

RESEARCH ARTICLE

Generic framework for a coherent integration of experience and exposure rating in reinsurance

Stefan Bernegger 

Independent Scholar, Guldenenstrasse 7B, CH-8127, Forch, Switzerland
Email: st.bernegger@gmail.com

Received: 3 October 2023; **Revised:** 11 April 2024; **Accepted:** 12 April 2024

Keywords: Experience rating; exposure rating; short-tail; long-tail; generative model; Bayesian inference; reduced statistics; model calibration; Monte Carlo simulation

Abstract

This article introduces a comprehensive framework that effectively combines *experience rating* and *exposure rating* approaches in reinsurance for both *short-tail* and *long-tail* businesses. The generic framework applies to all non-life lines of business and products emphasizing nonproportional treaty business. The approach is based on three pillars that enable a coherent usage of all available information. The first pillar comprises an exposure-based *generative model* that emulates the *generative process* leading to the observed claims experience. The second pillar encompasses a standardized *reduction procedure* that maps each high-dimensional claim object to a few weakly coupled *reduced random variables*. The third pillar comprises calibrating the generative model with retrospective *Bayesian inference*. The derived calibration parameters are fed back into the generative model, and the reinsurance contracts covering future cover periods are rated by projecting the calibrated generative model to the cover period and applying the future contract terms.

1. Introduction

There are two approaches to probability theory: the first is the *frequentist* interpretation, also known as *standard*, *orthodox*, or *classical* statistics, and the second is the *Bayesian* interpretation (Daykin *et al.*, 1994; Lee, 2012; Gelman, 2013; Klugman *et al.*, 2019; Murphy, 2022, 2023). This distinction also applies to the various statistical and modeling approaches.

The Bayesian interpretation has multiple advantages, but frequentist approaches are still popular and widely used for two main reasons. First, they are computationally ‘cheap’ compared to Bayesian methods, and second, they do not presuppose knowledge regarding priors. Bayesian inference is, nevertheless, widely applied in various domains of science (Parr *et al.*, 2022; Palmer, 2022; Harrison *et al.*, 2023), and it has become the foundation of machine learning (Murphy, 2022, 2023). Bayesian approaches are also gaining popularity in disciplines like insurance and reinsurance, which have traditionally been dominated by frequentist approaches (Bühlmann and Straub, 1970; Bühlmann and Gisler, 2005; Zhang *et al.*, 2012; Klugman *et al.*, 2019; Goffard *et al.*, 2023).

Reinsurance companies face multiple challenges when using models for assessing the assumed risks and for taking business decisions like risk selection, transactional pricing, reserving, investing, and corporate risk management, and they rely on a variety of methods and tools for managing these risks (Antal, 2009; Albrecher, 2017; Mildenhall, 2022). The actuarial models implemented by a reinsurer are based on simplified assumptions, and reinsurers know that the underlying assumptions are rarely fulfilled in the real world. One way to deal with the issue is to complement experience-based rating models with exposure-based models. The latter can incorporate all relevant features of the generative process that affect the rating. All such features must be calibrated, which is accomplished with market data. However, cedent-specific features (hidden variables) must also be reflected in the rating of reinsurance contracts.

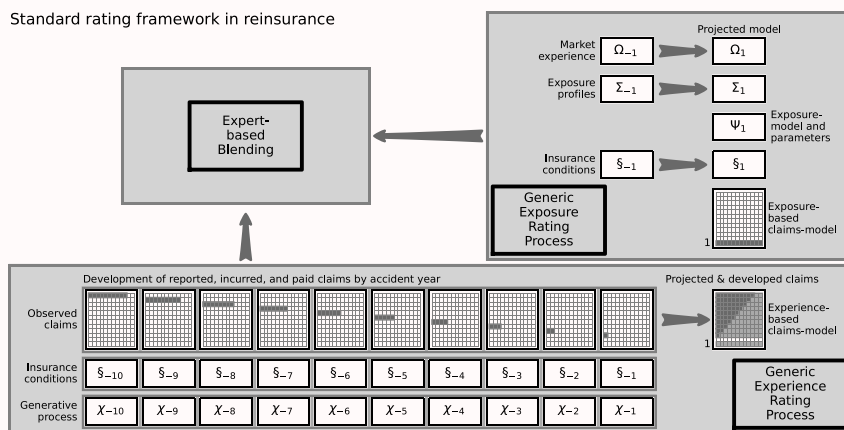


Figure 1. The standard rating framework in reinsurance comprises experience rating, exposure rating, and expert-based blending. Bottom panel: Experience rating is based on the observed claims reported by the cedent. The development of these claims with occurrence years $o \in \{-T, \dots, -1\}$ (with $T = 10$) results from applying policy conditions \S_o to the economic loss χ_o emanating from the generative process. The claims experience is projected to the future cover period $o = +1$ and developed. The statistics derived from the claims triangle for $o = +1$ is used to rate the reinsurance contracts. Upper right panel: Exposure rating combines the cedent’s most recent exposure profiles Σ_{-1} with market models Ω_{-1} and underwriting expertise. The profiles and the market models are projected to the cover period and processed in the exposure model Ψ_1 . The rating-relevant statistics for $o = +1$ are obtained by applying the anticipated insurance conditions \S_1 .

This is done via an ‘expert-based’ credibility weighting of experience-based rates and exposure-based rates. One such approach is to evaluate the experience-based rate for a ‘working layer’ and to use this rate for calibrating the exposure-based rates for high layers (Desmedt *et al.*, 2012).

The framework presented in this article attempts to resolve the calibration issue in an ‘optimal’ way. This is accomplished by using the claims experience of a cedent and *Bayesian inference* to adjust the calibration parameters within the exposure-based generative model. The approach is generic, and it coherently integrates all available information.

1.1. Motivation

This article claims that the integrated rating framework introduced in the following sections is superior to standard rating approaches applied by reinsurers (see Figures 1, 2, and Table 1) and that it can support reinsurers in optimizing their underwriting policy. These two claims are based on the following considerations:

- Agency, Markov blankets, and minimizing surprise:
 Agents use the available information about their external world to continuously update their internal representation of the external world (perceptual inference). This internal representation is used to derive an optimal ‘strategy’ for acting on the external world (active inference). Biological agents apply active inference to minimize surprise (Parr *et al.*, 2022), and this principle also applies to corporate agents.
 The various activities of a reinsurer, particularly the rating and underwriting processes, can be represented by a hierarchy of Markov blankets (a description of Markov blankets can be found in Murphy, 2023). At the top of this hierarchy, the reinsurer acts as a corporate agent. It receives



Figure 2. *Integrated rating framework. The exposure-based generative model attempts to emulate the generative process. The framework uses the market experience Ω_o , the exposure profiles Σ_o , the exposure models Ψ_o , and the policy conditions \S_o to generate samples with simulated claims for $o \in \{-T, \dots, -1\}$ (with $T = 10$). The respective observed claims reported by the cedent for $o \in \{-T, \dots, -1\}$ and Bayesian inference are used to calibrate the generative model. The calibrated exposure model is projected to the future cover period $o = +1$ to derive the statistics for rating the reinsurance contracts.*

various kinds of information from the external world and acts on the external world to pursue its corporate goals. At the bottom of the hierarchy, the pricing actuaries and the underwriters assess the individual contract’s risk and the expected loss costs by combining the submission data provided by cedents with market information and expertise. This assessment is combined with other criteria in the decision process on higher levels within the Markov blanket hierarchy.

- The role of rating in reinsurance:

The net premiums and the loss costs drive the bottom-line result of a reinsurer. A bias of a few percentage points in the loss-ratio estimates on an aggregate portfolio level can lead to a substantial systematic adverse development of reserves and erase the bottom-line profit anticipated when assuming the contracts. Reinsurers are, therefore, using various sources of information to model the potential future claims activity of the assumed contracts. This assessment is used for risk selection, setting terms and conditions, initial reserving, investment decisions, and corporate risk management.

The integrated rating framework introduced in this article attempts to minimize the bias and uncertainty in the loss-cost estimates on a contract level. This minimization is critical to a reinsurer’s attempt to implement an optimal underwriting policy.

- An integrated rating approach is superior to a standard experience-based rating:

An exposure-based generative model attempts to replicate the relevant features of the generative processes in the external world by considering the available information coherently. This includes past and anticipated future developments in the market and in the cedent’s portfolio. The prior (default) generative model is obtained by combining the idiosyncratic exposure of a cedent with loss models derived from market data. Bayesian inference is applied to the cedent’s claims experience and prior assumptions to reflect the cedent’s idiosyncratic features in the calibrated posterior generative model.

Hence, the calibrated generative model coherently reflects all available information about the respective market, the underlying client portfolio, and the client’s claims experience. Projecting the claims experience to the future cover period and developing the open claims (as done in

Table 1. Comparison of the integrated rating framework with the standard framework.

Evaluation criteria	Standard rating framework			Integrated framework
	Experience	Exposure	Blending	
Adequately reflecting all available information	+	+	++	+++
Reflecting the cedents experience	++	-	+	+++
Most recent experience in long-tail business	+	-	+	++
Reflecting the broader market experience	+	++	+	++
Reflecting changing insurance conditions	-	+	+	++
Adequate application of reinsurance conditions	++	+	++	+++
Adequate for rating higher layers	-	+	+	++
Reflecting changes in the portfolio composition	-	++	+	+++
Reflecting the future portfolio composition	+	++	+	++
Reflecting general trends	+	-	+	++
Decomposition into frequency, severity, and developments	++	+	++	+++
Providing uncertainty estimates	+	+	+	++
Supporting scenario analysis	-	+	+	++
Applicable for rating and risk management	-	++	-	++
Model uncertainty	high	high	high	medium
Initial effort	n/a	n/a	n/a	high
Update effort	medium	medium	medium	medium

experience-based rating) can thus not provide extra information about the claims process to be expected in the future cover period. An experience-based model lacks the extra information reflected in the generative model. Trending and developing past claims is not equivalent to the explicit approach implemented within generative models; therefore, standard experience-based rating approaches are inferior to an integrated rating framework.

- An integrated rating approach is superior to a standard exposure-based rating: Standard exposure-based rating approaches combine the most recent cedent exposure (projected to the cover period) with market loss models, but they do not reflect the cedent's claims experience. Standard exposure-based rating relies on the prior generative model, which is inferior to a calibrated posterior generative model.
- An integrated rating approach is superior to standard blending approaches: One blending approach (Desmedt *et al.*, 2012) uses the burning-cost rate of a reference layer (as derived with the help of an experience-based rating model) to calibrate the frequencies within the exposure-based model. An alternative approach applies a credibility-based blending of the experience-based and exposure-based rates. No blending approach can, however, overcome the inherent shortcomings of the underlying experience-based and exposure-based rating approaches; therefore, the blending of rates is inferior to an integrated rating framework.
- Caveat: The generic integration framework can be applied to any exposure-based rating model, provided that the probability distributions for the model features to be calibrated can be extracted (either analytically, via numerical approximations, or simulation) and that the derived calibration parameters can be fed back into the generative model. The integrated approach is only as strong as its weakest link, that is, as the weakest module within the exposure-based generative model. Applying the calibration framework to an exposure-based generative model composed of frequency and severity modules might, for

example, be appropriate for short-tail business. However, such a simplified model cannot adequately reflect the impact of the development of long-tail claims. An integrated rating approach relying on an oversimplistic generative model is superior to the respective exposure-based rating model, but it might be inferior to an experience-based rating approach that reflects the claims development.

1.2. Plan of the article

This article discusses the coherent integration of the available information in reinsurance rating using generative models. Section 2 describes the standard rating approaches applied in reinsurance. Section 3 introduces the concept of *generative models* and discusses how they can be used in an integrated framework. Section 4 covers appropriate statistics, while Section 5 discusses the calibration of the generative models with Bayesian inference. Finally, Section 6 analyzes the framework in more detail using ‘toy models’. Technical details are provided in the [Online Supplementary Material](#).

2. Experience and exposure rating in reinsurance

One major challenge for reinsurance companies is the lack of direct access to detailed information related to the covered portfolios, that is, they have no direct access to large datasets that relate the claim characteristics (frequency, severity, and development patterns) to the respective exposure information (objects at risk, policyholders, and policy terms), and the risk factors (covered perils and threads). Reinsurers thus rely on market information $\mathcal{D}^{\text{market}}$ provided by specialized organizations, for example, Verisk (Verisk, 2024), and on idiosyncratic submission data provided by their cedents. For nonproportional contracts, the submission data comprise the cedent’s claims experience $\mathcal{D}^{\text{claims}}$ in excess of a threshold and the cedent’s exposure profiles $\mathcal{D}^{\text{exposure}}$.

Reinsurers combine the various data sources when creating models for the claims activity expected to occur in the cedent portfolio during the future cover period; these models are used to derive the statistics required to rate the reinsurance contracts $\mathcal{D}^{\text{cover}}$. Reinsurers use an extensive suite of proprietary and vendor rating tools that are specifically designed for different markets, different lines of business, different perils, and distinct kinds of contracts. The rating is, however, either based on the claims experience of a cedent (experience rating), on the exposure of a cedent combined with market models (exposure rating), or on a combination of the two approaches. The ‘standard rating framework’ applied by a reinsurer is depicted in Figure 1, and a comparison of the various approaches is provided in Table 1. The integrated approach attempts to overcome the drawbacks of standard rating approaches by coherently combining all available information. The benefits come, however, at the cost of higher procedural and computational efforts.

A comprehensive description of rating models is beyond the scope of this article. A variety of approaches can, however, be found in books (Daykin *et al.*, 1994; Albrecher, 2017; Klugman *et al.*, 2019; Mildenhall, 2022), articles (Panjer, 1981; Sundt and Jewell, 1981; Schnieper, 1991; Bernegger, 1997; Riegel, 2010; Desmedt *et al.*, 2012; Zhang *et al.*, 2012), conference papers (White and Mrazek, 2004; Mata and Verheyen, 2005; Bernegger, 2012; Clark, 2014; Huang, 2014; Devlin, 2018), lecture notes (Antal, 2009), and technical publications by corporations and professional organizations (Guggisberg, 2004; Billeter and Salghetti-Drioli, 2016; Verisk, 2024; CAS, 2024; CARE, 2024).

2.1. Experience rating

Experience rating in reinsurance uses the claims experience $\mathcal{D}_t^{\text{claims}}$ (as reported by the cedent company ι) to infer a stochastic model $\Pr(\mathbf{x} \mid \mathcal{D}_t^{\text{claims}})$ for the expected future claims activity \mathbf{x} within the underlying portfolio. This model is used to evaluate the relevant statistics $\mathbb{E}[\mathbf{h}(\mathbf{x} \mid \mathcal{D}_t^{\text{cover}}) \mid \mathcal{D}_t^{\text{claims}}]$ for the cover $\mathcal{D}_t^{\text{cover}}$ provided to the cedent company ι during the future contractual period. The operator $\mathbf{h}: \mathbf{x}^{\text{insured}} \mapsto \mathbf{x}^{\text{cover}}$

projects the insured claim objects $\mathbf{x}^{\text{insured}}$ to the covered claim objects $\mathbf{x}^{\text{cover}}$ by applying the terms defined in $\mathcal{D}_t^{\text{cover}}$.

