



HAL
open science

A hierarchical reserving model for reported non-life insurance claims

Jonas Crevecoeur, Katrien Antonio

► **To cite this version:**

Jonas Crevecoeur, Katrien Antonio. A hierarchical reserving model for reported non-life insurance claims. Insurance: Mathematics and Economics, 2022, 10.1016/j.insmatheco.2022.02.005 . hal-04015725

HAL Id: hal-04015725

<https://hal.science/hal-04015725v1>

Submitted on 6 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A hierarchical reserving model for reported non-life insurance claims

Jonas Crevecoeur^{1,3,*} and Katrien Antonio^{1,2,3,4}

¹*Faculty of Economics and Business, KU Leuven, Belgium.*

²*Faculty of Economics and Business, University of Amsterdam, The Netherlands.*

³*LRisk, Leuven Research Center on Insurance and Financial Risk Analysis, KU Leuven, Belgium.*

⁴*LStat, Leuven Statistics Research Center, KU Leuven, Belgium.*

* *Corresponding author: jonas.crevecoeur@kuleuven.be, Faculty of Economics and Business, Naamsestraat 69, KU Leuven, Belgium.*

April 18, 2022

Abstract

Traditional non-life reserving models largely neglect the vast amount of information collected over the lifetime of a claim. This information includes covariates describing the policy (e.g. the value of the insured risk), claim cause (e.g. hail) as well as the detailed claim's history (e.g. settlement, payment, involvement lawyer). We present the hierarchical reserving model as a modular framework for integrating a claim's history and claim-specific covariates into the development process. Hierarchical reserving models decompose the joint likelihood of the development process over time. Moreover, they are tailored to the portfolio at hand by adding a layer to the model for each of the registered events (e.g. settlement, payment). Layers are modelled with classical techniques (e.g. generalized linear models) or machine learning methods (e.g. gradient boosting machines) and using claim-specific covariates. As a result of its flexibility, this framework incorporates many existing reserving models, ranging from aggregate models designed for runoff triangles to individual models using claim-specific covariates. This connection allows us to develop a data-driven strategy for choosing between aggregate and individual reserving; an important decision for reserving practitioners that is largely left unexplored in scientific literature. We illustrate our method with a case study on a real life insurance data set. This case study provides new insights in the covariates driving the development of claims and demonstrates the flexibility and robustness of the hierarchical reserving model over time.

JEL classification: G22

Keywords: individual claims reserving, gradient boosting models, covariate shift, model and variable selection, moving window evaluation

1 Introduction

Insurers set aside funds, the so-called reserve, for covering claims from past exposure years. This reserve is often split into a reserve for Incurred, But Not yet Reported (IBNR) claims and a reserve for Reported, But Not yet Settled (RBNS) claims. These separate reserves differ in the range of statistical tools that are available for modelling them. Since the claims that compose the IBNR reserve are not yet reported, claim-specific and policy(holder) covariates are unavailable for differentiating the cost per claim. Therefore, IBNR reserving mostly focuses on accurately estimating the number of unreported claims, followed by allocating a fixed cost per unreported claim. In RBNS reserving, the insurer is aware of the number of open claims as well as their characteristics and development so-far. This opens the possibility for reserving models that predict the future cost on a per claim basis. This paper focuses on predicting the RBNS reserve by modelling the development of reported claims.

Traditionally, the non-life insurance literature has been dominated by analytic models designed for aggregated data, such as the chain ladder method (Mack, 1993, 1999). These models compress the historical data on the development of claims over time in a two dimensional table, the so-called runoff triangle, by aggregating payments by occurrence and development year. Low data requirements, implementation simplicity and a straightforward interpretation of the predicted reserve justify the popularity of these models. However, by compressing the data valuable insights into the risk characteristics of individual claims are lost. This makes the reserve less robust against changes in the portfolio composition and extreme one-time events. In response to this, individual reserving methods designed for granular data available at the level of individual claims, have first been proposed in the nineties. Individual reserving remained largely unexplored for about two decades, with revived interest in recent years thanks to an increased focus on big data analytics.

