-\infty}^{+\infty} e^{-\lambda_{i\bullet} e^\delta + \delta(k_{i\bullet} + 1) - \frac{\delta^2}{2\sigma_\Delta^2}} d\delta}{\int_{-\infty}^{+\infty} e^{-\lambda_{i\bullet} e^\delta + \delta k_{i\bullet} - \frac{\delta^2}{2\sigma_\Delta^2}} d\delta}.$$

4.2.2. Multivariate Case

In the multivariate case, the expectation of the number of claims for the year $T_i + 1$ given the claim and near-claim history $\mathcal{H}_{T_i}^* = \{k_{i,1}, \dots, k_{i,T_i}, l_{i,1}, \dots, l_{i,T_i}\}$ is given by

$$E[N_{i,T_i+1} | \mathcal{H}_{T_i}^*] = \lambda_{i,T_i+1} E[\exp(\Delta_i) | \mathcal{H}_{T_i}^*],$$

where

$$E[\exp(\Delta_i) | \mathcal{H}_{T_i}^*] = \frac{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{\delta(k_{i\bullet} + 1) + \gamma l_{i\bullet} - \lambda_{i\bullet} e^\delta - \mu_{i\bullet} e^\gamma - \frac{\delta^2}{2\sigma_\Delta^2} - \frac{\gamma^2}{2\sigma_\Gamma^2}} d\delta d\gamma}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{\delta k_{i\bullet} + \gamma l_{i\bullet} - \lambda_{i\bullet} e^\delta - \mu_{i\bullet} e^\gamma - \frac{\delta^2}{2\sigma_\Delta^2} - \frac{\gamma^2}{2\sigma_\Gamma^2}} d\delta d\gamma}.$$

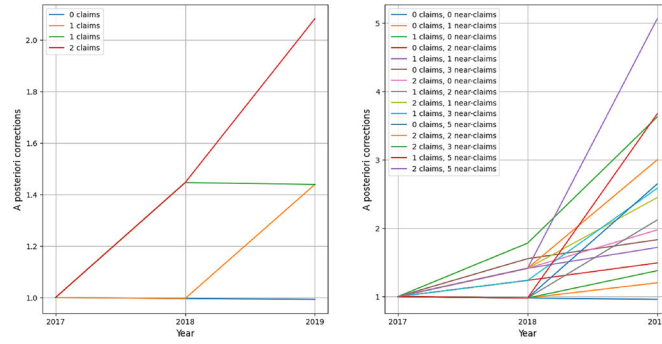
4.2.3. Comparison

Let us consider four different profiles corresponding to two different risk levels for claim and near-claim predicted frequencies, namely,

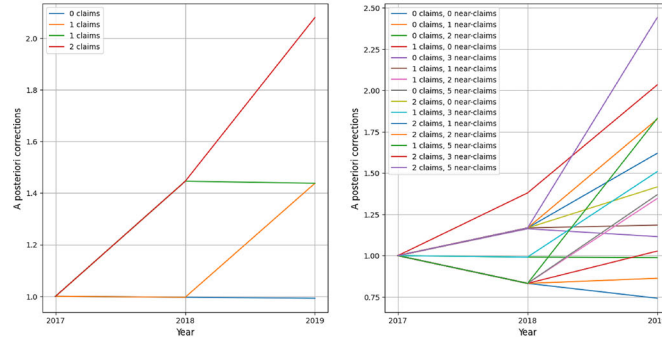
- Profile 1: Low claim and near-claim frequency
- Profile 2: Low claim and high near-claim frequency
- Profile 3: High claim and low near-claim frequency
- Profile 4: High claim and near-claim frequency.

The low (respectively high) frequency profile corresponds to the one-sixth-quantile (respectively five-sixths-quantile) of the distribution of fitted frequencies. For each profile, we let the number of claims and near-claim events vary and compute the corresponding corrections based on experience according to univariate and bivariate credibility models. The latter are displayed in [Figure 10](#) with claim numbers $\{0, 1, 2\}$ and near-claim numbers $\{0, 1, 2, 3, 5\}$. Corrections are applied for 2018 when 2017 history becomes available and for 2019 based on 2017 to 2018 history. For the univariate model, if the policyholder claim history is equal to 1, we represent the case where the claim occurred in 2017 and the case where the claim occurred in 2018, which leads mechanically to the same correction for 2019. For the near-claim history, we did not consider all possible combinations leading to the same near-claim history to avoid overloading figures, but corrections would be identical because as the random effects Δ_i and Γ_i do not depend on time. [Figure 10](#) clearly shows that near-claim events considerably enrich corrections based on experience, resulting in a much finer grid of reevaluations at the end of the period.