The experience rating process (depicted in the lower part of Figure 1) involves several steps, among which the *projection* and the *development* of past claims are most critical. The claims experience must be *projected* ('trended') to the future cover period, where the 'as-if' claims are used for rating the reinsurance contracts. This projection must reflect changes in the cedents portfolio and inflationary effects that impact the claims' frequency, severity, and development patterns. The development of the claims reported for past occurrence years is known up to the submission date t^{sub} . This development comprises the temporal evolution of the cumulative paid amount $P(t) := \sum_i p_i \cdot \mathbb{1}_{t_i \leq t}$ and the incurred loss amount $I(t) := \mathbb{E}[U | \mathcal{D}_t] = P(t) + O(t)$, where \mathcal{D}_t is all past information prior to time t , p_i is the amount paid at time t_i , and $O(t) := \mathbb{E}[\sum_i p_i \cdot \mathbb{1}_{t_i > t}]$ is the best estimate at time t for the outstanding payments. The ultimate loss amount $U := P(t) = I(t)$ for any $t \geq t^{\text{clos}}$ is thus known for closed claims and estimated for open claims. The claims triangle is completed by *developing* $P(t)$ and $I(t)$ either per individual claim or on a portfolio level (Schnieper, 1991).

The primary strength of experience rating is that it is based on the cedent's specific claims experience and that standardized actuarial approaches are applied across markets, lines of business, and products. There are, however, several weaknesses; some claims have not yet been reported; claims reported for old years are well developed, but they might no longer be representative; claims reported for the most recent years are representative, but they are not yet developed; the count of trended claims with an incurred loss amount exceeding the reporting threshold is often small; and experience rating cannot reflect significant portfolio changes or new risks covered by the underlying insurance policies.

2.2. Exposure rating

Reinsurers use an extensive suite of exposure-based rating models. These models combine the risk profiles provided by the cedent with market statistics, expert knowledge, and, depending on the specific nature of the model, other modeling features, for example, scientific models emulating natural perils. There are two kinds of exposure models: the first combines the exposure with probability distributions (exposure curves), and the second relies on *generative models* that simulate significant numbers of events applied to the exposure.

Models of the first kind are computationally cheap as they do not require Monte Carlo simulations. The models can, however, also be implemented as generative models, and there is no need to distinguish between the two kinds in the context of this article.

Models of the second kind attempt to replicate the corresponding *generative processes* in the world, for example, natural catastrophes, accidents, and defaulting corporations. These models can be used to rate individual reinsurance contracts (primary scope) and for corporate risk management (secondary scope). The primary scope is accomplished by exposing a contract to all relevant events and evaluating the respective loss amounts for the contract. The secondary scope is accomplished by aggregating the loss amounts of individual events across all contracts.

The strength of exposure-based rating models is that they reflect a cedent's (projected) future portfolio; they are aligned with the broader market experience; the same model is used for rating individual contracts and for corporate risk management. The weaknesses of exposure rating are that: a specific model is required for each business niche; calibrating the models is challenging; the models do not reflect the claims experience of a cedent; the models might not adequately reflect recent developments; and exposure models of the second kind are computationally expensive.

3. Integrated rating approach

3.1. Ensemble modeling

Ensemble forecast is fundamental in predictive science, that is, in attempts to make predictions about the chaotic world and in understanding and quantifying the uncertainty of such predictions. It is applied

in astronomy, meteorology, engineering, chemistry, biology, medicine, ecology, economics, conflict management, etc. (Palmer, 2022). The methodology is also applied in insurance and reinsurance, particularly to assess risks attributed to natural disasters (see Section 2.2).

3.2. Generic process

The basic idea underlying a fully integrated rating approach is presented in Figure 2. Standard experience-based rating models require that all claims incurred and reported in prior years (including their temporal development till the submission date) are projected to the future cover period. These ‘as-if’ claims are used to derive the required statistics. This projection (see Figure 1) is accomplished by scaling frequency, severity, and payments by respective exposure metrics and trend factors. This nontrivial step is completely avoided in the integrated rating approach, where the experience is related to the respective past exposure-based model. This alternative approach thus permits us to adequately consider all known (and presupposed) developments in the underlying portfolio and the broader market environment. Such developments can encompass portfolio size and composition, terms of insurance policies, reporting thresholds, court practices, claims inflation, technological developments, reserving practices, etc.

Hence, the approach requires that such an exposure-based model be available for each past year for which the claims experience is considered in the rating process. The aim is to use the exposure-based generative model to produce all the statistics needed to rate the reinsurance contracts that will be enforced in the future cover period. This process is briefly described in Section 2.2 and shown in the upper right of Figure 2. However, the generative model must be adequately calibrated before being used for this rating purpose. This calibration task is performed by running the generative model with the exposure in each prior year, applying the respective policy and contract terms, and comparing the simulated claims with the respective observed claims.

Relating the past claims experience to the respective past exposure-based loss model instead of projecting it to the future cover period is the first key feature of the integrated rating approach.

3.3. Claim generator

The purpose of *generative models* is to replicate real-world *generative processes* as closely as required for the respective scope. In the context of reinsurance rating, a loss model needs to generate all relevant features of the observed claims provided to the reinsurer via the cedent’s submission data. The format and content of the submission data vary by market, line of business, product, and cedent. There is, however, a generic claim format that captures the key features of nonlife insurance claims and the variations can be mapped to this format. The following features characterize the reported insurance claims:

- A tuple containing static claim features (administrative attributes), for example, policyholder, inception and expiration dates of the policy, policy ID, claim administrator, claim ID, event ID, status of the claim, etc.
- A tuple containing categorical attributes describing the nature of the claim, for example, line of business, geo-encoding, type of cover, risk category, kind of event causing the claim, etc.
- Critical dates like the occurrence (accident) date, the reporting date, and – depending on the status – the closure (settlement) date.
- Incurred pattern, that is, the development of the best estimate for the ultimate amount to be paid by the cedent company.
- Payment pattern, that is, the development of the aggregate payments (including the breakdown into components).

The critical dates and the development patterns (reported, incurred, and paid) matter from the perspective of quantitative analysis and model calibration. The reporting pattern is associated with *true* IBNR (incurred but not reported) and the incurred pattern with IBNER (incurred but not enough reserved). In contrast to IBNR, which is adverse, IBNER can either be adverse, adequate, or favorable, depending on whether the estimate of the incurred amount turns out to be too low, correct, or too high. The underlying payment pattern affects the amount and the timing of the payments of the loss portion covered by the reinsurer. The three development patterns depend on multiple factors (related to the policyholder, the cedent, and other agents), and they must be modeled and calibrated separately.

4. Reduced statistics

4.1. Challenge

The primary scope of a product-specific generative model used in exposure-based rating is to adequately emulate the relevant features of the generative process and provide all claim features that might affect the rating. This implies the generation of significant samples that populate the relevant regions in the sample space Ω , and the simulated claims must entail all rating-relevant features. The model must be calibrated, which is accomplished by running the generative model retrospectively. The simulated claims sets generated for the past observation period of length T (denoted by the occurrence years $o \in \{-T, \dots, -1\}$ defined relative to the submission year $o = 0$) are compared to the observed claims for this period.

The calibration tuple $\mathbf{C} := (\mathbf{c}_1, \dots, \mathbf{c}_M)$ containing the calibration parameters \mathbf{c}_ℓ is evaluated outside of the generative model and fed back into the model. The generative model can, therefore, be treated as an independent module interacting with the generic calibration framework. The iterative calibration procedure starts with default calibration parameters $\mathbf{c}_\ell^{(0)} = \mathbf{0}$ for all features ℓ to be calibrated. The generative model accepts the calibration tuple $\mathbf{C}^{(s)}$ and the simulation count $K^{(s+1)}$ as input, and it applies the calibration parameters $\mathbf{c}_\ell^{(s)}$ to the respective modules ℓ . The generative model provides $K^{(s+1)}$ sets containing simulated claims for the observation period T as an output. The calibration framework compares the simulated sets with the observed set and derives an updated calibration tuple $\mathbf{C}^{(s+1)}$ (see Section 5.2).

The challenge is to derive adequate statistics that permit the evaluation of the calibration parameters \mathbf{c}_ℓ . This challenge is addressed by introducing a generic claim object \mathbf{L} that covers the rating needs (see Figure 4):

$$\mathbf{L} := \left(\mathbf{D}, s, t^{occ}, t^{rep}, [t^{clo}], \{(I_i, t_i)\}_{1 \leq i \leq N^I}, \{(P_j, t_j)\}_{1 \leq j \leq N^P} \right) \tag{4.1}$$

where the tuple \mathbf{D} is a descriptor comprising the static claim features (see Section 3.3), s the status of the claim ('open', 'closed', or 're-opened'), t^{occ} the occurrence (accident) date, t^{rep} the reporting date, t^{clo} the closure date (in the case $s = \text{'closed'}$), N^I the count of incurred adjustments, I_i the incurred amount as estimated at time t_i , N^P the count of payments, and P_j the cumulative paid amount at time t_j^+ . The claim object \mathbf{L} is thus composed of multiple partially interdependent random variables.

Some cedents evaluate the development patterns at discrete points in time (e.g., on an annual basis). The observation period T is then decomposed into N^T discrete subperiods of length $\Delta t = T/N^T$ and the representation of the claims along the timeline is aligned with the time grid $\{t_0, \dots, t_{N^T}\}$. The patterns $I(t)$ and $P(t)$ are approximated with $\{(t_k, P_k, I_k)\}_{0 \leq k \leq N^T}$ where $t_k = t_0 + \Delta t \cdot k$, $I_k = I(t_k^+)$, and $P_k = P(t_k^+)$.

The simulated claims and the observed claims are represented by elements ω within the high-dimensional sample space Ω , and the generative model is calibrated if the observed claims are 'close' to the simulated claims, that is, if they are located within the highest density region(s) $HDR := \{H_i\} \subset \Omega$ defined as follows (see also Lee, 2012):

1. The sub-spaces H_i and H_j are disjoint: $i \neq j \rightarrow H_i \cap H_j = \{\}$.
2. The density within the HDR equals or exceeds the threshold density p_0 , that is, we have $p(\omega) \geq p_0$ for all $\omega \in HDR$ and $p(\omega) < p_0$ for all $\omega \in \Omega \setminus HDR$.

- The threshold density p_0 is determined by the condition $\Pr(\omega \in HDR \mid \omega \in \Omega) \stackrel{!}{=} P_{HDR}$, for example, $P_{HDR} = 95\%$.

Identifying the *HDR* is easy in the case of low-dimensional, uni-modal parametric probability distributions and extremely difficult in the case of high-dimensional, multi-modal, nonparametric probability distributions. The calibration criterion is formalized with the following sets containing the observed claims $\mathbf{L}_{o,i}^{obs}$ and the simulated claims $\mathbf{L}_{o,i,k}^{sim}$:

$$\begin{aligned} \mathcal{S}_{k=0} &:= \mathcal{S}^{obs} &:= \left\{ \left\{ \mathbf{L}_{o,i}^{obs} \right\}_{1 \leq i \leq N_o^{rep}} \right\}_{-T \leq o \leq -1} \\ \mathcal{S}_{k>0} &:= \mathcal{S}_k^{sim} &:= \left\{ \left\{ \mathbf{L}_{o,i,k}^{sim} \right\}_{1 \leq i \leq N_{o,k}^{rep}} \right\}_{-T \leq o \leq -1} \quad \text{for } k \in \{1, \dots, K\} \end{aligned} \tag{4.2}$$

where N_o^{rep} is the count of reported claims for the occurrence year o , K is the number of simulations runs, $N_{o,k}^{rep}$ is the count of simulated claims with $t_{o,i,k}^{rep} \leq t^{sub}$ generated in the k^{th} run for the occurrence year o , and t^{sub} is the submission date. The generative model is adequately calibrated if the observed set \mathcal{S}_0 cannot be identified as an outlier within the set of sets $\{\mathcal{S}_k\}_{0 \leq k \leq K}$. The challenge is defining appropriate metrics that permit quantifying this criterion and deriving the calibration tuple \mathbf{C} in the case of a mismatch.

4.2. Roadmap for a solution

The n -dimensional sample space Ω (with $n \approx 2 \cdot N^T$ in the grid representation) must be decomposed, that is, the high-dimensional tuples describing the patterns $I(t)$ and $P(t)$ must be projected to low-dimensional objects. This is accomplished by identifying the relevant features of the claims, that is, the high-dimensional claim objects \mathbf{L} (4.1) are projected to low-dimensional tuples \mathbf{Z} :

$$\begin{aligned} \zeta &: \mathbf{L} \mapsto \mathbf{Z} = (Z_1, \dots, Z_M) \\ \zeta_\ell &: \mathbf{L} \mapsto Z_\ell \end{aligned}$$

This is done with the calibration scope in mind, and this objective leads to the constraint $\dim(\mathbf{Z}) \stackrel{!}{=} \dim(\mathbf{C}) = M \ll \dim(\mathbf{L}) = n$, where the calibration tuple $\mathbf{C} = (c_1, \dots, c_M)$ contains the calibration parameters c_ℓ to be applied to the features ℓ within the generative model.

Remark 4.1. A calibration parameter c_ℓ (a scalar or a low-dimensional vector) is assigned to each feature ℓ of the generative model, but the actual calibration might be confined to a subset of these features.

Ideally, \mathbf{Z} is defined on an orthogonal basis while preserving the key features of the claim \mathbf{L} . The component Z_ℓ then strongly responds to the corresponding calibration parameter c_ℓ , and it is insensitive to all other calibration parameters (i.e., the non-diagonal elements in the Jacobi matrix $\partial Z_i(\mathbf{C})/\partial c_j$ are insignificant compared to the diagonal elements).

The simulation process implemented in the generative model aims to replicate the generative process leading to the observed claims. The projection ζ is, therefore, applied to the observed claims $\mathbf{L}_{o,j}^{obs}$ and the simulated claims $\mathbf{L}_{o,i,k}^{sim}$. The sets containing the observed and the simulated reduced claim components $Z_{\ell,\kappa} := \zeta_\ell(\mathbf{L}_\kappa)$ are used to calibrate the feature ℓ within the generative model (see Section 5.2). Any distortion in the observed variables is replicated in the simulated variables, and there is no need for the marginal random variables $Z_{\ell,\kappa}$ to be iid.