We identify three streams in the current literature on individual reserving. Following Norberg (1993, 1999), a first stream analyzes the events registered during a claim's development in continuous time. Lopez et al. (2016, 2019) adapt regression trees to the right-censoring present in continuous time reserving data. Covariates in these trees capture the heterogeneity in the claim size as well as in the time to settlement of reported claims. In Antonio and Plat (2014) hazard rates drive the time to events in the development of claims (e.g. a payment, or settlement) and a lognormal regression model is proposed for the payment size. Reserving in continuous time requires a time to event model that allows for multiple payments and multiple types of (recurrent) events. Since such models are complicated, many individual reserving models are defined in a more convenient discrete time framework, where the events in a claim's lifetime are registered in discrete time periods. A second stream of reserving methods models the reserve in discrete time by adapting models from insurance pricing, as such taking advantage of the detailed covariate information available within insurance companies. Since these covariates only become available at reporting, such models focus on the reserve for reported, but not settled (RBNS) claims, while using techniques from aggregated reserving to estimate the reserve for unreported claims. Larsen (2007) focuses on Generalized Linear Models (GLMs), Wüthrich (2018) considers regression trees and Wüthrich (2018) looks at neural networks for reserving. A third stream of papers aggregates the data into multiple runoff triangles. Martínez Miranda et al. (2012), Wahl et al. (2019) and Denuit and Trufin (2017, 2018) consider two, three and four triangles respectively. While the aggregation of the data makes these models easy to implement, covariate information of individual claims can not be used. The recent expansion in (individual) reserving methodology has resulted in a fragmented literature with few comparative studies and no unified approach with proven robustness and general applicability. The lack of a solid modelling framework hinders the implementation of individual reserving in-

surance practice. Moreover, providing data driven guidance on the choice between aggregate and individual reserving is a very important question that is largely unexplored in the reserving literature.

We aim to fill this gap in the literature by presenting the hierarchical reserving model as an intuitive framework for RBNS reserving with a focus on applicability in practice. This framework decomposes the joint likelihood of the development process of individual claims after reporting in discrete time. Hierarchical reserving models are tailored to the portfolio at hand by adding layers, which represent the events (e.g. settlement, positive or negative payments, changes in the incurred, ...) registered over the lifetime of a claim. This modular approach enables us to restate many existing reserving models, including models based on data aggregated into a runoff triangle, as special cases of the hierarchical reserving model. This intuitive model building process allows us to concentrate on the decisions made during the modelling process, such as model calibration and evaluation. These aspects of the modelling process have received little attention in individual reserving literature up to now with many papers following the model building steps applied in pricing or aggregate reserving. This results in a loss of performance as such methods do not consider the individual and censored structure of the data.

This paper is organized as follows. Section 2 introduces the hierarchical reserving model, presents best practices for calibrating this model to insurance data and explains how this model can be used to predict the future reserve. Section 3 investigates the connection between hierarchical reserving models proposed at individual claim level and certain aggregate reserving models. This results in a data driven strategy for choosing between aggregate and individual reserving. Section 4 demonstrates this methodology in a case study on a home insurance data set. This is a novel data set, which has not been used before in the literature on reserving. An R package accompanies the paper enabling researchers and practitioners to directly apply the hierarchical reserving model to their portfolios.

2 A hierarchical reserving model

It is common in insurance pricing to decompose the joint likelihood into a frequency and severity contribution (Henckaerts et al., 2018). Frees and Valdez (2008) extend this idea by splitting the severity contribution per claim type. In this spirit, we propose a hierarchical reserving model, which decomposes the joint likelihood of the claim development process over time and registered events (e.g. settlement, payment).

2.1 Notation and statistical model

We record the development of reported claims in discrete time over a period of τ years. For each reported claim k , r_k denotes the reporting year and the vector \mathbf{x}_k denotes the claim information available at the end of the reporting year. This information vector is static and consists of the circumstances of the claim, policy(holder) covariates and the claim development (e.g. initial reserve, payments) in the reporting year. In the years after reporting, so-called update vectors, denoted \mathbf{U}_k^j , describe the change in the claim development information in year $r_k + j - 1$. The length and components of \mathbf{U}_k^j depend on the events (e.g. claim settlement, change in the incurred, involvement of a lawyer) registered in the portfolio at hand. This paper defines a modular model building approach that can be tailored to the chosen structure in \mathbf{U}_k^j . These models, called hierarchical reserving models, are based on two fundamental assumptions

Hierarchical model assumptions

- (A1) All claims settle within d years after reporting;
- (A2) The development of a claim is independent of the development of the other claims in the portfolio.