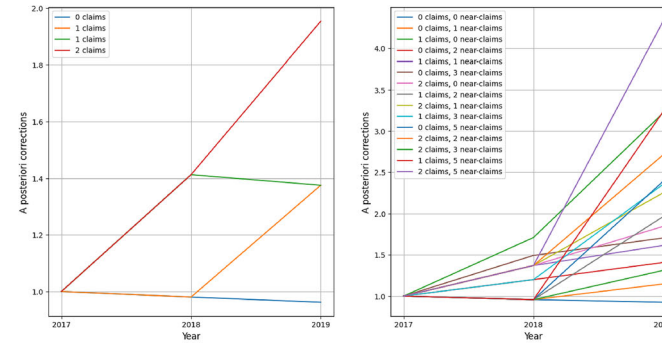
We observe that Profile 1 experiences the highest correction associated with an history of two claims at fault and five near-claim events observed, whereas Profile 4 experiences the lowest correction associated with an empty history of claims at fault and near-claim events irrespective of the univariate or bivariate model. This reflects the fact that a profile with a low



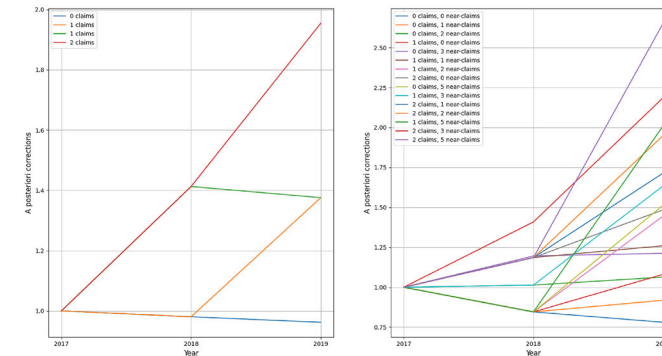
(a) Profile 1.



(b) Profile 2.



(c) Profile 3.



(d) Profile 4.

FIGURE 10. Experience Corrections Resulting from Univariate (Left) and Bivariate (Right) Credibility Models for Each of the Four Risk Profiles under Consideration: (a) Profile 1, (b) Profile 2, (c) Profile 3, and (d) Profile 4.

(respectively high) risk level and a contradictory claim and near-claim history will naturally experience the largest modification of their a priori frequency premium. Note also that for Profiles 1 and 2, the univariate correction is slightly softer. Moreover, we can see that the bivariate model better rewards policyholders than the univariate model in case of an empty near-claim history.

To end the comparison between both models, we show in Table 4 the estimated mean Poisson deviance and mean squared error obtained for 2018 and 2019. Note that the year 2017 is not displayed in Table 4 because no correction is applied with no history of claim at fault and near-claim events available. We observe that the use of near-claim events count improves the a posteriori claim frequency refinement on the MTPL portfolio in the sense that the estimated mean Poisson deviance and mean squared error are the lowest for the multivariate model.

5. BONUS–MALUS SCALE WITH NEAR-CLAIM EVENTS

Bonus–malus scales are commercial simplifications to credibility corrections. A bonus–malus scale consists of a finite number of levels, each with its own relative premium called “relativity”. The relativity associated to level ℓ is denoted as r_ℓ . The premium charged is then obtained by multiplying by $r_\ell\%$ a base premium depending on the policyholder’s characteristics.

Whereas classical bonus–malus scales are based on a single type of event (occurrence of claims at fault), this section integrates near-claim events as a second type of event. Notice that systems based on several types of events are commonly encountered in North America where elements of the policyholder’s driving record also impact insurance premiums. For instance, safe driver insurance plans in the United States encourage safe driving by rewarding drivers who do not cause an accident or incur a traffic law violation. Several types of events (major and minor at-fault accidents and traffic violations) each entail specific penalties: policyholders move up a certain number of levels in the bonus–malus scale according to the type of incident. Compared to Pitrebois, Denuit, and Walhin (2006), who developed multiple-event bonus–malus scales by assuming a multinomial partitioning scheme, we work here with the bivariate credibility model described in Section 4. Because it seems to be difficult in practice to penalize drivers for near-claim events, we adapt the $-1/\text{top}$ scale in Pitrebois, Denuit, and Walhin (2003) to account for both claim and near-claim counts. This scale is a pure no-claim discount scheme classifying drivers according to the number of claim-free periods. It is commonly encountered in North America. Here, we adapt it by letting the annual discount depend on the near-claim events, not the penalty. This is precisely explained next.