4.3. Tuple with reduced random variables

The calibration procedure is used for an overall adjustment of the frequencies, the severities, the lags, and the pattern counts. The patterns are thus characterized by respective representative statistics as captured

in the following reduced claim representation $\mathbf{Z}_\kappa = \zeta(\mathbf{L}_\kappa)$:

$$\mathbf{Z}_\kappa := (I_\kappa^{occ}, (N_\kappa^{inc}, N_\kappa^{rep}, \tau_\kappa^{rep}), (N_\kappa^{clo}, \tau_\kappa^{clo}), (N_\kappa^I, I_\kappa^*, \tau_\kappa^I), (N_\kappa^P, P_\kappa^*, \tau_\kappa^P)) \tag{4.3}$$

where the index ‘ $\kappa \equiv o, i$ ’ indicates the i^{th} claim in occurrence year o in the case of observed claims and ‘ $\kappa \equiv o, i, k$ ’ the i^{th} claim in occurrence year o in the k^{th} simulation run in the case of simulated claims.

All variables $Z_{\ell,o,i[k]}$ are defined in Table 2 on a *claim* level, aggregated into $Z_{\ell,o[k]}$ on an *annual* level, and into $Z_{\ell[k]}$ on a *period* level (Table 2 comprises the extra variable $N_{o[k]}^0$ indicating claim-free years).

Remark 4.2. The structure of the claim objects \mathbf{L}_κ and the constraints determine the structure of the reduced random variables \mathbf{Z}_κ . The reduced variables introduced in (4.3) and defined in Table 2 capture the key features of a claim. Some variables are, however, interdependent (see Section 5.3.1). A subset is thus chosen for calibration, and the remaining variables are used as test variables.

The relationship between \mathbf{Z}_κ and \mathbf{L}_κ , the temporal evolution of the incurred amount $I_\kappa(t)$, the cumulative paid amount $P_\kappa(t)$, and the two reduced random variables $\tau_\kappa^I(t)$ and $\tau_\kappa^P(t)$ are shown in Figures 3 and 4. The incurred lag $\tau_\kappa^I(t)$ evaluated at time $t > t_\kappa^{rep}$ can either be *larger*, *equal*, or *smaller* than the reporting lag τ_κ^{rep} . The three cases correspond to an *adverse*, *adequate*, and *favorable* IBNER development, respectively. The paid lag $\tau_\kappa^P(t)$ evaluated at time $t > t_\kappa^{rep}$ is always *larger* than the reporting lag τ_κ^{rep} (losses are paid post claims being made).

Replacing the detailed patterns of a claim by a tuple containing reduced and weakly coupled variables is the second key feature of the integrated rating approach.

5. Bayesian inference and model calibration

5.1. Bayesian inference

The Bayesian interpretation combines prior knowledge with observed data when estimating the parameters of parametric probability distributions. It distinguishes between the *intrinsic uncertainty* stemming from the stochastic process and the *epistemic uncertainty* stemming from the uncertainty in the parameter estimates (see Figures 6 and 7). This so-called *Bayesian inference* is particularly convenient when applied to conjugate priors (Fink, 1997; Gelman, 2013; Klugman *et al.*, 2019; Parr *et al.*, 2022; Palmer, 2022; Murphy, 2023). Depending on the kind of distributions, a closed form might be available for the *posterior predictive distribution* (Fink, 1997). However, Bayesian inference is also applicable in situations that deviate from the ideal case, that is, when the observed random variables are not iid or when dealing with nonparametric probability distributions (Parr *et al.*, 2022; Palmer, 2022; Murphy, 2023).

The observed random variables \mathbf{L}_κ^{obs} are complex objects emanating from real-world generative processes that evolve over time. They cannot be used directly for evaluating the calibration tuple \mathbf{C} to be applied to the exposure-based generative model; the model is, therefore, calibrated with reduced random variables \mathbf{Z}_κ^{obs} (see Section 4.3 and Table 2).

Applying Bayesian inference requires that the ‘conditional distributions’ and the respective ‘prior distributions’ are represented (or approximated) by parametric distributions. The parameters to be calibrated are identified, the respective prior distributions are specified, and the posterior estimates are fed back into the generative model. This is accomplished with scaling factors applied to corresponding parameters that determine the frequency, the severity, and the development patterns. The incurred lag is, however, either positive or negative; thus, a ‘translate’ parameter is used to calibrate the incurred pattern.

A Monte Carlo simulation generates claim objects \mathbf{L}_κ^{sim} that replicate all relevant features of the observed claims \mathbf{L}_κ^{obs} . The simulated claims and the observed claims are reduced and denoted $\mathbf{X}_\kappa := \zeta(\mathbf{L}_\kappa^{sim})$ and $\mathbf{Y}_\kappa := \zeta(\mathbf{L}_\kappa^{obs})$, respectively. Data slices $\mathcal{X}_{\ell,o}$ and $\mathcal{Y}_{\ell,o}$ containing the simulated random

Table 2. Comprehensive list with reduced random variables \mathbf{Z} on a claim level, an annual level, and a period level.

Claim feature ℓ	claim level: $Z_{\ell,o,i[k]}$	annual level: $Z_{\ell,o[k]}$	mean: $\hat{Z}_{\ell,o[k]}$	period level: $Z_{\ell[k]}$	mean: $\hat{Z}_{\ell[k]}$
Claim-free		$N_o^0 := \mathbb{1}_{N_o^{rep}=0}$		$N^0 := \sum_o N_o^0$	$\hat{\lambda}^0 := \frac{N^0}{T}$
Occurrence	$t_{o,i}^{occ}$	$o := \lfloor t_{o,i}^{occ} - t^{sub} \rfloor - 1$			
Incurred count	$N_{o,i}^{inc} := 1$	$N_o^{inc} := \sum_i N_{o,i}^{inc}$		$N^{inc} := \sum_o N_o^{inc}$	$\hat{\lambda}^{inc} := \frac{N^{inc}}{T}$
Reported count	$N_{o,i}^{rep} := \mathbb{1}_{t_{o,i}^{rep} \leq t^{sub}}$	$N_o^{rep} := \sum_i N_{o,i}^{rep}$		$N^{rep} := \sum_o N_o^{rep}$	$\hat{\lambda}^{rep} := \frac{N^{rep}}{T}$
Reporting lag	$\tau_{o,i}^{rep} := t_{o,i}^{rep} - t_{o,i}^{occ}$	$\tau_o^{rep} := \sum_i \tau_{o,i}^{rep} \cdot \frac{I_{o,i}^*}{I_o^*}$		$\tau^{rep} := \sum_o \tau_o^{rep} \cdot \frac{I_o^*}{I^*}$	
Closed count	$N_{o,i}^{clo} := \mathbb{1}_{t_{o,i}^{clo} \leq t^{sub}}$	$N_o^{clo} := \sum_i N_{o,i}^{clo}$		$N^{clo} := \sum_o N_o^{clo}$	$\hat{\lambda}^{clo} := \frac{N^{clo}}{T}$
Closure lag	$\tau_{o,i}^{clo} := t_{o,i}^{clo} - t_{o,i}^{occ}$	$\tau_o^{clo} := \sum_i \tau_{o,i}^{clo} \cdot \frac{I_{o,i}^*}{I_o^*}$		$\tau^{clo} := \sum_o \tau_o^{clo} \cdot \frac{I_o^*}{I^*}$	
Incurred adj.	$N_{o,i}^l$	$N_o^l := \sum_i N_{o,i}^l$	$\hat{n}_o^l := \frac{N_o^l}{N_o^{rep}}$	$N^l := \sum_o N_o^l$	$\hat{n}^l := \frac{N^l}{N^{rep}}$
Incurred	$I_{o,i}^* := I_{o,i}(t^{sub})$	$I_o^* := \sum_i I_{o,i}^*$	$\hat{I}_o^* := \frac{I_o^*}{N_o^{rep}}$	$I^* := \sum_o I_o^*$	$\hat{I}^* := \frac{I^*}{N^{rep}}$
Incurred lag	$\tau_{o,i}^l := \int_{t_{o,i}^{occ}}^{t^{sub}} \left(1 - \frac{I_{o,i}(t)}{I_{o,i}^*}\right) dt$	$\tau_o^l := \sum_i \tau_{o,i}^l \cdot \frac{I_{o,i}^*}{I_o^*}$		$\tau^l := \sum_o \tau_o^l \cdot \frac{I_o^*}{I^*}$	
Payments	$N_{o,i}^p$	$N_o^p := \sum_i N_{o,i}^p$	$\hat{n}_o^p := \frac{N_o^p}{N_o^{rep}}$	$N^p := \sum_o N_o^p$	$\hat{n}^p := \frac{N^p}{N^{rep}}$
Paid	$P_{o,i}^* := P_{o,i}(t^{sub})$	$P_o^* := \sum_i P_{o,i}^*$	$\hat{P}_o^* := \frac{P_o^*}{N_o^{rep}}$	$P^* := \sum_o P_o^*$	$\hat{P}^* := \frac{P^*}{N^{rep}}$
Paid lag	$\tau_{o,i}^p := \int_{t_{o,i}^{occ}}^{t^{sub}} \left(1 - \frac{P_{o,i}(t)}{P_{o,i}^*}\right) dt$	$\tau_o^p := \sum_i \tau_{o,i}^p \cdot \frac{I_{o,i}^*}{I_o^*}$		$\tau^p := \sum_o \tau_o^p \cdot \frac{I_o^*}{I^*}$	

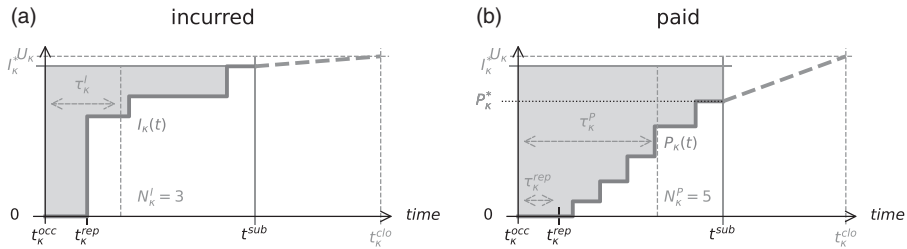


Figure 3. Claim representation and reduced variables ($\kappa \equiv o, i$ or $\kappa \equiv o, i, k$). The available information, that is, the occurrence date t_k^{occ} , the reporting date t_k^{rep} , the submission date t^{sub} , the incurred pattern $I_\kappa(t)$ (a), and the paid pattern $P_\kappa(t)$ (b) are shown in black. The derived reduced variables, that is, the reporting lag τ_κ^{rep} , the incurred amount I_κ^* , the incurred lag τ_κ^I , the count of incurred adjustments N_κ^I , the cumulative paid amount P_κ^* , the paid lag τ_κ^P , and the count of payments N_κ^P are shown in gray (see definitions in Table 2). The temporal development is taken from the ‘increase’ case shown in Figure 4 by evaluating the patterns at time $t = 75$ [months].

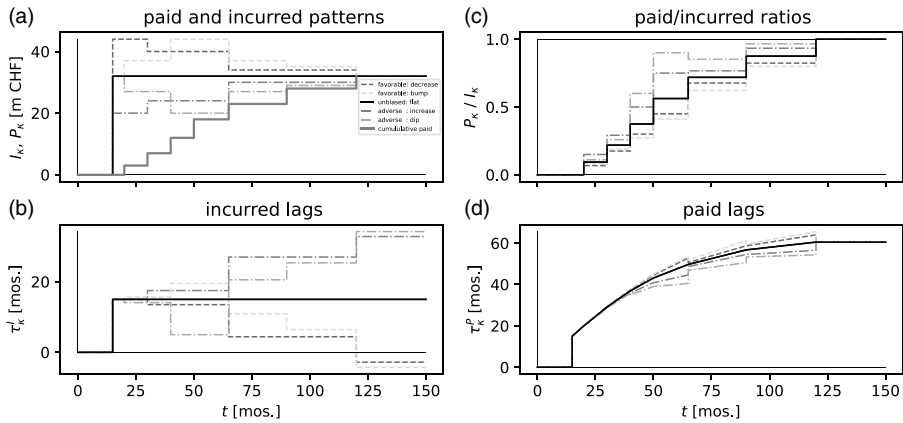


Figure 4. Patterns and lags as defined in Table 2 and Figure 3 (t in months [mos.] since t_k^{occ} and $\kappa \equiv o, i$ or $\kappa \equiv o, i, k$). (a) The solid gray curve shows the temporal evolution of the cumulative paid amount $P_\kappa(t)$. The other curves depict five examples for the development of the incurred amount $I_\kappa(t)$. Two adverse cases are represented by ‘dash-dotted’ lines, the adequate case by a ‘solid’ line, and two favorable cases by ‘dashed’ lines. (b) Temporal evolution of the respective incurred lags $\tau_\kappa^I(t)$. The incurred lags are initially equal to the reporting lag $\tau_\kappa^{rep} = 15$ [months], and they subsequently increase in the ‘adverse’ cases, remain stable in the ‘adequate’ case, and decrease in the ‘favorable’ cases. (c) Temporal evolution of the respective paid-over-incurred ratios $P_\kappa(t)/I_\kappa(t)$. (d) Temporal evolution of the paid lags $\tau_\kappa^P(t)$ for the five cases of incurred patterns.

variables $X_{\ell,o,i,k}$ and the observed random variables $Y_{\ell,o,j}$ on a claim level, respectively, are extracted for the feature ℓ and the occurrence year o .

Marginal probability distributions $f_{\ell,o}(x | \mathbf{c}_\ell) := f_\ell(x; \boldsymbol{\theta}_{\ell,o}(\mathbf{c}_\ell))$ are derived by fitting the parametric distributions $f_\ell(x; \boldsymbol{\theta}_{\ell,o}^0)$ to the simulated distributions $g(x | \mathcal{X}_{\ell,o})$ and by linking the parameter tuple $\boldsymbol{\theta}_{\ell,o}$ to the calibration parameter \mathbf{c}_ℓ (see Section 5.2). Applying Bayesian inference permits us to derive posterior distributions $p(\mathbf{c}_\ell)$ and maximum a posteriori (MAP) estimates for the parameters \mathbf{c}_ℓ given the observations $\mathcal{Y}_{\ell,o}$. The calibration is either performed with calibration scalars c_ℓ or with calibration vectors, for example, $\mathbf{c}_\ell := (a_\ell, b_\ell)$ in the case of a linear-trend calibration. Analogous procedures are applied if the calibration is performed on an annual level or a period level.