Although upper limit d on the the settlement delay is not necessarily limited to the length of the observation window, we implicitly assume $d = \tau$ for notational convenience. Given these assumptions, the set of observed claim updates after reporting, say \mathcal{R}^{obs} , for a portfolio of n claims is

$$\mathcal{R}^{\text{obs}} = \{\mathbf{U}_k^j \mid k = 1, \dots, n, j = 2, \dots, \tau_k\},$$

with $\tau_k = \min(d, \tau - r_k + 1)$ the number of observed development years since reporting for claim k . The associated likelihood is

$$\mathcal{L}(\mathcal{R}^{\text{obs}}) = \prod_{k=1}^n f(\mathbf{U}_k^2, \dots, \mathbf{U}_k^{\tau_k} \mid \mathbf{x}_k),$$

where we use assumption (A2) to write the likelihood as a product of claim-specific likelihood contributions. Inspired by [Frees and Valdez \(2008\)](#), we introduce a hierarchical structure in this likelihood by applying the law of conditional probability twice. First, we include the temporal dimension by splitting the likelihood in chronological order

$$\mathcal{L}(\mathcal{R}^{\text{obs}}) = \prod_{k=1}^n \prod_{j=2}^{\tau_k} f(\mathbf{U}_k^j \mid \mathbf{U}_k^2, \dots, \mathbf{U}_k^{j-1}, \mathbf{x}_k).$$

By conditioning on past events, we acknowledge that the future development of a claim depends on its development in previous years. Second, we split the likelihood by the events registered in the vector \mathbf{U}_k^j

$$\mathcal{L}(\mathcal{R}^{\text{obs}}) = \prod_{k=1}^n \prod_{j=2}^{\tau_k} \prod_{l=1}^s f(\mathbf{U}_{k,l}^j \mid \mathbf{U}_k^{(2)}, \dots, \mathbf{U}_k^{j-1}, \mathbf{U}_{k,1}^j, \dots, \mathbf{U}_{k,l-1}^j, \mathbf{x}_k), \quad (1)$$

where s is the length of the update vector \mathbf{U}_k^j . In the remainder of this paper, we refer to these events registered over the lifetime of a claim as the layers of the hierarchical model. The order of the layers is an important model choice, since the outcome of a layer becomes a covariate when modelling higher indexed layers. Since the assumptions (A1-A2) are common in reserving literature, most discrete time reserving models can be seen as a special case of our hierarchical reserving framework. Notice that in contrast with the chain ladder method, the hierarchical framework includes the full history of the claim and thus allows for non-Markovian models.

When applying the hierarchical claim development model to a specific portfolio, we extend assumptions (A1-A2) with an additional assumption, which tailors the structure of the update vector \mathbf{U}_k^j to the portfolio at hand. For example, in the case study covered in [Section 4](#), we model \mathbf{U}_k^j with a three-layer hierarchical model.

Hierarchical layers

- (A3) The update vector \mathbf{U}_k^j for claim k in development year j has three layers
- $$\mathbf{U}_k^j = (\mathbf{U}_{k,1}^j, \mathbf{U}_{k,2}^j, \mathbf{U}_{k,3}^j) = (\mathbf{C}_k^j, \mathbf{P}_k^j, \mathbf{Y}_k^j):$$

- C_k^j is the settlement indicator which is one when claim k settles in development year j and zero otherwise. Conditional on past events, the settlement indicator follows a Bernoulli distribution with

$$C_k^j | \mathbf{U}_k^2, \dots, \mathbf{U}_k^{j-1}, \mathbf{x}_k \sim \text{Bernoulli} \left(p \left(\mathbf{U}_k^2, \dots, \mathbf{U}_k^{j-1}, \mathbf{x}_k \right) \right).$$

- P_k^j is the payment indicator which is one when there is a payment for claim k in development year j and zero otherwise. Conditional on past events, the payment indicator follows a Bernoulli distribution with

$$P_k^j | \mathbf{U}_k^2, \dots, \mathbf{U}_k^{j-1}, C_k^j, \mathbf{x}_k \sim \text{Bernoulli} \left(q \left(\mathbf{U}_k^2, \dots, \mathbf{U}_k^{j-1}, C_k^j, \mathbf{x}_k \right) \right).$$