The $-1/\text{top}$ bonus–malus scale under consideration has $m + 1$ levels (numbered 0 to m , where m denotes the maximum level). Policyholders are classified according to the number of claim-free periods since the last event (0, 1, ..., $m - 1$ or at least m). New policyholders enter the system at the top level m where they pay the base premium (so that $r_m = 100\%$). Each event-free period (no claim reported or near-claim event recorded) is rewarded by one bonus class. If no claim is reported but at least one near-claim event is recorded, then the policyholder stays in his or her current level, thus losing the discount. If a claim is reported, all of the discounts are lost and the policyholder is transferred to level m .

The policyholder’s trajectory in the scale is modeled by a sequence $\{L_t, t = 0, 1, \dots\}$ of random variables valued in $\{0, 1, \dots, m\}$, with $L_0 = m$. Movements in the scale occur on a regular basis (yearly or monthly, say) and the policyholder occupies level L_t from time t until time $t + 1$. Once the number N_t of reported claims at fault during $(t - 1, t)$ and the number Z_t of near-claim events are both known, this information is used to reevaluate the position in the scale. Random variables L_t obey to the recursion

TABLE 4
Models’ Performance for 2018 and 2019 Calendar Years

Year Model	2018			2019		
	A priori	Univariate	Multivariate	A priori	Univariate	Multivariate
Mean Poisson deviance	0.298	0.287	0.281	0.313	0.217	0.212
Mean squared error	0.102	0.101	0.099	0.123	0.118	0.116

$$L_t = \begin{cases} \max\{L_{t-1} - 1, 0\} \text{ if } N_t = Z_t = 0 \\ L_{t-1} \text{ if } N_t = 0 \text{ and } Z_t \geq 1 \\ \text{mif } N_t \geq 1. \end{cases}$$

This stochastic recursive equation shows that the future trajectory in the scale is independent of the levels occupied in the past, provided that the present level is given. This conditional independence property is at the heart of Markov models.

We follow Pitrebois, Denuit, and Walhin (2003) to determine relativities r_ℓ associated with each level $\ell \in \{0, 1, \dots, m\}$. Let us drop the policy index i and let us consider a generic policyholder with $\Delta = \delta$ and $\Gamma = \gamma$, taken at random from the portfolio. Given $\Delta = \delta$ and $\Gamma = \gamma$, all random variables N_t and Z_t are mutually independent, obeying Poisson distributions with respective means denoted as λ and μ . Let $p_{\ell_1 \ell_2}(\lambda, \mu)$ be the probability of moving from level ℓ_1 to level ℓ_2 for that policyholder; that is, the probability that $L_{t+1} = \ell_2$ given that $L_t = \ell_1$, with $\ell_1, \ell_2 \in \{0, 1, \dots, m\}$. These probabilities are gathered in the transition matrix $\mathbf{P}(\lambda, \mu)$ assumed to be regular (that is, there exists some power of $\mathbf{P}(\lambda, \mu)$ whose entries are all strictly positive). For the bonus–malus scale considered in this article, we get

$$\mathbf{P}(\lambda, \mu) = \begin{pmatrix} \exp(-\lambda) & 0 & 0 & \cdots & 0 & 0 & 1 - \exp(-\lambda) \\ \exp(-\lambda - \mu) & \exp(-\lambda)(1 - \exp(-\mu)) & 0 & \cdots & 0 & 0 & 1 - \exp(-\lambda) \\ 0 & \exp(-\lambda - \mu) & \exp(-\lambda)(1 - \exp(-\mu)) & \cdots & 0 & 0 & 1 - \exp(-\lambda) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \exp(-\lambda - \mu) & \exp(-\lambda)(1 - \exp(-\mu)) & 1 - \exp(-\lambda) \\ 0 & 0 & 0 & \vdots & 0 & \exp(-\lambda - \mu) & 1 - \exp(-\lambda - \mu) \end{pmatrix}.$$

The probability $p_{\ell_1 \ell_2}^{(k)}(\lambda, \mu)$ that $L_{t+k} = \ell_2$ given $L_t = \ell_1$ of being transferred from level ℓ_1 to level ℓ_2 in k steps is then obtained from the k th power of the one-step transition matrix $\mathbf{P}(\lambda, \mu)$.