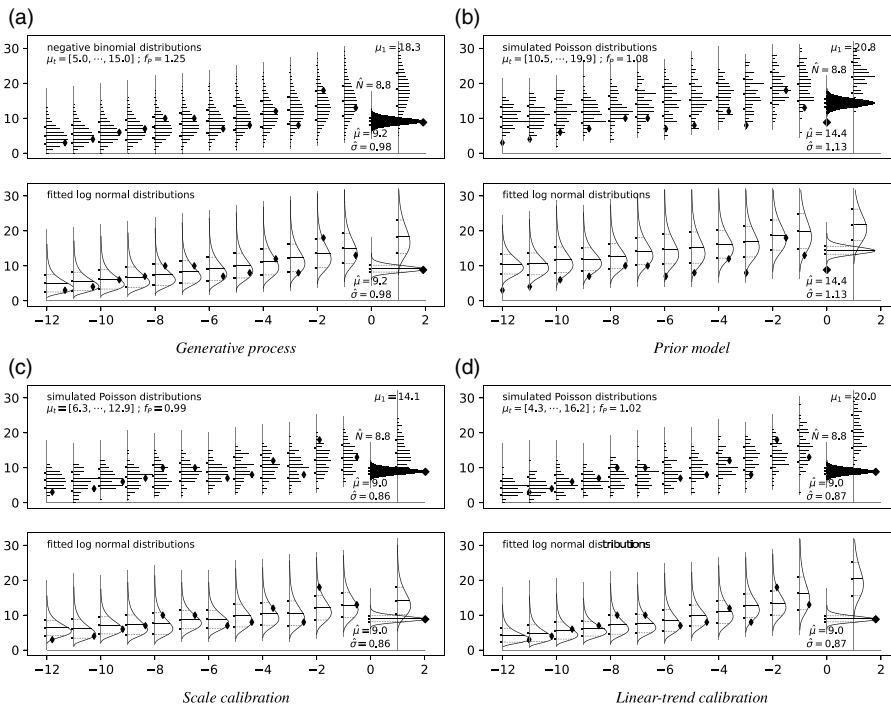


Figure 5. Calibration of the annual ‘observations’ N_o (depicted by diamonds). The pmfs of the models for the occurrence years $o \in \{-T, \dots, -1\}$ (with $T = 12$), the resulting pmf of the average over the period (plotted at $o = 0$), and the projection (plotted at $o = +1$) are shown in the upper part of the four charts. The fitted log N distributions are shown in the lower part, and ticks indicate the means and standard deviations. The distributions shown in (c) and (d) are obtained by rerunning the, respectively, calibrated models.

Assigning one or two calibration parameters to each of the relevant reduced variables and applying Bayesian inference to evaluate MAP estimates for the calibration parameters is the third key feature of the integrated rating approach.

5.2. Generic calibration process

5.2.1. Overview

Bayesian inference presupposes parametric probability distributions for the observed random variables and prior distributions for the calibration parameters. The probability distributions implemented within a generative model (considered an independent module) are not directly accessible by the generic calibration framework. The calibration framework furthermore involves probability distributions for multiple occurrence years o . Hence, the calibration is performed with the help of proxy distributions $f_i(x; \theta_{\ell,o})$ fitted to the simulated probability distributions $g(x | \mathcal{X}_{\ell,o}^c)$ derived from the sets $\mathcal{X}_{\ell,o}^c$ containing the simulated claim-level random variables (respective probability distributions can be derived in an analogous way from annual-level sets $\mathcal{X}_{\ell,o}^a$ or from the period-level sets \mathcal{X}_{ℓ}^p):

$$g(x | \mathcal{X}_{\ell,o}^c) := \Pr(X = x | \mathcal{X}_{\ell,o}^c) = \frac{1}{\sum_{k=1}^K N_{o,k}^{rep}} \cdot \sum_{k=1}^K \sum_{i=1}^{N_{o,k}^{rep}} \mathbb{1}_{x=X_{\ell,o,i,k}} \tag{5.1}$$

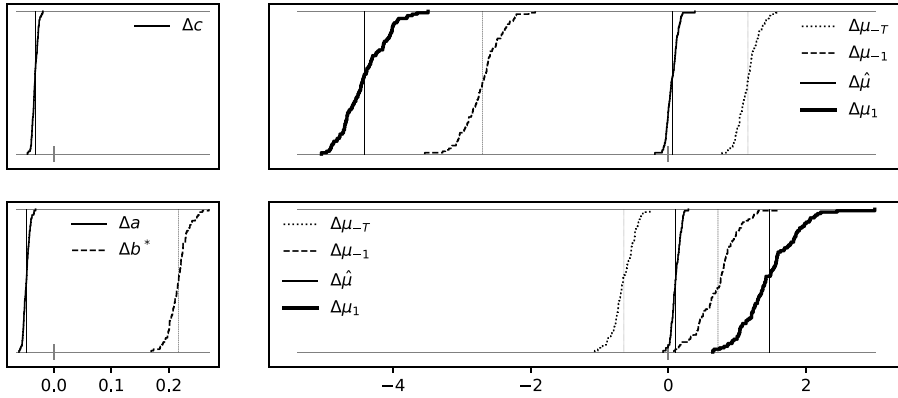


Figure 6. *Conditional calibration statistics. The calibration of the sample model presented in Figure 5 is run 100 times, conditional on the given set of observations $\{N_o\}_{-T \leq o \leq -1}$. The simulated differential cdfs $F(\Delta X)$ (with $\Delta X^{(m)} := X^{(m)} - \mathbb{E}[X]$) are depicted in the four panels. Top: Deviation statistics for the scale-only calibration. Bottom: Deviation statistics for the linear-trend calibration. Left: Deviation statistics for the scale parameters a and c , and the trend parameter $b^* = b \cdot (T - 1)$, respectively. Right: Resulting deviation statistics for the overall mean $\hat{\mu}$, the calibrated means at $o = -T$ and $o = -1$, and the projected mean at $o = +1$.*

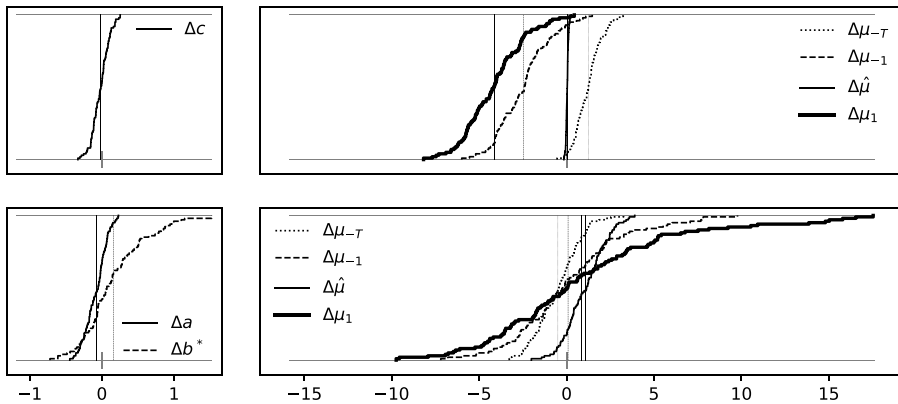


Figure 7. *Unconditional calibration statistics. Deviation statistics as shown in Figure 6 (note the different scales), but without conditioning on $\{N_o\}_{-T \leq o \leq -1}$, that is, a new set of observations $\{N_o^{(m)}\}_{-T \leq o \leq -1}$ is drawn for each simulation run $m \in \{1, \dots, 100\}$.*

Conditional annual probability distributions $f_{\ell,o}(x | \mathbf{c}_\ell) := f_\ell(x; \boldsymbol{\theta}_{\ell,o}(\mathbf{c}_\ell))$ are obtained by linking the parameter tuples $\boldsymbol{\theta}_{\ell,o}$ to the calibration parameter \mathbf{c}_ℓ .

Remark 5.1. The selection of appropriate distribution families $f_\ell()$ for the model features ℓ depends on the variables $Z_{\ell,\kappa} = \zeta_\ell(\mathbf{L}_\kappa)$ being discrete or continuous, the support, the shape, the number of modes, and the tail behavior. The calibration is, however, determined by the HDR (see definition in Section 4.1) of the probability distribution. In the case of uni-modal distributions, it is thus not overly sensitive to the selection if the support and the first few moments are preserved (see Sections 6.3 and 6.6).

The calibration is performed for each feature ℓ with the respective observed random variables $Y_{\ell,o,j}$, the simulated random variables $X_{\ell,o,i,k}$, and the priors $\pi(\mathbf{c}_\ell)$:

1. Annual data slices $\mathcal{Y}_{\ell,o} := \{Y_{\ell,o,j}\}_{1 \leq j \leq N_{o,j}^{obs}}$ and $\mathcal{X}_{\ell,o} := \{X_{\ell,o,i,k}\}_{1 \leq i \leq N_{o,k}^{sim}, 1 \leq k \leq K}$ containing the observed reduced components $Y_{\ell,o,j} = \zeta_{\ell}(\mathbf{L}_{o,j}^{obs})$ and the simulated reduced components $X_{\ell,o,i,k} = \zeta_{\ell}(\mathbf{L}_{o,i,k}^{sim})$, respectively, are extracted for the feature ℓ .
2. A parametric probability distribution family $f_{\ell}(\cdot)$ is used to specify conditional annual probability distributions $f_{\ell,o}(x | \mathbf{c}_{\ell}) := f_{\ell}(x; \boldsymbol{\theta}_{\ell,o}(\mathbf{c}_{\ell}))$.
3. Annual prior parameters $\boldsymbol{\theta}_{\ell,o}^0$ are obtained by fitting the distributions $f_{\ell}(x; \boldsymbol{\theta}_{\ell,o}^0)$ to the simulated marginal distributions $g(x | \mathcal{X}_{\ell,o})$ (5.1).

Bayesian inference applied to the set of observations $\mathcal{Y}_{\ell} := \{\mathcal{Y}_{\ell,o}\}_{-T \leq o \leq -1}$ is used to update the prior distribution $\pi(\mathbf{c}_{\ell})$ for the calibration parameter \mathbf{c}_{ℓ} , that is, to derive the conditional posterior distribution $p(\mathbf{c}_{\ell} | \mathcal{Y}_{\ell})$. This is accomplished by postulating $\mathbb{E}_{\pi}[\mathbf{c}_{\ell}] \stackrel{\dagger}{=} \mathbf{0}$ and linking the calibration parameter \mathbf{c}_{ℓ} to the parameters $\boldsymbol{\theta}_{\ell,o}$ with the condition $\boldsymbol{\theta}_{\ell,o}(\mathbf{c}_{\ell} = \mathbf{0}) \stackrel{\dagger}{=} \boldsymbol{\theta}_{\ell,o}^0$ for all o (see Section 5.2.4). The tuple $\boldsymbol{\theta}_{\ell,o} = (\vartheta_{\ell,o}, [\varphi_{\ell,o}, \varsigma_{\ell,o}, \varrho_{\ell,o}, \dots])$ is assumed to comprise only one parameter $\vartheta_{\ell,o}$ that is linked to the calibration parameter \mathbf{c}_{ℓ} , that is, $\boldsymbol{\theta}_{\ell,o}(\mathbf{c}_{\ell}) = (\vartheta_{\ell,o}(\mathbf{c}_{\ell}), [\varphi_{\ell,o}^0, \varsigma_{\ell,o}^0, \varrho_{\ell,o}^0, \dots])$ with $\vartheta_{\ell,o}(\mathbf{c}_{\ell} = \mathbf{0}) \stackrel{\dagger}{=} \vartheta_{\ell,o}^0$.

A scalar calibration parameter c_{ℓ} is used to jointly adjust the scale (or location) across all occurrence years o via conditional parameters $\vartheta_{\ell,o}(c_{\ell})$ (see Section 5.2.4). A two-dimensional calibration vector $\mathbf{c}_{\ell} = (a_{\ell}, b_{\ell})$ is alternatively used to adjust the parameters $\vartheta_{\ell,o}(c_{\ell,o})$ by postulating a linear trend $c_{\ell,o} = a_{\ell} + b_{\ell} \cdot (o - o_0)$ (see Section 5.2.8).

Remark 5.2. The procedure can – in principle – calibrate temporal developments that involve more than two parameters, for example, nonlinear trends or periodic processes. This is, however, rarely feasible in practice due to the scarcity of data and low signal-to-noise ratios.

The generic calibration procedure can be applied independently to each feature ℓ of the generative model (i.e., to the frequency generator, the severity generator, and the pattern generators) if the following conditions are (nearly) fulfilled:

- C1: The features ℓ_1 and ℓ_2 of a claim are independent: $\ell_1 \neq \ell_2 \rightarrow Y_{\ell_1,o,j_1} \perp\!\!\!\perp Y_{\ell_2,o,j_2}$.
- C2: The occurrence years o_1 and o_2 are independent: $o_1 \neq o_2 \rightarrow Y_{\ell,o_1,j_1} \perp\!\!\!\perp Y_{\ell,o_2,j_2}$.
- C3: The N_o^{rep} claims reported for the occurrence year o are iid: $Y_{\ell,o,j} \stackrel{iid}{\sim} f_{\ell}(\boldsymbol{\theta}_{\ell,o})$.

The conditional marginal probability distributions $f_{\ell,o}(x | \mathbf{c}_{\ell})$ can be used to calibrate the distinctive features ℓ of the generative model (C1), probability distributions $f_{\ell}(x; \boldsymbol{\theta}_{\ell,o}^0)$ can be fitted to the simulated distributions $g(x | \mathcal{X}_{\ell,o})$ derived for the occurrence years o (C2), and the conditional likelihood for observing the random variables in the sets $\mathcal{Y}_{\ell,o}$ can be evaluated (C3).

Remark 5.3. Modified calibration procedures are applied in cases where these conditions are not fulfilled (see Section 5.3).

5.2.2. Calibration in three steps

The generic calibration procedure is first described for random variables defined on a *claim* level and then adapted to the *annual* and *period* levels (see definitions in Table 2).

1st step: A Monte Carlo simulation is run K times, and in run k the generative model draws the random counts $N_{o,k}^{inc}$, and it generates the random claims $\mathbf{L}_{o,j,k}^{sim}$ for $j \in \{1, \dots, N_{o,k}^{inc}\}$ and $o \in \{-T, \dots, -1\}$. The $N_{o,k}^{rep}$ reported claims $\mathbf{L}_{o,i,k}^{sim}$ (characterized by $t_{o,i,k}^{rep} \leq t^{sub}$) are projected to the reduced random tuples $\mathbf{X}_{o,i,k} := (X_{1,o,i,k}, \dots, X_{\ell,o,i,k}, \dots, X_{M,o,i,k})$ (a refined calibration approach that involves the incurred claims instead of the reported claims is described in Section 5.3.1). Following data slices $\mathcal{X}_{\ell,o}^c$ containing the simulated random variables $X_{\ell,o,i,k}$ for the feature $\ell \in \{1, \dots, M\}$ in occurrence year $o \in \{-T, \dots, -1\}$ are extracted:

$$\mathcal{X}_{\ell,o}^c := \left\{ \{X_{\ell,o,i,k}\}_{1 \leq i \leq N_{o,k}^{rep}} \right\}_{1 \leq k \leq K} \tag{5.2}$$

Respective data slices $\mathcal{X}_{\ell,o}^a$ and \mathcal{X}_{ℓ}^p containing the aggregated random variables $X_{\ell,o,k}$ and $X_{\ell,k}$ can be extracted on an *annual* and a *period* level, respectively.