- Y_k^j is the payment size, given that there was a payment in development year j . Conditional on past events, the payment size is gamma distributed with mean

$$E(Y_k^j | \mathbf{U}_k^2, \dots, \mathbf{U}_k^{j-1}, C_k^j, P_k^j, \mathbf{x}_k, j) = \mu(\mathbf{U}_k^2, \dots, \mathbf{U}_k^{j-1}, C_k^j, P_k^j, \mathbf{x}_k)$$

and variance

$$\sigma^2(Y_k^j | \mathbf{U}_k^2, \dots, \mathbf{U}_k^{j-1}, C_k^j, P_k^j, \mathbf{x}_k) = \theta \cdot \mu(\mathbf{U}_k^2, \dots, \mathbf{U}_k^{j-1}, C_k^j, P_k^j, \mathbf{x}_k).$$

As such, we structure the development of claims with a simple three-layer hierarchical model. Conditioning on the settlement status in past years, allows us to train the model on the development of open claims only, whereas without settlement indicator, the model would predict new payments for already settled claims. Moreover, by choosing settlement as the first layer of the hierarchical model, settlement becomes a covariate when modelling later layers. The gamma distribution for the sizes is frequently used in insurance pricing literature when modelling attritional losses (Henckaerts et al., 2020). Choosing a strictly positive distribution assumes that there are no recoveries in the portfolio. In portfolios in which recoveries are common, additional layers should be added to the hierarchical model to allow for negative payments.

2.2 Hierarchical model calibration

The hierarchical claim development framework makes no assumption with respect to the statistical modelling technique that is used to model the individual layers. The case-study in Section 4 illustrates the proposed hierarchical reserving model by calibrating both a Generalized Linear Model (GLM) as well as a Gradient Boosting Model (GBM) to the layers outlined in (A3). Although standard procedures are available for calibrating these models, special attention is required for the variable selection process or (hyper) parameter turning steps. In reserving the historical, observed data contains mainly records from early development years, whereas the future predictions are more oriented towards later years. This imbalance between the training and prediction data set poses a model risk when covariates exhibit a different effect on the first development years versus the later development years. In machine learning literature this phenomenon is known as a covariate shift (Sugiyama et al., 2007b). Following Sugiyama et al. (2007a), we correct for a potential covariate shift by maximizing a weighted likelihood in which weights depend on the development year, i.e.

$$\mathcal{L}^{\text{weighted}}(\mathcal{R}^{\text{obs}}) = \prod_{k=1}^n \prod_{j=2}^{\tau_k} w_j \prod_{l=1}^s f \left(U_{k,l}^j | U_k^{(2)}, \dots, U_k^{(j-1)}, U_{k,1}^j, \dots, U_{k,l-1}^j, \mathbf{x}_k \right), \quad (2)$$

where w_j is the weight assigned to an observation from development year j . Following Sugiyama et al. (2007a), we define these weights as the ratio of the number of records from development year j in the prediction data set to the number of records from development year j in the training data set. For typical reserving data sets this ratio is observed and can be computed as

$$w_j = \frac{\sum_{i=d-j+2}^d n_i}{\sum_{i=1}^{d-j+1} n_i},$$

where n_i is the number of reported claims in reporting year i . These weights increase in j assigning more weight to observations from later development years. When selecting covariates or tuning (hyper) parameters, we maximize (2) in a 5-fold cross validation scheme. For this, we calibrate predictive models per layer l and allocate observations at the level of a claim k and a development year j (see (2)) to different folds.

2.3 Predicting the future development of claims

Algorithm 1 simulates the development of reported claims beyond the observation window τ . In line with the hierarchical structure of the model, development years are simulated in chronological order and within a development year, this simulation algorithm respects the order of the layers. The simulation order is important, since simulated values from previous development years and lower indexed layers become inputs for later development years and higher indexed layers.

Algorithm 1: Simulating the future development of reported claims

Input: the observed development of reported claims

Output: simulation of the future development of reported claims

foreach claim k **do**

for development year j **in** $\tau + 1 - r_k \dots d$ **do**

for hierarchical layer l **in** $1 \dots s$ **do**

 Simulate $U_l^j \mid \mathbf{U}_k^{(2)}, \dots, \mathbf{U}_k^{(j-1)}, U_{k,1}^j, \dots, U_{k,l-1}^j, \mathbf{x}_k$

end

end

end

Following this algorithm, the simulated data has the same hierarchical layered structure as the input data set, which enables us to derive aggregated quantities for the events registered in the update vector \mathbf{U}_k^j . For example, given the specific hierarchical structure in assumption (A3), we obtain estimates for the number of open claims, the number of payments and the total payment size. Prediction intervals for these reserving quantities are obtained by running Algorithm 1 many times.