It is natural to expect that the system stabilizes in the long run, in that each policyholder ultimately oscillates around an equilibrium level corresponding to the expected number λ of reported claims at fault and expected number μ of near-claim events recorded by the telematics device. Formally, define the stationary distribution $\pi_\ell(\lambda, \mu)$ as

$$\pi_{\ell_2}(\lambda, \mu) = \lim_{k \rightarrow \infty} p_{\ell_1 \ell_2}^{(k)}(\lambda, \mu).$$

Here, $\pi_\ell(\lambda, \mu)$ can be regarded as the fraction of the time a policyholder with expected claim frequency λ and expected near-claim frequency μ spends in level ℓ , once the steady state has been reached. The stationary distribution can be obtained by solving a system of linear equations. It can also be obtained by matrix calculation. Specifically, let $\mathbf{I}$ be a column vector with all components equal to 1 and let $\mathbf{E}$ be the $(m+1) \times (m+1)$ matrix consisting of $m+1$ column vectors $\mathbf{I}$. Denoting as $\boldsymbol{\pi}(\mu)$ the vector of stationary probabilities is then given by

$$\boldsymbol{\pi}^\top(\lambda, \mu) = \mathbf{I}^\top (\mathbf{I} - \mathbf{P}(\lambda, \mu) + \mathbf{E})^{-1}. \quad (5.1)$$

The number of levels can be selected so that a policyholder drawn at random from the portfolio reaches level 0 in a given number of periods, on average, or derived from some other meaningful condition. Notice that basing calculation of r_ℓ on the stationary distribution is a conservative strategy with the –1/top scale.

Let w_k be the weight of the k th risk class with expected claim frequency λ_k per period and expected near-claim frequency μ_k per period. A very accurate risk classification (like the one produced by GBM in this article) can result in a large number of risk classes k to consider. Model points or reference risk profiles can then be created to ease calculations (applying Classification And Regression Trees [CARTs], for instance). In this section, we consider all risk profiles so that we actually sum over all policyholders ($k = i$), taking individual exposure to risk as weight w_k .

Let L be the level occupied by a randomly selected policyholder once the steady state has been reached. The distribution of L can be written as

$$P[L = \ell] = \sum_k w_k \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \pi_\ell(\lambda_k \exp(\delta), \mu_k \exp(\gamma)) f_\Sigma(\delta, \gamma) d\delta d\gamma. \quad (5.2)$$

Here, $P[L = \ell]$ represents the proportion of the policyholders in level ℓ . Starting from the seminal work by Norberg (1976), there are several ways to compute relativities r_ℓ in the presence of risk classification since Taylor (1997) pointed out the interaction between a priori and a posteriori premium evaluations. Pitrebois, Denuit, and Walhin (2003) suggested minimizing the expected square difference between the true relative premium and the relativity r_L applicable to the policyholder once the steady state has been reached. Relativities are then obtained from

$$r_\ell = \frac{E[\exp(\Delta)|L = \ell]}{E[\exp(\Delta)]},$$

where

$$E[\exp(\Delta)|L = \ell] = \frac{\sum_k w_k \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp(\delta) \pi_\ell(\lambda_k \exp(\delta), \mu_k \exp(\gamma)) f_\Sigma(\delta, \gamma) d\delta d\gamma}{\sum_k w_k \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \pi_\ell(\lambda_k \exp(\delta), \mu_k \exp(\gamma)) f_\Sigma(\delta, \gamma) d\delta d\gamma}, \quad (5.3)$$

as demonstrated by Pitrebois, Denuit, and Walhin (2003). It is easily seen that $E[r_L] = 1$, resulting in financial equilibrium once the steady state is reached. This approach indeed reduces to the formulas derived by Norberg (1976) in the absence of a priori risk classification, but the resulting optimal relativities may not minimize the expected squared difference between the true premium in the bivariate credibility model and the actual premium paid by the policyholder occupying level L in the scale, as pointed out by Tan et al. (2015). See also Oh, Lee et al. (2020). The relativities minimizing the expected squared difference between the true premium and the actual premium as derived in Tan et al. (2015) are expressed in our case as