- 2nd step: A parametric distribution family $f_{\ell}(x; \theta_{\ell})$ is selected for each model feature ℓ . Parametric probability distributions $f_{\ell,o}^0(x) := f_{\ell}(x; \theta_{\ell,o}^0)$ are fitted to the respective simulated distributions $g_{\ell,o}(x) := g(x | \mathcal{X}_{\ell,o})$ derived with (5.1) from the simulated sets $\mathcal{X}_{\ell,o}$. The parameters $\theta_{\ell,o}^0 = (\vartheta_{\ell,o}^0, [\varphi_{\ell,o}^0, \varsigma_{\ell,o}^0, \varrho_{\ell,o}^0, \dots])$ are obtained with, for example, the method of moments (MM) fit or the maximum likelihood estimator (MLE) fit (see Example 6.1).
- 3rd step: Conditional parameters $\theta_{\ell,o}(\mathbf{c}_{\ell}) = (\vartheta_{\ell,o}(\mathbf{c}_{\ell}), [\varphi_{\ell,o}^0, \varsigma_{\ell,o}^0, \varrho_{\ell,o}^0, \dots])$ are linked to the calibration parameter \mathbf{c}_{ℓ} with the constraint $\vartheta_{\ell,o}(\mathbf{c}_{\ell} = \mathbf{0}) \stackrel{!}{=} \vartheta_{\ell,o}^0$. The conditional probability distributions

$$f_{\ell,o}(x | \mathbf{c}_{\ell}) := f_{\ell}(x; \vartheta_{\ell,o}(\mathbf{c}_{\ell}), [\varphi_{\ell,o}^0, \varsigma_{\ell,o}^0, \varrho_{\ell,o}^0, \dots]) \tag{5.3}$$

derived for the occurrence years $o \in \{-T, \dots, -1\}$ are jointly representing the model feature ℓ to be calibrated with the parameter \mathbf{c}_{ℓ} . The conditional probability distributions $f_{\ell,o}(x | \mathbf{c}_{\ell})$ are used as a proxy for determining the probabilities of the observations $Y_{\ell,o,i}$.

Bayes' theorem is used to evaluate the conditional likelihood of \mathbf{c}_{ℓ} given the prior distribution $\pi(\mathbf{c}_{\ell})$ and the set \mathcal{Y}_{ℓ} containing the observations, to derive the MAP estimate $\mathbf{c}_{\ell,MAP}$, and the posterior distribution $p(\mathbf{c}_{\ell})$.

A straightforward approach is to calibrate the scale (or the location) of the distributions $f_{\ell,o}(x | \mathbf{c}_{\ell})$ with a scalar parameter $c_{\ell} \in \mathbb{R}$. A two-dimensional calibration vector $\mathbf{c}_{\ell} = (a_{\ell}, b_{\ell}) \in \mathbb{R}^2$ is used in the case of a linear-trend calibration (see Remark 5.2). The T conditional annual distributions $f_{\ell,o}(x | \mathbf{c}_{\ell,o})$ are thus interlinked via $c_{\ell,o} = c_{\ell}$ in the case of a scale (or location) calibration and via $c_{\ell,o} = a_{\ell} + b_{\ell} \cdot (o - o_0)$ in the case of a linear-trend calibration (see Section 5.2.8).

5.2.3. Calibration levels and iteration

A scale (or translate) calibration scalar c_{ℓ} can be derived from the set \mathcal{Y}_{ℓ}^c containing all *claim*-level observations $Y_{\ell,o,j}$, or from the set \mathcal{Y}_{ℓ}^a containing the observations $Y_{\ell,o}$ aggregated on an *annual* level, or from the set \mathcal{Y}_{ℓ}^p containing the single observation Y_{ℓ} aggregated on the *period* level. A trend calibration vector $\mathbf{c}_{\ell} = (a_{\ell}, b_{\ell})$ can be derived either from \mathcal{Y}_{ℓ}^c or from \mathcal{Y}_{ℓ}^a .

The generative model is supposed to accept the calibration tuple $\mathbf{C}^{(s)}$ and the simulation count $K^{(s+1)}$ as an input for the iteration $s + 1$ (with $s \geq 0$) and to provide the simulated set of sets $\{\mathcal{S}_k^{sim(s+1)} | \mathbf{C}^{(s)}\}_{1 \leq k \leq K^{(s+1)}}$ (4.2) as an output. The calibration framework uses the derived sets $\mathcal{X}_{\ell}^{(s+1)}$ (5.2) containing the simulated reduced random variables and the set \mathcal{Y}_{ℓ} containing the respective observations to derive the calibration parameters $\mathbf{c}_{\ell}^{(s+1)}$ for the model features $\ell \in \{1, \dots, M\}$. The generative model is adjusted by feeding the calibration tuple $\mathbf{C}^{(s+1)}$ back into the generative model where the calibration parameter $\mathbf{c}_{\ell}^{(s+1)}$ is applied to the generator(s) of the feature ℓ .

Remark 5.4. The generic iterative calibration framework can interact with the generative model via standardized interfaces. The generative model can thus be considered an independent module from the perspective of the calibration framework. The reverse is also the case; hence, the generic calibration framework can be implemented as a service that can be called by any exposure-based generative model capable of simulating claims for the observation period.

5.2.4. Link function

The calibration of the scale (or the location) of the annual probability distributions $f_{\ell}(x | \mathbf{c}_{\ell})$ with the scalar c_{ℓ} is accomplished by combining a *normal* prior $\pi_{\ell}(c_{\ell}) := \mathcal{N}(c_{\ell}; 0, \sigma_{c_{\ell}})$ with a link function

$\psi_\ell: (c_\ell, \vartheta_\kappa^0) \mapsto \vartheta_\kappa(c_\ell)$ that maps the set $\{\vartheta_\kappa^0\}$ containing the prior parameters to the set $\{\vartheta_\kappa(c_\ell)\}$ containing the adjusted parameters (the superscripts and subscripts indicated by $\kappa \in \{c_{\ell,o}, a_{\ell,o}, p_\ell\}$ depend on the calibration level). The prior distributions of the annual parameters ϑ_κ are thus decomposed into a joint (normal) prior $\pi_\ell(c_\ell)$ and annual link functions $\psi_\kappa(c_\ell) := \psi_\ell(c_\ell, \vartheta_\kappa^0)$. The reasons for this decomposition are the need to jointly calibrate the annual distributions $f_\kappa(x | c_\ell)$, the need to feed the calibration parameters c_ℓ back into the generative model, and parsimony.

The decomposition permits us to choose normal priors $\pi_\ell(c_\ell) = \mathcal{N}(c_\ell; 0, \sigma_\ell^2)$, provided that the link function ψ_ℓ adequately maps the range of validity \mathbb{R} of the calibration parameter c_ℓ to the range of validity of the respective model parameters ϑ_κ . Depending on the constraints on ϑ_κ , either a *linear* linking, an *exponential* linking, or a *logistic* (sigmoid) linking is applied (see examples in Section 6):

$$\vartheta_\kappa(c) := \psi_\ell(c, \vartheta_\kappa^0) := \begin{cases} \vartheta_\kappa^0 + c & : \text{ linear} & \psi_\ell : \mathbb{R} \rightarrow \mathbb{R} \\ \vartheta_\kappa^0 \cdot e^c & : \text{ exponential} & \psi_\ell : \mathbb{R} \rightarrow \mathbb{R}^+ \quad \text{for all } \vartheta_\kappa^0 \in \mathbb{R}^+ \\ \frac{\vartheta_\kappa^0 \cdot e^c}{\vartheta_\kappa^0 \cdot e^c + (1 - \vartheta_\kappa^0) \cdot e^{-c}} & : \text{ logistic} & \psi_\ell : \mathbb{R} \rightarrow [0, 1] \quad \text{for all } \vartheta_\kappa^0 \in [0, 1] \end{cases} \tag{5.4}$$

5.2.5. Claim-level observations and Bayes

The MAP estimate $c_{\ell,MAP}^c := \underset{c \in \mathbb{R}}{\operatorname{argmax}} \Pr(c | \mathcal{Y}_\ell^c)$ derived from the set containing the *claim*-level ('c') observations \mathcal{Y}_ℓ^c is obtained by applying Bayes' theorem and maximizing the log-likelihood function $\mathcal{L}^c(c)$:

$$\begin{aligned} \Pr(c | \mathcal{Y}_\ell^c) &= \Pr(\mathcal{Y}_\ell^c | c) \cdot \frac{\Pr(c)}{\Pr(\mathcal{Y}_\ell^c)} \propto \prod_{o,i} f_{\ell,o}^c(Y_{\ell,o,i} | c) \cdot \pi_\ell^c(c) \\ \Rightarrow \mathcal{L}^c(c) &= \ln(\Pr(c | \mathcal{Y}_\ell^c)) = \text{const.} + \sum_{o,i} \ln(f_{\ell,o}^c(Y_{\ell,o,i} | c)) + \ln(\pi_\ell^c(c)) \end{aligned}$$

where $f_{\ell,o}^c(x | c)$ is defined in (5.3) and the superscript 'c' indicates the calibration on a *claim* level. The MAP estimate $c_{\ell,MAP}^c$ is derived from the condition $\partial \mathcal{L}^c(c) / \partial c \stackrel{!}{=} 0$:

$$\sum_{o=-T}^{-1} \sum_{i=1}^{N_o^{rep}} \frac{\partial f_{\ell,o}^c(Y_{\ell,o,i} | c)}{f_{\ell,o}^c(Y_{\ell,o,i} | c)} \stackrel{!}{=} \frac{c}{\sigma_{c_\ell^c}^2} \Bigg|_{c=c_{\ell,MAP}^c} \tag{5.5}$$

5.2.6. Annual-level observations and Bayes

The calibration of the model features can also be performed on an *annual* level ('a') with the observed random variables $Y_{\ell,o}$. The *annual* probability distributions $f_{\ell,o}^a(x)$ are fitted to the simulated probability distributions $g_{\ell,o}^a(x) := g(x | \mathcal{X}_{\ell,o}^a)$ determined by the sets $\mathcal{X}_{\ell,o}^a := \{X_{\ell,o,k}\}_{1 \leq k \leq K}$. The MAP estimate $c_{\ell,MAP}^a$ is derived from the condition $\partial \mathcal{L}^a(c) / \partial c \stackrel{!}{=} 0$:

$$\sum_{o=-T}^{-1} \frac{\partial f_{\ell,o}^a(Y_{\ell,o} | c)}{f_{\ell,o}^a(Y_{\ell,o} | c)} \stackrel{!}{=} \frac{c}{\sigma_{c_\ell^a}^2} \Bigg|_{c=c_{\ell,MAP}^a} \tag{5.6}$$

5.2.7. Period-level observations and Bayes

A third option is to calibrate the feature ℓ on an aggregated *period* level ('p') with the observed random variable Y_ℓ . The probability distribution $f_\ell^p(x)$ is fitted to the simulated distribution $g_\ell^p(x) := g(x | \mathcal{X}_\ell^p)$

determined by the set $\mathcal{X}_\ell^p := \{X_{\ell,k}\}_{1 \leq k \leq K}$. The MAP estimate $c_{\ell,MAP}^p$ is derived from the condition $\partial \mathcal{L}_\ell^p(c)/\partial c \stackrel{!}{=} 0$:

$$\frac{\partial}{\partial c} \frac{f_\ell^p(Y_\ell | c)}{f_\ell^p(Y_\ell | c)} \stackrel{!}{=} \frac{c}{\sigma_{c_\ell^p}^2} \Bigg|_{c=c_{\ell,MAP}^p} \tag{5.7}$$

5.2.8. Linear trend and Bayes

The approach derived for the calibration scalar c_ℓ (i.e., $c_{\ell,o} = c_\ell$ for all o) can be extended to a calibration vector \mathbf{c}_ℓ by postulating $c_{\ell,o} = a_\ell + b_\ell \cdot (o - o_0)$, that is, $\mathbf{c}_\ell = (a_\ell, b_\ell)$. The two parameters a_ℓ and b_ℓ are assumed to be independent with priors $\pi_a(a_\ell) = \mathcal{N}(a_\ell; 0, \sigma_{a_\ell}^2)$ and $\pi_b(b_\ell) = \mathcal{N}(b_\ell; 0, \sigma_{b_\ell}^2)$.

If calibrating a linear trend of feature ℓ on the *claim* level (indicated by the superscript ‘ c ’), following system of equations for the MAP estimates $a_{\ell,MAP}^c$ and $b_{\ell,MAP}^c$ is derived from the two conditions $\partial \mathcal{L}_\ell^c(c(o; a, b))/\partial a \stackrel{!}{=} 0$ and $\partial \mathcal{L}_\ell^c(c(o; a, b))/\partial b \stackrel{!}{=} 0$:

$$\begin{aligned} \sum_{o=-T}^{-1} \sum_{i=1}^{N_o^{rep}} \frac{\frac{\partial}{\partial c} f_{\ell,o}^c(Y_{\ell,o,i}|c(o; a, b))}{f_{\ell,o}^c(Y_{\ell,o,i}|c(o; a, b))} \cdot 1 & \stackrel{!}{=} \frac{a}{\sigma_{a_\ell^c}^2} \\ \sum_{o=-T}^{-1} \sum_{i=1}^{N_o^{rep}} \frac{\frac{\partial}{\partial c} f_{\ell,o}^c(Y_{\ell,o,i}|c(o; a, b))}{f_{\ell,o}^c(Y_{\ell,o,i}|c(o; a, b))} \cdot (o - o_0) & \stackrel{!}{=} \frac{b}{\sigma_{b_\ell^c}^2} \end{aligned} \Bigg|_{(a,b)=(a_{\ell,MAP}^c, b_{\ell,MAP}^c)} \tag{5.8}$$

Respective trend parameters $a_{\ell,MAP}^a$ and $b_{\ell,MAP}^a$ (indicated by the superscript ‘ a ’) can also be derived from observations on an *annual* level:

$$\begin{aligned} \sum_{o=-T}^{-1} \frac{\frac{\partial}{\partial c} f_{\ell,o}^a(Y_{\ell,o}|c(o; a, b))}{f_{\ell,o}^a(Y_{\ell,o}|c(o; a, b))} \cdot 1 & \stackrel{!}{=} \frac{a}{\sigma_{a_\ell^a}^2} \\ \sum_{o=-T}^{-1} \frac{\frac{\partial}{\partial c} f_{\ell,o}^a(Y_{\ell,o}|c(o; a, b))}{f_{\ell,o}^a(Y_{\ell,o}|c(o; a, b))} \cdot (o - o_0) & \stackrel{!}{=} \frac{b}{\sigma_{b_\ell^a}^2} \end{aligned} \Bigg|_{(a,b)=(a_{\ell,MAP}^a, b_{\ell,MAP}^a)} \tag{5.9}$$

5.2.9. Smoothing

The simulated annual probability distributions $g_{\ell,o}(x)$ defined in (5.1) can be quite volatile unless a large number of simulated claim objects is generated. The resulting volatility of the fitted parametric distributions $f_\ell(x; \theta_{\ell,o}^0)$ can be reduced by replacing the parameters $\theta_{\ell,o}^0$ evaluated independently for each occurrence year o by smoothed parameters $\tilde{\theta}_{\ell,o} := (\tilde{\nu}_{\ell,o}, [\tilde{\varphi}_{\ell,o}, \tilde{\zeta}_{\ell,o}, \tilde{q}_{\ell,o}, \dots])$. This can, for example, be accomplished by blending the parameters $\theta_{\ell,o}^0$ and the parameters $\theta_{\ell,o}^*$, that is, by setting $\tilde{\theta}_{\ell,o} = \text{diag}(\alpha_{\ell,1}, \alpha_{\ell,2}, \dots) \times \theta_{\ell,o}^0 + \text{diag}(1 - \alpha_{\ell,1}, 1 - \alpha_{\ell,2}, \dots) \times \theta_{\ell,o}^*$ where $\alpha_\ell := (\alpha_{\ell,1}, \alpha_{\ell,2}, \dots)$ is the vector containing the credibility weights to be applied to the components of θ .