2.4 Implementation in R

We have developed a package called `hirem` (Crevecoeur, 2020) for defining and calibrating hierarchical reserving models as well as simulating the future development of claims. In this package, layers can be estimated with generalized linear models (GLMs) or gradient boosting models (GBMs). The case-study of Section 4 uses the implementation from Southworth (2015) of the `gbm` package, which adds the gamma loss function to the original package developed by Greenwell et al. (2018).

3 Bridging aggregate and individual reserving

Most claim reserving models used in insurance companies are based on data aggregated into runoff triangles. We start from data registered at the level of individual claims and illustrate how aggregate reserving models can be retrieved as special cases of the hierarchical reserving model. Section 3.1 investigates the simplified case of a hierarchical model with independent layers. Section 3.2 extends these results and allows a simple, but common dependency structure between the layers. The results of these sections offer valuable insights and statistical tools for choosing between aggregate and individual reserving. Section 3.3 demonstrates the universality of our framework by constructing hierarchical reserving models inspired by recent literature contributions on aggregate reserving with multiple runoff triangles. The hierarchical reserving model as a unified framework for RBNS reserving facilitates model comparison and offers new insights as to how these models could be extended to data registered at the level of individual claims.

3.1 From individual hierarchical reserving models with independent layers to aggregate reserving models

In contrast with traditional reserving models, the hierarchical reserving model proposed in Section 2.1 analyses the development of claims from development year two since reporting onwards. For this, our approach collects all information registered during the reporting year of a claim k in a vector \mathbf{x}_k . This vector not only includes claim covariates (e.g. the cause of the accident), but also covariates structuring the development of the claim in the reporting year (e.g. the amount paid during the reporting year). In this section, we denote by \mathbf{U}_k^1 the claim development information that becomes available during the reporting year. Next to this, \mathbf{x}_k refers in this section to the remaining static claim covariates that become available at reporting. Introducing \mathbf{U}_k^1 brings our notation more in line with traditional reserving practice and enables us to model the development of claims from development year one onwards.

We construct for each of the layers l in the update vector \mathbf{U}_k^j a runoff triangle $(X_l^{ij})_{1 \leq i, j \leq d}$ with cells

$$X_l^{ij} = \sum_{k:r_k=i} U_{k,l}^j.$$

In line with our focus on modelling RBNS claims, we aggregate by reporting year and development year since reporting instead of the traditional set-up where aggregation goes per occurrence year and development year since occurrence. As such, we model the development of claims since reporting. Although, we denote the reporting year of claim k by r_k , we keep the traditional index i for the rows, i.e. the reporting years, in the runoff triangle.

Let us now assume that the individual updates depend multiplicatively on the reporting year and the development year, i.e.

$$\mathbb{E}(U_{k,l}^j) = \alpha_{r_k,l} \cdot \beta_{j,l} \quad \text{and} \quad \sum_{j=1}^d \beta_{j,l} = 1, \quad (3)$$

for each layer l and where $\alpha_{r_k,l}$ is the effect of reporting year r_k and $\beta_{j,l}$ is the effect of development year j . When we aggregate these individual updates into a runoff triangle, the cell values follow a similar multiplicative relation, i.e.

$$\mathbb{E}(X_l^{ij}) = \mathbb{E} \left(\sum_{k:r_k=i} U_{k,l}^j \right) = n_i \cdot \alpha_{i,l} \cdot \beta_{j,l} := \tilde{\alpha}_{i,l} \cdot \beta_{j,l}, \quad (4)$$

where n_i , the number of reported claims in reporting year i , is observed. As a result, we can calibrate individual hierarchical reserving models that only depend multiplicatively on reporting year and development year using data aggregated into runoff triangles.

Matching (3) with the original hierarchical reserving model specification in (1), we rephrase the expected value for the updates U_k^j at the individual level in full generality as

$$E(U_{k,l}^j) = \alpha_{r_k,l} \cdot \beta_{j,l} \cdot \phi \left(U_k^1, \dots, U_k^{j-1}, U_{k,1}^j, \dots, U_{k,l-1}^j, \mathbf{x}_k \right),$$

where $\phi(\cdot)$ represents the effect of all other covariates. When we add a distributional assumption for U_k^j , choosing between an aggregate or individual reserving model reduces to testing for $\phi(\cdot) = 1$. Since the models with and without $\phi(\cdot)$ are nested, a likelihood ratio test can be used for this.