$$r_\ell = \frac{E[\Lambda^2 \exp(\Delta)|L = \ell]}{E[\Lambda^2|L = \ell]} - \frac{\alpha}{2E[\Lambda^2|L = \ell]}$$

with

$$\alpha = \frac{\left(\sum_{\ell=0}^m \frac{E[\Lambda^2 \exp(\Delta)|L=\ell] P[L=\ell]}{E[\Lambda^2|L=\ell]} \right) - 1}{\sum_{\ell=0}^m \frac{P[L=\ell]}{2E[\Lambda^2|L=\ell]}},$$

where

$$E[\Lambda^2 \exp(\Delta)|L = \ell] = \frac{\sum_k w_k \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \lambda_k^2 \exp(\delta) \pi_\ell(\lambda_k \exp(\delta), \mu_k \exp(\gamma)) f_\Sigma(\delta, \gamma) d\delta d\gamma}{\sum_k w_k \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \pi_\ell(\lambda_k \exp(\delta), \mu_k \exp(\gamma)) f_\Sigma(\delta, \gamma) d\delta d\gamma}$$

and

$$E[\Lambda^2|L = \ell] = \frac{\sum_k w_k \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \lambda_k^2 \pi_\ell(\lambda_k \exp(\delta), \mu_k \exp(\gamma)) f_\Sigma(\delta, \gamma) d\delta d\gamma}{\sum_k w_k \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \pi_\ell(\lambda_k \exp(\delta), \mu_k \exp(\gamma)) f_\Sigma(\delta, \gamma) d\delta d\gamma}.$$

When implemented in practice, it is very likely that the relativity r_m associated to the highest level in the scale differs from 100%. It then suffices to adapt the intercept in the frequency model to re-normalize it by dividing all relativities by r_m and change the intercept β_0 into $\beta_0 + \ln r_m$ in the Poisson regression model producing frequency premiums. Relativities r_ℓ can be made linear in ℓ to entail the same discount when policyholders get one level down.

The relativities based on the MTPL insurance portfolio considered in this article are computed based on the risk classes produced by the GBM. More precisely, we consider the a priori claim and near-claim frequencies for each policyholder of the portfolio entering the computation of stationary probabilities as well as integrals for each individual. As part of this approach, integrals involved in (5.3) have been approximated by the Simpson's rule using `python` package `scipy`. We parallelized the

TABLE 5
Characteristics of the $-1/\text{Top}$ Bonus–Malus Scale Integrating Both Claim and Near-Claim Events

Level	$P[L = \ell]$	Pitrebois et al. (2003)		Tan et al. (2015)	
		Relativity r_ℓ	Norm. relativity	Relativity r_ℓ	Norm. relativity
5	0.2353457	1.3967499	1.0	1.5944239	1.0
4	0.0872465	1.196706	0.8567790	1.2557890	0.787613
3	0.0601058	1.1234858	0.8043572	1.1453890	0.7183717
2	0.0466335	1.0735371	0.7685966	1.0696886	0.6708935
1	0.0382324	1.0349771	0.7409896	1.0107743	0.6339433
0	0.5324361	0.7695047	0.5509252	0.6720503	0.4215004

TABLE 6
Characteristics of the $-1/\text{Top}$ Bonus–Malus Scale Integrating Claims Only

Level	$P[L = \ell]$	Pitrebois et al. (2003)		Tan et al. (2015)	
		Relativity r_ℓ	Norm. relativity	Relativity r_ℓ	Norm. relativity
5	0.0623249	1.4247254	1.0	1.634531	1.0
4	0.0549567	1.353324	0.9498841	1.5052132	0.9208839
3	0.0489845	1.2971699	0.9104702	1.4075126	0.8611110
2	0.0440318	1.2509817	0.8780512	1.3291116	0.8131456
1	0.0398543	1.2118683	0.8505979	1.2637949	0.7731850
0	0.7498477	0.8933913	0.6270621	0.8502649	0.5201889

computation of the integrals for each individual, reducing the computation time of the relativities for our MTPL insurance portfolio to 128 minutes on a 12-core machine.

The relativities associated to each level of the scale are reported in Table 5. We observe that based on the transition rules considered, level 5 is occupied by 23.5% of the portfolio. Level 0 is occupied by 53.2% of the portfolio benefiting from the maximum discount, whereas other levels accommodate a lower proportion of policyholders. We observe that after reaching the lowest level of the scale, the discount is about 45% as per Pitreboi, Denuit, and Walhin's (2003) relativities and 58% with Tan et al.'s (2015) relativities.