The parameters $\theta_{\ell,o}^0$ are obtained by independently fitting each distribution $f_\ell(x; \theta_{\ell,o}^0)$ to the respective simulated distribution $g_{\ell,o}(x)$. The parameters $\theta_{\ell,o}^* = \theta_\ell^*(o; \phi_\ell)$ are instead obtained by jointly fitting the set $\{f_\ell(x; \theta_\ell^*(o, \phi_\ell))\}$ to the set $\{g_{\ell,o}(x)\}$ with the help of the hyperparameters $\phi_\ell = (\varphi_{\ell,1}, [\varphi_{\ell,2}, \dots])$.

Example 5.5. *Discrete Panjer class (see Appendix B in the Online Supplementary Material) with annually calibrated means λ_o and a shared f_P , that is, $\alpha = (\alpha_\lambda, \alpha_{f_P}) = (1, 0)$:*

$$\begin{aligned} f_o(k) & := \text{Panjer}(k; \lambda_o, f_P) \\ \text{with: } \lambda_o & = \mathbb{E}[N_o] \\ \text{and: } f_P & = \frac{\sum_o \mathbb{V}[N_o]}{\sum_o \mathbb{E}[N_o]} \end{aligned} \tag{5.10}$$

5.3. Calibration of interdependent features

The generative model is assumed to contain M partially dependent calibration parameters \mathbf{c}_ℓ , which are used to calibrate M distinctive features ℓ within the generative model (depending on the feature to be calibrated, \mathbf{c}_ℓ is either a scalar or a vector).

5.3.1. Decomposition of known interdependencies

The optimal approach is to decompose known interdependencies into independent features. The calibration of the variables N_o^{inc} , τ_o^{rep} , and N_o^{rep} can, for example, be accomplished with two independent calibration parameters (an equivalent approach is applicable to the variables $I_{o,i}^*$, $\tau_{o,i}^P$, and $P_{o,i}^*$).

The count and the features of the IBNR claims are not known, that is, the annual counts N_o^{inc} of the incurred claims and the reporting lags $\tau_{o,j}^{rep}$ of the IBNR claims (for which $t_{o,j}^{rep} > t^{sub}$) are not observed. These features are, however, modeled within the generative framework, that is, the counts $N_{o,k}^{inc}$ and the reporting lags $\tau_{o,i,k}^{rep}$ are accessible for all simulated claims.

The frequencies $\lambda_o^{rep} := \mathbb{E}[N_o^{rep}] \approx \lambda_o^{inc} \cdot \rho_o(\tau = -o)$ of the claims reported for the occurrence years o depend on the frequencies $\lambda_o^{inc} := \mathbb{E}[N_o^{inc}]$ of incurred claims and the reporting patterns $\rho_o(\tau)$ evaluated at $\tau = t^{sub} - t_{o,j}^{occ} \approx -o$. The frequencies λ_o^{inc} and the patterns $\rho_o(\tau)$ are approximated by the simulated frequencies $\hat{\lambda}_o^0$ and the simulated patterns $\hat{\rho}_o(\tau)$ and considered independent, that is, they can be calibrated with respective scalar parameters c_λ and c_ρ .

Parametric reporting patterns $\rho(\tau; \tau_o^0, [\dots])$ are fitted to the simulated reporting patterns $\hat{\rho}_o(\tau)$. Conditional estimates $\hat{\lambda}_o^{rep}(\mathbf{c})$ for the frequencies of reported claims given the calibration parameter $\mathbf{c} := (c_\lambda, c_\rho)$ are obtained as follows:

$$\hat{\lambda}_o^{rep}(\mathbf{c}) = \hat{\lambda}_o^0 \cdot e^{c_\lambda} \cdot \rho(\tau = -o; \tau_o^0 \cdot e^{c_\rho}, [\dots])$$

These frequencies are used as parameters in the discrete probability distributions $f_N(k; \hat{\lambda}_o^{rep}(\mathbf{c}), [\dots])$ used to evaluate the conditional likelihood of the observations N_o^{rep} given \mathbf{c} . The MAP estimate $\mathbf{c}_{MAP} := \operatorname{argmax}_{\mathbf{c} \in \mathbb{R}^2} \Pr(\mathbf{c} | \{N_o^{rep}\}_{-T \leq o \leq -1})$ is obtained numerically by applying Bayes' theorem in combination with independent normal priors $\pi_\lambda(c_\lambda)$ and $\pi_\rho(c_\rho)$.

5.3.2. Multivariate normal approach MVN

One calibration approach is to generate a substantial number K of simulated variables $\mathbf{X}_k = (X_{1,k}, \dots, X_{M,k})$, define $\Xi_{i,k} := X_{i,k}$ and $\Upsilon_i := Y_i$ in the case of linearly linked components i , define $\Xi_{j,k} := \ln(X_{j,k})$ and $\Upsilon_j := \ln(Y_j)$ in the case of exponentially linked components j , fit a multivariate normal (MVN) distribution $\mathcal{N}(\boldsymbol{\xi}; \boldsymbol{\mu}^0, \boldsymbol{\Sigma}^0)$ to the simulated set $\{\Xi_k\}_{1 \leq k \leq K}$, specify the prior $\pi(\boldsymbol{\mu}^0) = \mathcal{N}(\boldsymbol{\mu}^0; \mathbf{v}^0, \boldsymbol{\Psi})$, derive the posterior $p(\boldsymbol{\mu}^0 | \mathbf{Y}) = \mathcal{N}(\boldsymbol{\mu}^0; \mathbf{v}^*, \boldsymbol{\Psi})$, and obtain $\mathbf{C}_{MAP} = \mathbf{v}^* - \mathbf{v}^0$ (Lee, 2012; Murphy, 2023). This approach does, however, rely on an explicit quantification of the prior means \mathbf{v}^0 and the prior correlation matrix $\boldsymbol{\Psi}$.

5.3.3. EM algorithm and Gibbs sampler

An alternative approach, which relies on the implicit specification of the correlations within the generative model, is to evaluate the calibration tuple \mathbf{C}_{MAP} iteratively with the *expectation-maximization (EM) algorithm* (Lee, 2012; Murphy, 2023). Preliminary estimates for the calibration parameters are used to generate marginal distributions, and Bayesian inference is used to update the parameters with the maximum likelihood values. The iterative process is initialized with the prior values $\mathbf{C}^{(s=0)} = (\mathbf{0}, \dots, \mathbf{0})$. The parameter tuple $\mathbf{C}^{(s)}$ evaluated at step $s \geq 1$ with the EM algorithm is determined by the observations $\{\mathbf{Y}\}$ and the simulations $\{\mathbf{X}^{(s)} | \mathbf{C}^{(s-1)}\}$ (see also Remark 5.3 and Section 5.2.3), that is, the process is governed by a Monte Carlo Markov chain (MCMC). The update of the parameter tuple $\mathbf{C}^{(s)}$ in step s can be done in several ways. In the case of relevant couplings, the *basic Gibbs sampler* is applied

(Lee, 2012; Murphy, 2023), where the iteration is run in cycles (with M steps per cycle), and where a single element c_{ℓ_i} is updated in an iteration step. Alternatively, a *block Gibbs sampler* might be applied where multiple or all parameters are updated in an iteration step.

Finding a converging solution with the basic Gibbs sampler can involve a large number of iteration steps. It is thus computationally expensive (or even intractable), and paying the price only makes sense if the set containing the observational data ensures a robust solution. This condition is rarely fulfilled in the reinsurance rating process where the submission data comprises a small number (typically 10 to 100 per year) of medium and large partially developed claims. The intrinsic uncertainty in the observational data and the epistemic uncertainty in the generative model are thus significant (see Figures 6 and 7). These irreducible uncertainties translate into respective uncertainties in the parameter estimates. The calibration process aims to find a robust solution within these uncertainties and reflect the uncertainties in the posterior parameter distributions and the posterior predictive model.

5.3.4. Combined approach

A third option is to combine the two approaches. The variables \mathbf{Z} are decomposed into weakly coupled components Z_k and strongly coupled pairs (Z_i, Z_j) . The MVN approach is applied to the sub-spaces in each iteration step of the EM algorithm.

5.4. Overall process and calibration of priors

The rating of reinsurance contracts is an integral part of the (annual) renewal process. Underwriters and pricing actuaries are then absorbed by the interaction of the reinsurer with its cedents, and they cannot calibrate the generative model for each incoming submission. However, the model calibration and the rating of the future covers with the calibrated model are two distinct processes that can be performed at separate times of the year. The calibration of the generative models can be performed outside of the renewal season by considering all information available at that time.

A differentiation is made between the initial setup, the initial calibration, subsequent updates of the setup, and calibration updates. The initial setup involves capturing the relevant market information, exposure profiles, insurance conditions, and claims experience for the reporting period and specifying the prior distributions (see Figure 2). The initial calibration is an interactive and iterative process performed by an expert.

The calibrated model is fed with the most recent exposure profiles (and other submission data) during the renewal process, and the model is used to rate the future contracts. A subsequent (periodic) recalibration can be performed after capturing any missing information from recent submission updates. This recalibration can either be performed (semi-) automatically in the case of minor changes in the underlying portfolio, or it is done analogously as the initial calibration in the case of significant changes in the underlying portfolio or if the generative model has been modified substantially.

A similar process can be implemented on a portfolio level to periodically update the prior assumptions, that is, with the help of a hierarchical model (see, e.g., Sections 3.6 and 3.7 in Murphy 2023). The reinsurer can use the set $\{C_\ell\}_{1 \leq \ell \leq N^{ced}}$ containing the calibration tuples C_ℓ derived for the N^{ced} cedents operating in the same market to update the hyperparameters used to specify the priors.

6. Sample calibration

6.1. Calibration case study with a toy model

Example 6.1. *The calibration of the frequencies λ_o of the annual claim-counts distributions $N_o^{rep} \sim f_N(\lambda_o, f_{P,o})$ with a scale parameter c can be performed either on an annual level or on a period level and a linear trend with the parameter $c_\ell = (a_\ell, b_\ell)$ can be calibrated on an annual level (the Panjer factor f_p*

Table 3. Parameters and statistics for the generative process and the generative models shown in Figure 5.

Model	Calibration			Annual			Average	Projected	
	a or c	b	o_0	μ_{-12}	μ_{-1}	f_p	\hat{N}	μ	μ_1
Generative process				5.0	15.0	1.25		9.2	18.3
Actual pmf				5.0	15.0	1.25	8.8	9.2	18.3
Prior model				10.0	20.0	1.00		14.5	22.7
Simulated pmf				10.5	19.9	1.08		14.4	20.8
Level calibration	-0.49			6.2	12.3	1.00		8.9	14.0
Simulated pmf				6.3	12.9	0.99		9.0	14.1
Linear-trend cal.	-0.54	0.057	-6.5	4.3	16.0	1.00		9.0	20.3
Simulated pmf				4.3	16.2	1.02		9.0	20.0
Random variables used for the ‘toy model’	N_o μ_o^0			(3, 4, 6, 7, 10, 10, 7, 8, 12, 8, 18, 13) (10.5, 10.5, 11.7, 11.9, 12.8, 14.2, 14.8, 15.2, 16.2, 16.9, 18.8, 19.9)					

is defined in Section 6.5.1 and in Appendix B). The calibration of the sample models specified in Table 3 is shown in Figure 5.

The parameters and critical statistics for the generative process and the three generative models illustrated in Figure 5 for the discrete random variable $N_o := N_o^{rep}$ are summarized in Table 3. The first line of each entry contains the model parameters and the expected statistics. The second line contains the statistics derived from the actual pmf (probability mass function) or the simulated pmf, respectively. The selected random variables listed at the bottom are used as input for the ‘toy model’ (see Appendix C.1 in the [Online Supplementary Material](#) for details).

The model used to emulate the generative process is shown in panel (a), and the initial generative model is shown in panel (b). The simulated annual and aggregate pmfs are fitted with $\log N$ distributions. The scale calibration (see (A.1) in the [Online Supplementary Material](#)) based on aggregated *period*-level statistics is shown in panel (c), and the linear-trend calibration (see (A.3) in the [Online Supplementary Material](#)) based on *annual* statistics is shown in panel (d).

6.2. Calibration statistics

Additional insight is gained by running the calibration procedure multiple times and analyzing the resulting statistics. This is either done on a conditional basis, that is, by keeping the ‘observations’ $\{N_o\}_{-T \leq o \leq -1}$ fixed as shown in Figure 6, or on an unconditional basis, that is, by drawing a new set of observations $\{N_o^{(m)}\}_{-T \leq o \leq -1}$ in each calibration run m , as shown in Figure 7.

The systematic bias observed in Figure 6 for the deviation statistics $F(\Delta X)$ stems from the different trend assumptions and stochastic processes specified in the underlying models. The calibration of the scale (determined by a and c) is more robust than the calibration of the trend (determined by b^*); therefore, the bias and the uncertainty of the overall mean $\hat{\mu}$ are minor compared to the bias and uncertainty of the means at $o = -T$, $o = -1$, and $o = +1$ (determined by b^*). The scale-only calibration (top panels) preserves the optimistic trend assumption; therefore, it leads to a positive bias for the calibrated mean at $o = -T$, a negative bias for the calibrated mean at $o = -1$, and a substantial negative bias for the projected mean at $o = +1$. The linear-trend calibration (bottom panels) overcompensates the underestimated trend in the initial model; hence, it leads to a (moderate) positive bias for the projected mean at $o = +1$.