3.2 From individual hierarchical reserving models with dependent layers to aggregate reserving models

The reserving models constructed in Section 3.2 treat each layer independent of the others. This results in simple aggregated models, where each layer is estimated from a single runoff triangle, independent from the other layers. However, in most multi-layer hierarchical structures some dependence between the layers is inevitable and offering a simple framework to include these dependencies is one of the main motivations for the hierarchical reserving model. This section investigates the special, but common setting of a two-layer hierarchical model in which layer one is a binary random variable and layer two is zero whenever layer one equals zero. As an example, think of layer one as a payment indicator and layer two as the payment size. When there is no payment, the payment size is zero.

Again focusing on the multiplicative structure of reporting and development year, we structure the expected values of the layers as

$$\begin{aligned} E(U_{k,1}^j) &= \alpha_{r_k,1} \cdot \beta_{j,1} \cdot \phi \left(U_k^1, \dots, U_k^{j-1}, \mathbf{x}_k \right) \\ E(U_{k,2}^j) &= \begin{cases} \alpha_{r_k,2} \cdot \beta_{j,2} \cdot \psi \left(U_k^1, \dots, U_k^{j-1}, \mathbf{x}_k \right) & U_{k,1}^j = 1 \\ 0 & U_{k,1}^j = 0 \end{cases}, \end{aligned}$$

where $\sum_{j=1}^d \beta_{j,l} = 1$ for $l \in \{1, 2\}$. When $\phi(\cdot)$ and $\psi(\cdot)$ are both equal to one, the claim development depends only on reporting year and development year in a multiplicative way. We then retrieve

$$\begin{aligned} E(X_1^{ij}) &= E \left(\sum_{k:r_k=i} U_{k,1}^j \right) = n_i \cdot \alpha_{i,1} \cdot \beta_{j,1} := \tilde{\alpha}_{i,1} \cdot \beta_{j,1} \\ E(X_2^{ij} | X_1^{ij}) &= E \left(\sum_{k:r_k=i} U_{k,2}^j \right) = X_1^{ij} \cdot \alpha_{i,2} \cdot \beta_{j,2}. \end{aligned} \quad (5)$$

When calibrating the model for the second layer, the observed upper triangle of the first layer acts as an exposure term. When estimating the future reserve, this exposure term, X_1^{ij} , should be estimated using the model proposed for the first layer. If we interpret the first layer as a payment indicator and the second layer as the payment size then the number of payments becomes the exposure for the total payment size. Similar to Section 3.1, statistical tests for $\phi(\cdot) = \psi(\cdot) = 1$ offer data driven tools for choosing between individual and aggregate reserving.

3.3 Hierarchical reserving models inspired by aggregate reserving models proposed for multiple runoff triangles

As a result of the weak assumptions underlying the hierarchical reserving model, many existing reserving models can be restated as special cases of our framework. As a unifying framework, the hierarchical reserving model facilitates model comparison and allows extending the calibration and simulation strategy developed in this paper to other models. In the case of models designed for aggregate data, the hierarchical reserving framework in addition offers insights as to how these models could be extended to data registered at the level of individual claims.

As an illustration of the generality of our framework, we construct hierarchical reserving models inspired by recent contributions on aggregate reserving using data structured in multiple triangles. We discuss two examples of such models, namely the double chain ladder (Martínez Miranda et al., 2012) and the collective reserving model (Wahl et al., 2019). As motivated in Section 1, we limit our analysis to the RBNS part of these aggregate models.