We also compare the relativities when near-claim events are taken into account to those of the standard $-1/\text{top}$ scale based on claims only. The levels L_t occupied in the scale then obey the recursion

$$L_t = \begin{cases} \max\{L_{t-1} - 1, 0\} & \text{if } N_t = 0 \\ \text{mif } N_t \geq 1. \end{cases}$$

Table 6 shows that policyholders are more concentrated in the lower levels of the bonus–malus scale when only claims at fault are considered. As expected, Table 6 shows that the univariate system grants less discount. The normalized relativities are larger for all levels 0 to 4 compared to Table 5. The maximum discount is reduced from 45% to about 37% as per Pitrebois, Denuit, and Walhin's (2003) relativities and from 58% to 48% with Tan et al.'s (2015) relativities.

6. CONCLUSION

Based on telematics signals recorded in an MTPL insurance portfolio, we proposed a simple and effective definition of near-claims based on deceleration events. As for claim frequency, we use traditional insurance explanatory variables to explain the near-claim frequency. Both responses disclose similar partial dependencies with numerical explanatory variables but exhibit differences with high-dimensional categorical explanatory variables. Then, we integrated the relevant information about

near-claim events related to driving experience with the help of a bivariate Poisson-Lognormal credibility model. We demonstrated the added value of the bivariate model as the corrections resulting from the bivariate model adapt more dynamically to driving experience over the contract duration.

Then, we designed a commercial simplification as a –1/top bonus–malus scale integrating both claim and near-claim events but penalizing policyholders only for reporting claims at fault. More sophisticated transition rules could be considered. For instance, near-claim events could be used to transfer policyholders to upper levels in the scale.

The information used in this article is limited to the number of claims at fault and one type of near-claim event. Following Pitrebois, Denuit, and Walhin (2006), we could distinguish between different types of claims and near-claim events. Considering claims at fault, accidents with bodily injuries could be counted separately from accidents with material damage only. This is a convenient, though indirect, way to account for the severity of the claims because the former ones cost much more on average. Also, several types of near-claim events could be considered. Multivariate credibility models joining claim frequencies and severities have also been proposed and applied to the design of bonus–malus scales, with important implication, such as accounting for the bonus hunger phenomenon, for instance. See, for example, Park, Kim, and Ahn (2018), Oh, Shi, and Ahn (2020), and Oh, Kim, and Ahn (2022). Including claim severities in a posteriori premium corrections is challenging, due, for instance, to the sometimes long settlement procedures in motor third-party liability insurance so that the final claim cost is not known at the end of the occurrence period when the amount of premium must be updated based on past experience. But severities certainly bring some added information worthy of inclusion in the analysis. Therefore, this could be a promising extension of the model proposed in the present article for future research.

ACKNOWLEDGEMENTS

The authors thank their colleagues from P&V Assurances, especially Nikias De Feyter and Carolina Aldana, for interesting discussions that greatly contributed to the success of this research project. The authors also thank the anonymous referees and the editor for their constructive comments that helped to improve this article.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author(s).

FUNDING

The authors gratefully acknowledge the financial support of P&V Assurances through the research chair "Telematics and Prevention."