The variation in the unconditional case (Figure 7) increases by about one order of magnitude compared to the variation in the conditional case. However, the variation of the scale parameters a and c in the unconditional case is insignificant compared to the variation of the parameter b^* . The scale-only calibration (top panels) thus leads to projections with a significant systematic bias but low uncertainty. In contrast, the linear-trend calibration (bottom panels) leads to projections with a small systematic bias but a significant uncertainty.

6.3. Features of the calibration

The key features of the calibration process are presented in Figure 5. The annual pmfs (derived from the generative model) are represented by *negative binomial* distributions $f_o(k) := \mathcal{NB}(k; r_o, p_o)$. They can be adequately approximated by $\log\mathcal{N}$ distributions if $\Pr(N_o = 0) \ll 1.0$. The parameters r_o are assumed to increase with time, while the parameters $p_o = 1/f_p$ are kept constant (see (5.10) and Appendix B.3 in the [Online Supplementary Material](#)). The respective overall distribution of the average over the period is shown on the right.

The probability distributions underlying the generative process leading to the observations $\{N_o\}_{-T \leq o \leq -1}$ are not known. These observations are, instead, used to adjust the respective calibration parameter c_ℓ within the generative model. This is accomplished by fitting parametric distributions to the simulated distributions (see Section 6.6). The calibration of the scale (or the location) with a scalar c_ℓ (i.e., $c_{\ell,o} = c_\ell$ for all o) is performed with statistics on a *claim* level, on an *annual* level, or on a *period* level. The statistics on a *claim* level and on an *annual* level can also be used to calibrate a linear trend (i.e., $c_{\ell,o} = a_\ell + b_\ell \cdot (o - o_0)$).

Remark 6.2. The calibration procedure can be expanded to higher-dimensional vectors c_ℓ . Given the interdependencies between parameters, the uncertainties, and the need to provide an adequate prior for each parameter, it is rarely possible to calibrate any feature beyond a linear trend.

6.4. Continuous distributions

Most elements within the set of reduced random variables (see Section 4.3) are nonnegative, that is, they have a support \mathbb{R}^+ or \mathbb{Z}^+ . The underlying modules within the generative model are, therefore, calibrated via an exponential linking (5.4). The incurred lag can, however, be positive or negative, that is, it has a support \mathbb{R} , and the underlying modules are calibrated via a linear linking.

In the case of continuous processes, $\log\mathcal{N}$ distributions are fitted to the simulated distributions in the exponential-linking case and \mathcal{N} distributions are fitted to the simulated distributions in the linear-linking case. Combining these distributions with \mathcal{N} priors for the calibration parameters and plugging the pdfs (probability density function) into (5.5), (5.6), (5.7), (5.8), and (5.9) leads to the MAP estimates for c_{MAP}^c , c_{MAP}^a , c_{MAP}^p and (a_{MAP}^c, b_{MAP}^c) , (a_{MAP}^a, b_{MAP}^a) , respectively (see (A.1) and (A.3) in the [Online Supplementary Material](#)).

Remark 6.3. The conjugate prior distribution of a $\log\mathcal{N}$ (or \mathcal{N}) distribution with known parameter σ is a \mathcal{N} distribution and the posterior predictive distribution is also a $\log\mathcal{N}$ (or \mathcal{N}) distribution. The fitted $\log\mathcal{N}$ (or \mathcal{N}) distributions are used as a proxy for calibration purposes; they are not used within the generative model. The posterior predictive $\log\mathcal{N}$ (or \mathcal{N}) distributions can thus not be fed back into the generative model; therefore, ‘scale’ (or ‘translate’) calibration parameters are used as a proxy.

6.5. Discrete: The Panjer class

6.5.1. Exponential linking

The three distribution families within the *Panjer* class (Panjer, 1981; Sundt and Jewell, 1981) are suitable to fit a broad variety of discrete processes (see Appendix B in the [Online Supplementary Material](#)). The

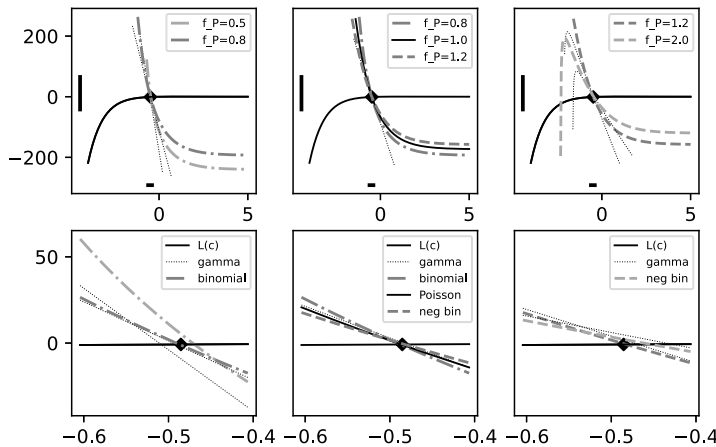


Figure 8. Fitting comparison. Representation of $L(c) = c/e^c$ and of $R_D(c) = R(c | N_o, \mu_o^0, f_P, \sigma_c)$ for various distributions D . The random variables N_o and μ_o^0 are taken from the ‘toy model’ (see Table 3) and combined with $f_P \in \{0.5, 0.8, 1.0, 1.2, 2.0\}$. The distribution D is selected depending on f_P : $D = \mathcal{P}$ if $f_P = 1$, $D = \mathcal{NB}$ if $f_P > 1$, and $D = \mathcal{B}$ if $f_P < 1$ (see Table B.1 in the [Online Supplementary Material](#)). The approximation $D = \mathcal{G}a$ (dotted lines) is evaluated for all selections of f_P . The diamonds depict the root $c_{MAP} = -0.48$ evaluated with $D = \log N$ (see (A.1) in the [Online Supplementary Material](#)). The top panels show $L(c)$ and $R_D(c)$ for the wide range $c \in [-5, +5]$ around the prior $c = 0.0$, and the bottom panels show the curves for the narrow range $c \in [-0.6, -0.4]$.

class comprises the *Poisson* family \mathcal{P} , the *negative binomial* family \mathcal{NB} , and the *binomial* family \mathcal{B} . The variance-over-mean ratio (Panjer factor) $f_P = \mathbb{V}[N]/\mathbb{E}[N]$ is used to discriminate between the three families: $f_P = 1$ for \mathcal{P} , $f_P > 1$ for \mathcal{NB} , and $f_P < 1$ for \mathcal{B} .

The calibration of the discrete process $N_{o,i}^{inc}, N_{o,i}^{rep}, N_{o,i}^{clo}, N_{o,i}^I$, and $N_{o,i}^P$ with support \mathbb{Z}^+ is performed in an analogous way as the calibration of continuous processes with support \mathbb{R}^+ . A parametric *Panjer*-class distribution is fitted to the respective simulated discrete distribution and used as a proxy when calibrating the underlying discrete process within the generative model. The calibration parameter c_ℓ with normal prior $\pi_\ell(c_\ell)$ is linked to the frequency parameter (i.e., λ for \mathcal{P} , n for \mathcal{B} , and r for \mathcal{NB}) via an exponential linking. The \mathcal{N} distribution is not a conjugate prior of the *Panjer*-class families, but the MAP estimate of the calibration parameters with a \mathcal{N} prior can be found iteratively (see formulas in Table B.1 in the [Online Supplementary Material](#)).

A discrete process can also be calibrated with a continuous distribution if the frequency is sufficiently large (see Appendix A.3 in the [Online Supplementary Material](#)). A gamma distribution $\mathcal{G}a$ can, for example, approximate the *Panjer*-class distributions without the need to discriminate between the three families. Another alternative is to derive the calibration parameters with the help of a $\log N$ approximation (see Figure 5). A comparison of the calibration of discrete processes performed with several discrete and continuous distributions is shown in Table 5 and in Figure 8.

Example 6.4. A comparison of the calibration scalars c_{MAP}^a and c_{MAP}^p derived with the help of different discrete and continuous parametric distribution families for three *Panjer* class models defined in Table 4 is provided in Table 5 and discussed in Section 6.6.

6.5.2 Logistic linking

A generative model might also comprise processes involving a fixed support $\{0, \dots, n\}$ or processes involving a fixed threshold r for a discrete random variable N . Such features can be calibrated via a logistic linking of the calibration parameter c_ℓ to a probability parameter $p_o \in [0, 1]$, for example,

Table 4. Parameters used for the model comparison.

Model	T	λ_{-T}^{ref}	λ_{-1}^{ref}	$\widehat{\ln \lambda_o^{ref}}$	β	K	f_P^{NB}	f_P^P	f_P^B
Generative process	10	10.0	12.5	2.414	5.0	1	1.5	1.0	0.6
Generative model	10	5.0	10.0	1.956	5.0	50	1.5	1.0	0.6

Table 5. Comparison of the calibration parameters derived on an annual and a period level with different approximations for the three model cases NB , P , and B .

Model		Calibration parameters c_{MAP}^a and c_{MAP}^p									
Case	f_P^{sim}	NB^a	NB^p	$P^a = P^p$	B^a	B^p	Ga^a	Ga^p	$\log N^a$	$\log N^p$	
NB $f_P: 1.5$ $c^{mod}: 0.458$	1.527	0.423	0.447	0.447	–	–	0.322	0.451	0.345	0.452	
	1.452	0.342	0.344	0.343	–	–	0.342	0.349	0.406	0.350	
	1.598	0.404	0.386	0.385	–	–	0.437	0.391	0.465	0.391	
	1.431	0.304	0.312	0.311	–	–	0.286	0.316	0.387	0.317	
	1.510	0.559	0.569	0.569	–	–	0.547	0.573	0.691	0.572	
NB mean	1.504	0.406	0.412	0.411	–	–	0.387	0.416	0.459	0.416	
P $f_P: 1.0$ $c^{mod}: 0.458$	0.961	–	–	0.386	0.387	0.386	0.361	0.391	0.485	0.392	
	0.921	–	–	0.543	0.547	0.543	0.498	0.547	0.596	0.546	
	0.906	–	–	0.357	0.355	0.357	0.380	0.362	0.419	0.363	
	1.140	0.433	0.436	0.436	–	–	0.398	0.441	0.490	0.441	
	0.989	–	–	0.294	0.294	0.293	0.207	0.298	0.268	0.299	
P mean	0.983	–	–	0.403	–	–	0.369	0.408	0.452	0.408	
B $f_P: 0.6$ $c^{mod}: 0.458$	0.579	–	–	0.333	0.331	0.333	0.330	0.338	0.380	0.338	
	0.639	–	–	0.418	0.429	0.418	0.380	0.422	0.408	0.422	
	0.629	–	–	0.528	0.577	0.528	0.469	0.533	0.561	0.533	
	0.622	–	–	0.408	0.411	0.407	0.403	0.411	0.479	0.412	
	0.629	–	–	0.527	0.533	0.527	0.522	0.531	0.542	0.531	
B mean	0.619	–	–	0.443	0.456	0.443	0.421	0.447	0.474	0.447	

$N \sim \mathcal{B}(n, p_o(c_\ell))$ or $N \sim \mathcal{NB}(r, p_o(c_\ell))$. The respective equations to be fulfilled by the calibration parameter $c_{\ell,MAP}$ via a logistic linking (5.4) to the annual probability parameters p_o are listed in the two bottom rows of Table B.1 in the [Online Supplementary Material](#).

One B example (Bernoulli case) is a random indicator variable $\mathbb{1}_{EN}$ for ‘El Niño’ years (i.e., $n = 1$). Another B example is an upper bound on the random count of payments per year, for example, $0 \leq N_{o,i}^P \leq n = 12$ for all (o, i) .

A NB example is a hypothetical dual-trigger contract that covers the economic loss of $N_T = N_{TF} + N_{TT}$ claims stemming from TF events or TT events (i.e., the first trigger must be ‘True’ and the second trigger can be ‘False’ or ‘True’) as long as the count N_{TT} of TT events does not exceed the fixed threshold $r \in \mathbb{R}^+$: $N_{TF} \sim \mathcal{NB}(r, p)$, and $N_{TT} = \lfloor r \rfloor$.

6.6. Method comparison

The various calibration approaches briefly described above (see details in Appendices A and B in the [Online Supplementary Material](#)) are applied to the NB case (i.e., $f_P > 1$), the P case (i.e., $f_P \approx 1$), and the B case (i.e., $f_P < 1$). The parameters used to generate the set $\{N_o^{obs}\}$ containing the observations stemming

from the generative process and the parameters used to generate the set $\{N_{o,k}^{sim}\}$ representing the generative model are listed in Table 4. The reference frequencies $\{\lambda_o^{ref}\}$ for the occurrence year o are derived from λ_{-T}^{ref} and λ_{-1}^{ref} via an exponential interpolation. The simulated frequencies $\lambda_o^{sim} \sim \mathcal{Ga}(\alpha_o, \beta)$ with $\alpha_o = \lambda_o^{ref} \cdot \beta$ are used to simulate the K random variables $N_{o,k}^{sim} \sim \mathcal{Panjer}(\lambda_o^{sim}, f_p)$. The derived calibration parameters c_{MAP}^a (annual level) and c_{MAP}^p (period level) obtained with the various approaches are compared in Table 5. Five simulations are run for each of the three cases $f_p = 1.5$ (\mathcal{NB} case), $f_p = 1.0$ (\mathcal{P} case), and $f_p = 0.6$ (\mathcal{B} case), respectively. The case, the respective model Panjer factor f_p , and the calculated calibration parameter c^{mod} (the difference of the mean log frequencies $\ln \lambda_o^{ref}$) are indicated in column 1 for each of the three cases. The simulated Panjer factors f_p^{sim} per run are shown in column 2.

The respective calibration parameters c_{MAP}^a and c_{MAP}^p are shown in columns 3–11. The calibration parameters are only evaluated with the \mathcal{NB} approach when $f_p^{sim} > 1$, and with the \mathcal{B} approach when $f_p^{sim} < 1$. The parameters derived with the \mathcal{NB} approach in the \mathcal{NB} case, with the \mathcal{P} approach in the \mathcal{P} case, and with the \mathcal{B} approach in the \mathcal{B} case, are shown in bold (see Appendix C.2 in the [Online Supplementary Material](#) for details).

Using the \mathcal{P} approximation (with $c_{MAP}^a = c_{MAP}^p$) as a reference indicates that the scale calibration parameters c_{MAP}^p derived from aggregate period statistics with the different approaches are remarkably close. The scale calibration parameters c_{MAP}^a derived from annual statistics vary between the approaches, but the variations are not substantial.

Hence, the model comparison indicates that a logN approximation based on annual statistics might be used to derive linear-trend parameters a_{MAP}^a and b_{MAP}^a with the help of (A.3) in the [Online Supplementary Material](#). Attempting to derive these two parameters with other models, for example, the \mathcal{Ga} family or the \mathcal{Panjer} class yields more complex equations that must be solved numerically (see Table B.1 in the [Online Supplementary Material](#)).