Double chain ladder The double chain ladder (DCL) (Martínez Miranda et al., 2012) extends the chain ladder method to two runoff triangles to obtain separate estimates for the IBNR and RBNS reserve. Since we only consider the development of claims after reporting, we focus on the triangle of claim sizes and construct a one-layer hierarchical model. DCL structures the expected payment size for a claim k in development year j since reporting, denoted $U_{k,1}^j$, as

$$\mathbb{E}(U_{k,1}^j) = \tilde{\pi}_j \cdot \tilde{\mu}_j \cdot \gamma_{i_k},$$

where i_k denotes the occurrence year of the claim and $\tilde{\pi}_j$ and $\tilde{\mu}_j$ are the payment probability in development year j and average payment size in development year j respectively. The coefficient γ_{i_k} adjusts the size of the payments from occurrence year i_k for inflation. Letting inflation depend on the occurrence year is natural in DCL, which aggregates runoff triangles by occurrence year and development year since occurrence. Since runoff triangles based on the hierarchical reserving model aggregate by reporting year and development year since reporting, it is in our framework more natural to model inflation per reporting year. If we change the occurrence year effect γ_{i_k} by a reporting year effect γ_{r_k} , the individual updates become

$$\mathbb{E}(U_{k,1}^j) = \tilde{\pi}_j \cdot \tilde{\mu}_j \cdot \gamma_{r_k}.$$

Aggregating these individual updates, we retrieve

$$\mathbb{E}(X_1^{ij}) = \mathbb{E} \left(\sum_{k:r_k=i} U_{k,1}^j \right) = n_i \cdot \tilde{\pi}_j \cdot \tilde{\mu}_j \cdot \gamma_i.$$

This is the same model as (3), when we rewrite $\alpha_i = n_i \cdot \gamma_i$ and $\beta_j = \tilde{\pi}_j \cdot \tilde{\mu}_j$.

The collective reserving model Extending the earlier work of Verrall et al. (2010) and Martínez Miranda et al. (2012), the collective reserving model (Wahl et al., 2019) structures the claim development after reporting in two layers. These layers represent the number of payments and the size per payment. Inspired by Wahl et al. (2019)'s aggregate model, we structure the individual updates as

$$\begin{aligned} \mathbb{E}(U_{k,1}^j) &\sim \text{Poisson}(\lambda_j), && \text{(number of payments)} \\ \mathbb{E}(U_{k,2}^j | U_{k,1}^j) &= \mu(i_k, r_k, j) \cdot U_{k,1}^j, && \text{(payment size)} \end{aligned}$$

where a claim can have multiple payments in the same year, each with an average size $\mu(i_k, r_k, j)$, which depends on the occurrence year, reporting year and development year since reporting. When $\mu(i_k, r_k, j) = \alpha_{r_k} \cdot \beta_j$, the model aggregates to

$$\begin{aligned} E(X_1^{ij}) &= n_i \cdot \lambda_j, \\ E(X_2^{ij} | X_1^{ij}) &= X_1^{ij} \cdot \alpha_i \cdot \beta_j, \end{aligned}$$

where n_i denotes the number of claims reported in reporting year i . This representation is almost identical to (5), with the estimated effect of reporting year i for the number of payments replaced by the observed count n_i .

4 Case study: European home insurance portfolio

This case study models the RBNS reserve for a European home insurance portfolio. This insurance reimburses damages to the insured property and its contents resulting from a wide range of causes including fire damage, water damage and theft. For reasons of confidentiality we can not disclose the size of the portfolio and the associated reserve. Therefore, we express the performance of the investigated reserving methods via a percentage error measure, comparing the actual and predicted reserve.

4.1 Data characteristics

We observe the development of individual claims over a seven year period from January, 2011 until December, 2017. Figure 1 structures individual payments by the reporting date of the claim (vertical axis) and the number of days elapsed since reporting (horizontal axis). Every dot represents a single payment and one claim can have multiple payments. A triangular structure appears, since the claim development after December, 2017 is censored. Home insurance is a short tailed business line, with many payments in the first years after reporting. The black grid in Figure 1 visualizes how individual payments would be aggregated when constructing a yearly runoff triangle. As shown in this triangle, extreme weather events cause sudden spikes in the number of reported claims. This has a large impact on the stability of the runoff triangle in classical, aggregate reserving. Therefore, insurers most often reserve these claims separately based on expert opinion. In this paper we analyze the robustness of various hierarchical reserving methods by predicting the future reserve with and without extreme weather events. Table 1 provides a detailed description of the available covariates. We group these covariates into four categories. Policy covariates identify the policy or policyholder entering the claim. These covariates are available when pricing the contract. Claim covariates describe the static characteristics of the claim. These covariates become available in the reporting year of the claim. Development covariates describe the yearly evolution of the claim and layer covariates constitute the layers of the hierarchical reserving model.