ORCID

Julien Trufin  <http://orcid.org/0000-0001-6690-9020>

REFERENCES

- Apley, D. W., and J. Zhu. 2016. Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society Series B* 82:1059–86. [10.1111/rssb.12377](https://doi.org/10.1111/rssb.12377)
- Ayuso, M., M. Guillen, and J. P. Nielsen. 2019. Improving automobile insurance ratemaking using telematics: Incorporating mileage and driver behaviour data. *Transportation* 46:735–52. [10.1007/s11116-018-9890-7](https://doi.org/10.1007/s11116-018-9890-7)
- Boucher, J.-P., S. Côté, and M. Guillen. 2017. Exposure as duration and distance in telematics motor insurance using generalized additive models. *Risks* 5: 54. [10.3390/risks5040054](https://doi.org/10.3390/risks5040054)
- Cheung, E. C., W. Ni, R. Oh, and J. K. Woo. 2021. Bayesian credibility under a bivariate prior on the frequency and the severity of claims. *Insurance: Mathematics and Economics* 100:274–95. [10.1016/j.insmatheco.2021.06.003](https://doi.org/10.1016/j.insmatheco.2021.06.003)
- Chevalier, A., K. Coxon, A. J. Chevalier, E. Clarke, K. Rogers, J. Brown, S. Boufous, R. Ivers, and L. Keay. 2016. Predictors of older drivers' involvement in rapid deceleration events. *Accident Analysis & Prevention* 98:312–19. [10.1016/j.aap.2016.10.010](https://doi.org/10.1016/j.aap.2016.10.010)
- Corradin, A., M. Denuit, M. Detyniecki, V. Grari, M. Sammarco, and J. Trufin. 2022. Joint modeling of claim frequencies and behavioral signals in motor insurance. *ASTIN Bulletin* 52:33–54. [10.1017/asb.2021.24](https://doi.org/10.1017/asb.2021.24)
- De Ceunynck, T. 2017. Defining and applying surrogate safety measures and behavioural indicators through site-based observations. PhD diss., Department of Technology and Society, Lund University.
- Denuit, M., M. Guillen, and J. Trufin. 2019. Multivariate credibility modeling for usage-based motor insurance pricing with behavioral data. *Annals of Actuarial Science* 13:378–99. [10.1017/S1748499518000349](https://doi.org/10.1017/S1748499518000349)

- Denuit, M., and Y. Lu. 2021. Wishart-Gamma random effects models with applications to nonlife insurance. *Journal of Risk and Insurance* 88:443–81. [10.1111/jori.12327](#)
- Dionne, G., and C. Vanasse. 1992. Automobile insurance ratemaking in the presence of asymmetrical information. *Journal of Applied Econometrics* 7:149–65. [10.1002/jae.3950070204](#)
- Duval, F., J.-P. Boucher, and M. Pigeon. 2022. How much telematics information do insurers need for claim classification? *North American Actuarial Journal* 26:570–90. [10.1080/10920277.2021.2022499](#)
- Friedman, J. H. 2008. Greedy function approximation: A gradient boosting machine. *Annals of Statistics* 29:1189–232.
- Fu, C., and T. Sayed. 2021. Comparison of threshold determination methods for the deceleration rate to avoid a crash (DRAC)-based crash estimation. *Accident Analysis & Prevention* 153:106051. [10.1016/j.aap.2021.106051](#)
- Gao, G., S. Meng, and M. Wüthrich. 2019. Claims frequency modeling using telematics car driving data. *Scandinavian Actuarial Journal* 2019:143–62. [10.1080/03461238.2018.1523068](#)
- Gao, G., S. Meng, and M. Wüthrich. 2022. What we can learn from telematics car driving data: A survey. *Mathematics and Economics* 104:185–99. [10.1016/j.insmatheco.2022.02.004](#)
- Gao, G., H. Wang, and M. Wüthrich. 2022. Boosting Poisson regression models with telematics car driving data. *Machine Learning* 111:243–72. [10.1007/s10994-021-05957-0](#)
- Gomez-Deniz, E. 2016. Bivariate credibility bonus–malus premiums distinguishing between two types of claims. *Insurance: Mathematics and Economics* 70:117–24. [10.1016/j.insmatheco.2016.06.009](#)
- Guillen, M., J. Nielsen, and A. M. Pérez-Marín. 2021. Near-miss telematics in motor insurance. *Journal of Risk and Insurance* 88:569–89. [10.1111/jori.12340](#)
- Guillen, M., J. P. Nielsen, A. M. Pérez-Marín, and V. Elpidorou. 2020. Can automobile insurance telematics predict the risk of near-miss events? *North American Actuarial Journal* 24:141–52. [10.1080/10920277.2019.1627221](#)
- Norberg, R. 1976. A credibility theory for automobile bonus system. *Scandinavian Actuarial Journal* 1976:92–107. [10.1080/03461238.1976.10405605](#)
- Oh, R., J. H. Kim, and J. Y. Ahn. 2022. Designing a bonus–malus system reflecting the claim size under the dependent frequency–severity model. *Probability in the Engineering and Informational Sciences* 36:963–87. [10.1017/S0269964821000188](#)
- Oh, R., K. S. Lee, S. C. Park, and J. Y. Ahn. 2020. Double-counting problem of the bonus–malus system. *Insurance: Mathematics and Economics* 93:141–55. [10.1016/j.insmatheco.2020.04.008](#)
- Oh, R., P. Shi, and J. Y. Ahn. 2020. Bonus–malus premiums under the dependent frequency–severity modeling. *Scandinavian Actuarial Journal* 2020:172–95. [10.1080/03461238.2019.1655477](#)
- Park, S. C., J. H. Kim, and J. Y. Ahn. 2018. Does hunger for bonuses drive the dependence between claim frequency and severity? *Insurance: Mathematics and Economics* 83:32–46. [10.1016/j.insmatheco.2018.09.002](#)
- Pechon, F., J. Trufin, and M. Denuit. 2018. Multivariate modelling of household claim frequencies in motor third-party liability insurance. *ASTIN Bulletin* 48:969–93. [10.1017/asb.2018.21](#)
- Pinquet, J. 1998. Designing optimal bonus–malus systems from different types of claims. *ASTIN Bulletin* 28:205–20. [10.2143/AST.28.2.519066](#)
- Pitrebois, S., M. Denuit, and J.-F. Walhin. 2003. Setting a bonus–malus scale in the presence of other rating factors: Taylor’s work revisited. *ASTIN Bulletin* 33:419–36. [10.1017/S0515036100013544](#)
- Pitrebois, S., M. Denuit, and J.-F. Walhin. 2006. Multi-event bonus–malus scales. *Journal of Risk and Insurance* 73:517–28. [10.1111/j.1539-6975.2006.00186.x](#)
- So, B., J.-P. Boucher, and E. Valdez. 2021. Synthetic dataset generation of driver telematics. *Risks* 9:58. [10.3390/risks9040058](#)
- Sun, S., J. Bi, M. Guillen, and A. M. Pérez-Marín. 2021. Driving risk assessment using near-miss events based on panel Poisson regression and panel negative binomial regression. *Entropy* 23:829. [10.3390/e23070829](#)
- Tan, C. I., J. Li, J. S. H. Li, and U. Balasooriya. 2015. Optimal relativities and transition rules of a bonus–malus system. *Insurance: Mathematics and Economics* 61:255–63. [10.1016/j.insmatheco.2015.02.001](#)
- Taylor, G. 1997. Setting a bonus–malus scale in the presence of other rating factors. *ASTIN Bulletin* 27:319–27. [10.2143/AST.27.2.542055](#)
- Tzougas, G., and A. P. di Cerchiara. 2021. The multivariate mixed negative binomial regression model with an application to insurance a posteriori ratemaking. *Insurance: Mathematics and Economics* 101:602–25. [10.1016/j.insmatheco.2021.10.001](#)
- Verschuren, R. M. 2022. Frequency–severity experience rating based on latent Markovian risk profiles. *Insurance: Mathematics and Economics* 107:379–92. [10.1016/j.insmatheco.2022.09.007](#)
- Zheng, L., T. Sayed, and M. Essa. 2018. Validating the bivariate extreme value modeling approach for road safety estimation with different traffic conflict indicators. *Accident Analysis & Prevention* 123:314–23. [10.1016/j.aap.2018.12.007](#)