6.7. Example of a complete calibration

An example of an iterative calibration of seven features is shown in Figures 9 and 10 (see Appendix C.3 in the [Online Supplementary Material](#) for details). A marginal linear-trend calibration is applied to the annual claim counts N_o^{rep} , the average annual claims severities \hat{I}_o^* , the average annual lags τ_o^{rep} , τ_o^l , τ_o^p , and the average annual counts \hat{n}_o^l and \hat{n}_o^p . The distributions of the annual counts of closed claims N_o^{clo} and the distributions of the annual average paid amounts \hat{P}_o^* are strongly dependent ‘test variables’ (they are monitored but not calibrated). The distributions drawn at $o = 0$ show the respective aggregate distributions on a period level. The different linear-trend assumptions underlying the model used to emulate the generative process and the initial generative model, respectively, can be recognized in the pmfs and pdfs shown in Figure 9. The two models are sampled 200 times, and parametric distributions are fitted to the simulated marginal distributions.

The calibration procedure is iterated five times after the initial sampling (iteration 0). The marginal linear-trend parameters $a_{\ell,MAP}^a$ and $b_{\ell,MAP}^a$ are evaluated in the first step of each iteration, the respective parameters within the generative model are adjusted in the second step, and the adjusted model is used to generate 200 new samples in the third step. A robust solution is obtained by reducing the prior parameters $\sigma_{a_{\ell}^a}$ and $\sigma_{b_{\ell}^a}$ after each iteration. The resulting marginal distributions derived from the calibrated model are shown in Figure 10. The residual offsets of corresponding modes are small relative to the respective standard deviations, that is, the highest density regions of corresponding distributions are overlapping. The calibrated generative model is thus well aligned with the (hidden) ‘true’ model used to emulate the generative process.

7. Conclusions

A generic framework that permits calibrating the relevant features within exposure-based generative models with the claims experience \mathcal{Y} of a cedent has been introduced. Depending on the generative

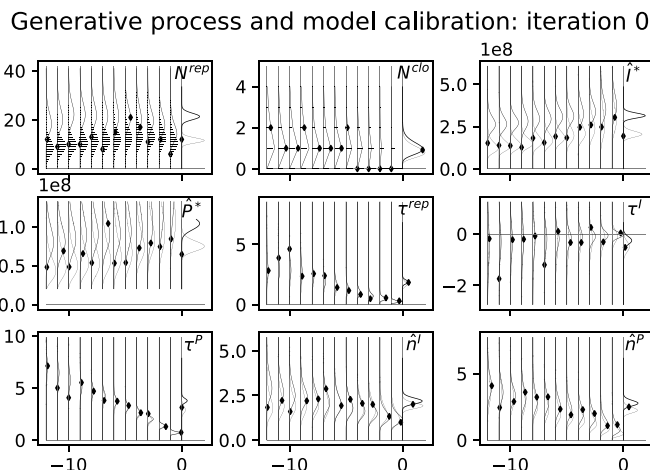


Figure 9. Generative process and initial model. Temporal evolution of the probability distributions (shown at $o \in \{-T, \dots, -1\}$ with $T = 12$) and the aggregated distribution (shown at $o = 0$) of nine reduced variables. The underlying (hidden) pmfs (horizontal lines) and pdfs (continuous lines) of the generative process are drawn with thin lines. The pdfs of the underlying generative model are drawn with thicker lines, and the diamonds represent the ‘observations’ drawn from the generative process. Large offsets of the corresponding modes (large relative to the respective standard deviations) can be identified for the claims count N^{rep} , the incurred amount \hat{I}^* , and the (derived) cumulative paid amount \hat{P}^* .

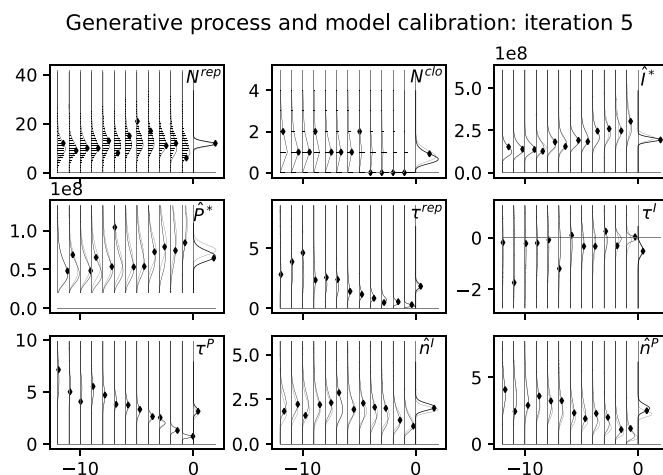


Figure 10. Generative process and calibrated model. Same representation as in Figure 9. The linear-trend parameters within the generative model are calibrated in five iteration steps with seven marginal distributions. The residual offsets between the corresponding modes of the (hidden) true model and the calibrated generative model are minor compared to the standard deviations. Note: The count of closed claims N^{clo} and the paid amount \hat{P}^* are dependent ‘test variables’ that are monitored but not calibrated.

model, the calibration is accomplished by applying the maximum a posteriori (MAP) point estimates C_{MAP} or by applying the posterior probability distribution $Pr(C | \mathcal{Y})$.

Generative models attempt to emulate all relevant features of the generative process underlying the claims experience of a cedent, that is, the simulated claims are represented in the same way as the observed claims in the submission data. The detailed features of the simulated high-dimensional claim objects are required for the rating process, but these objects cannot be directly used for calibration. The calibration is thus performed with low-dimensional *reduced variables* and simulated marginal probability distributions. Rating and calibration are two distinct processes performed with the same generative model; the two processes can, however, be performed at different points in time.

Standard experience-based rating approaches project the past claims experience to the future, and the incomplete claims are developed. In contrast, the integrated framework projects the exposure models to the past and attempts to replicate the characteristics of the claims experience. A calibrated generative model is, therefore, coherently reflecting all information available to the reinsurer. The calibrated model is projected to the cover period and used for rating, while the standard experience-based rating is obsolete.

The generic calibration framework interacts with the generative model via standardized interfaces, and it is applicable to all nonlife lines of business. In the case of short-tail business, it is used to calibrate the claims frequency and the claims severity. In the case of long-tail lines, it is used to calibrate the claims frequency, the claims severity, the reporting pattern, the incurred pattern, and the payment pattern.

The marginal probability distributions encountered in reinsurance ratings are assumed to be unimodal and heavy-tailed. The calibration of such distributions with *Bayesian inference* is determined by the highest density region (*HDR*), and it only depends on relative probabilities. It is thus sufficient to adequately fit the simulated distributions in the *HDR*, and there is no need to normalize the potentials defining the shape of the distributions. Continuous potentials as, for example, defined by the normal (\mathcal{N}) family, the log-normal ($\log\mathcal{N}$) family, or the gamma ($\mathcal{G}\alpha$) family, can, therefore, be used to calibrate discrete processes if the frequencies are not too small. This is of particular interest when attempting to calibrate the linear trend of a feature.

Interdependencies between the calibration features, for example, the (unknown) count of incurred claims, the reporting lag, and the count of reported claims, depend on the characteristics of the underlying generative processes. Such dependencies are considered in the calibration by reflecting the characteristics within the generative model and by iterating the calibration procedure with an expectation-maximization (EM) algorithm.

The set of calibration parameters derived during a certain period for cedents operating in the same market can be used to update the hyperparameters used to specify the prior distributions in a subsequent period.

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/asb.2024.17>.

Data availability statement. Replication data and code can be found in GitHub: <https://github.com/Steivan/GIRF>.

Acknowledgments. The author thanks the editor and the referees for their constructive feedback.

Competing interests. The author declares none.

References

- Albrecher, H. (2017) *Reinsurance: Actuarial and Statistical Aspects*, 1st ed. Wiley. https://books.google.ch/books/about/Reinsurance.html?id=HZjObwAACAAJ&redir_esc=y
- Antal, P. (2009) *Mathematical Methods in Reinsurance*. https://www.actuaries.ch/fr/downloads/aid!b4ae4834-66cd-464b-bd27-1497194efc96/id!96/Reinsurance_Script2009_v2.pdf

- Bernegger, S. (1997) The Swiss Re exposure curves and the MBBEFD distribution class. *ASTIN Bulletin*, **27**(1), 99–111. <https://doi.org/10.2143/AST.27.1.563208>
- Bernegger, S. (2012) Casualty Exposure Rating: ISO Approach vs European ILF. https://www.researchgate.net/publication/262000842_Casualty_Exposure_Rating_ISO_Approach_vs_European_ILF
- Billeter, S. and Salghetti-Drioli, F. (2016) Liability Risk Drivers: Bringing a forward-looking perspective into liability modelling. https://www.swissre.com/dam/jcr:6524ff3e-c35d-4f65-a99c-aa11604990b2/SwissRe_ForwardLookingModeling.pdf
- Bühlmann, H. and Gisler, A. (2005) *A Course in Credibility Theory and its Applications*. Springer. <http://catalogue.bnf.fr/ark:/12148/cb44685059k>
- Bühlmann, H. and Straub, E. (1970) Glaubwürdigkeit für Schadensätze. *Bulletin/Association of Swiss Actuaries*. <https://doi.org/10.5169/seals-967024>
- CARe. (2024) Casualty Actuaries in Reinsurance (CARe). <https://www.casact.org/membership/special-interest-sections/casualty-actuaries-reinsurance-care>
- CAS. (2024) Casualty Actuarial Society (CAS). <https://www.casact.org/publications-research>
- Clark, D.R. (2014) Basics of Reinsurance Pricing. https://www.casact.org/sites/default/files/2021-03/8_Clark.pdf
- Daykin, C.D., Pentikäinen, T. and Pesonen, M. (1994) *Practical Risk Theory for Actuaries*. Chapman & Hall. https://www.routledge.com/Practical-Risk-Theory-for-Actuaries/Daykin-Pentikainen-Pesonen/p/book/9780412428500?gclid=Cj0KCQjwmiCoBhDxARIsABXkXlIxp1gQLNcb_tdN8zWedJROfwjqU_J4r5MGemYdk5BBOMJ_ZJNXQLcaAjGLEALw_wcB
- Desmedt, S., Snoussi, M., Chenut, X. and Walhin, J.F. (2012) Experience and exposure rating for property per risk excess of loss reinsurance revisited. *ASTIN Bulletin*, **42**(1), 233–270. <https://doi.org/10.2143/AST.42.1.2160742>
- Devlin, S. (2018) Introduction to Exposure Rating. https://www.casact.org/sites/default/files/presentation/reinsure_2018_presentations_b-1_fishman_2.pdf
- Fink, D. (1997) A Compendium of Conjugate Priors. https://www.researchgate.net/publication/238622435_A_Compendium_of_Conjugate_Priors
- Gelman, A. (2013) *Bayesian Data Analysis*, 3rd ed. CRC Press. <http://www.stat.columbia.edu/gelman/book/>
- Goffard, P.-O., Piette, P. and Peters, G.W. (2023) *Market-based insurance ratemaking* (Working Papers No. hal-04297811). HAL. <https://ideas.repec.org/p/hal/wpaper/hal-04297811.html>
- Guggisberg, D. (2004) Exposure rating. <https://www.swissre.com/dam/jcr:7137dac0-83a6-4cfa-80a4-93d33c35562f/exposure-rating-brochure.pdf>
- Harrison, W.J., Bays, P.M. and Rideaux, R. (2023) Neural tuning instantiates prior expectations in the human visual system. *Nature Communications*, **14**(1), 5320. <https://doi.org/10.1038/s41467-023-41027-w>
- Huang, E. (2014). Moving beyond history: A loss driver approach to projecting and quantifying casualty exposure. https://www.casact.org/sites/default/files/presentation/reinsure_2014_handouts_paper_3349_handout_2145_0.pdf
- Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2019) *Loss Models: From Data to Decisions*, 5th ed. Wiley. <https://www.wiley.com/en-us/Loss+Models:+From+Data+to+Decisions,+5th+Edition-p-9781119523789>
- Lee, P.M. (2012) *Bayesian Statistics: An Introduction*, 4th ed. Wiley. <https://www.wiley.com/en-us/Bayesian+Statistics%3A+An+Introduction%2C+4th+Edition-p-9781118332573>
- Mata, A.J. and Verheyen, M.A. (2005) An improved method for experience rating reinsurance treaties using exposure rating techniques. In *CASUALTY ACTUARIAL SOCIETY FORUM Spring 2005 Including the Reinsurance Call Papers*, p. 171. https://www.casact.org/sites/default/files/database/forum_05spforum_05spf171.pdf
- Mildenhall, S.J. (2022) *Pricing Insurance Risk: Theory and Practice*. John Wiley & Sons, Incorporated. <https://www.wiley.com/en-ae/Pricing+Insurance+Risk%3A+Theory+and+Practice-p-9781119755678>
- Murphy, K.P. (2022) *Probabilistic Machine Learning: An Introduction*. The MIT Press. <https://probml.github.io/pml-book/book1.html>
- Murphy, K.P. (2023) *Probabilistic Machine Learning: Advanced Topics*. MIT Press. <https://probml.github.io/pml-book/book2.html>
- Palmer, T.N. (2022) *The Primacy of Doubt: From Climate Change to Quantum Physics, How the Science of Uncertainty can Help Predict and Understand Our Chaotic World*, 1st ed. Oxford University Press. <https://global.oup.com/academic/product/the-primacy-of-doubt-9780192843593?cc=ch&lang=en&>
- Panjer, H.H. (1981) Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, **12**(1), 22–26. <https://doi.org/10.1017/S0515036100006796>
- Parr, T., Pezzulo, G. and Friston, K.J. (2022) *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. MIT Press. <https://doi.org/10.7551/mitpress/12441.001.0001>
- Riegel, U. (2010) On fire exposure rating and the impact of the risk profile type. *ASTIN Bulletin: The Journal of the IAA*, **40**(2), 727–777. <https://doi.org/10.2143/AST.40.2.2061135>
- Schnieper, R. (1991) Separating true IBNR and IBNER claims. *ASTIN Bulletin*, **21**(1), 111–127. <https://doi.org/10.2143/AST.21.1.2005404>
- Sundt, B. and Jewell, W.S. (1981) Further results on recursive evaluation of compound distributions. *ASTIN Bulletin*, **12**(1), 27–39. <https://doi.org/10.1017/S0515036100006802>
- Verisk. (2024) Underwriting and Rating Solutions. <https://www.verisk.com/insurance/>
- White, S. and Mrazek, K. (2004) Advanced Exposure Rating: Beyond the Basics. https://www.casact.org/sites/default/files/presentation/reinsure_2004_handouts_whitemrazek.pdf
- Zhang, Y., Dukic, V. and Guszczka, J. (2012) A Bayesian non-linear model for forecasting insurance loss payments. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, **175**(2), 637–656. <http://www.jstor.org/stable/23251253>