APPENDIX A. EXPLANATORY VARIABLES

TABLE A.1
Explanatory Variables Description

Related to	Variable	Type	Unique values	Description
Driver	Postcode longitude	Float	–	Longitude of the policyholder's municipality
	Postcode latitude	Float	–	Latitude of the policyholder's municipality
	Main driver age	Integer	–	Age of the main driver insured
	Vehicle mileage limit	Binary	2	If the policyholder drives more than 10k km yearly
	Contract paysplit	Categorical	4	Premium payment frequency
	Vehicle use	Categorical	4	Use of the vehicle
	Number of drivers	Categorical	5	Number of drivers insured
	Mosaic code	Categorical	9	Policyholder's municipality socio-demographic code
Vehicle	Vehicle power	Float	–	Power of the vehicle (in kWh)
	Vehicle weight	Float	–	Weight of the vehicle (in kg)
	Vehicle age	Float	–	Age of the vehicle
	Vehicle value	Float	–	Value of the vehicle (in EUR)
	Vehicle adas	Binary	2	If the vehicle is equipped with any driving assisting system
	Vehicle garage	Binary	2	If the vehicle is parked in a garage
	Vehicle fuel	Categorical	7	Vehicle fuel
	Vehicle segmentation	Categorical	7	Vehicle segmentation category
	Vehicle seats	Categorical	8	Number of seats in the vehicle
	Vehicle make	Categorical	32	Make of the vehicle