In Figure 2 a treemap visualizes the available claims grouped into the 12 risk categories as coded in the covariate `catnat`. Each claim is represented by a rectangle, where the size of this rectangle visualizes the amount paid for that claim by the end of December, 2017. Water and fire damage are the most important insurance covers in this portfolio. Together these risks generate more than half of the total claim cost. Fire claims are typically larger than non-fire claims. Although less than 5% of all claims are related to fire, these claims represent more than 25% of the total cost. The large difference between the average size of fire claims versus the average size of non-fire claims, motivates us to build separate reserving models for fire claims on

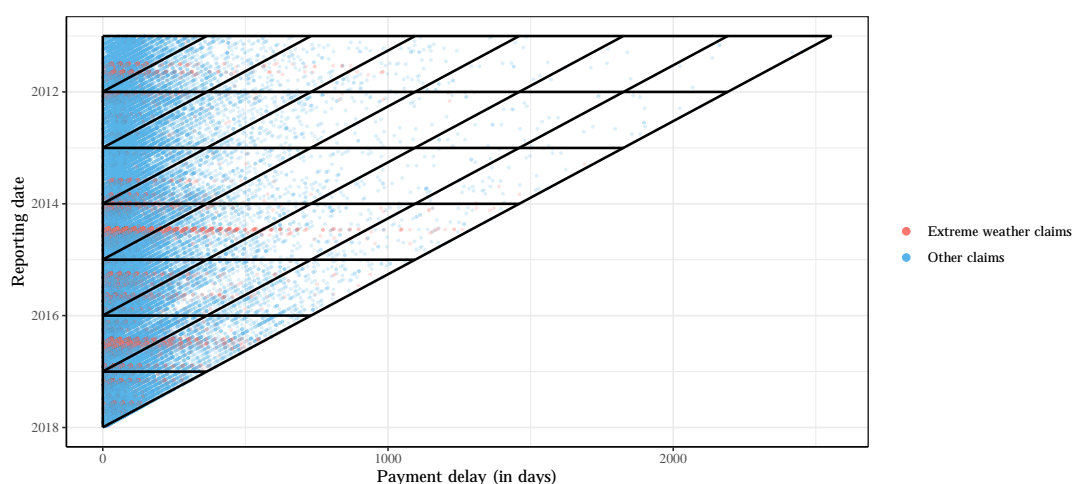


Figure 1: Payments structured by reporting date and payment delay in days. Every dot represents a single payment and one claim can have multiple payments. A grid indicates how individual payments would be aggregated when constructing a yearly runoff triangle. Claims resulting from extreme weather (e.g. a storm) are colored red.

Policy covariates

<code>valuables</code>	Objects were declared with a value exceeding the standard cover: yes or no
<code>age.insured</code>	The age of the policyholder
<code>profession</code>	Profession of the policyholder, 16 categories
<code>sex</code>	Gender of the policyholder: male or female
<code>construction.year</code>	The year in which the building was constructed
<code>property.value</code>	The value of the property in Euro

Claim covariates

<code>acc.date</code>	Date on which the accident occurred
<code>rep.date</code>	Date on which the claim was reported to the insurer
<code>rep.delay</code>	Delay in days between the occurrence and reporting of the claim
<code>rep.month</code>	Calendar month in which the claim was reported (Jan - Dec)
<code>coverage</code>	The main coverage applicable to the claim: theft, building or contents
<code>catnat</code>	The cause of the claim, grouped in 12 categories
<code>extreme.weather</code>	Claim is the result of extreme weather (e.g. storm): yes or no
<code>initial.reserve</code>	Expert estimate of the initial reserve at the end of the reporting year

Development covariates

<code>dev.year</code>	The number of years elapsed since the reporting of the claim
<code>calendar.year</code>	Number of years elapsed between the start of the portfolio and <code>dev.year</code>
<code>size.last.year</code>	Total amount paid in the previous development year
<code>total.amount.paid</code>	Total amount paid in all previous development years

Layer covariates

<code>close</code>	The claim closes in the current development year: yes or no
<code>payment</code>	A payment occurs in the current development year: yes or no
<code>size</code>	Total amount paid in the current development year

Table 1: List of covariates available in the home insurance data set. A level NA (not available) identifies the records with no registered value for a covariate.

the one hand and non-fire claims on the other hand. Estimating separate reserves for risks with

a different development pattern is a common approach in traditional reserving. Alternatively, we can distinguish fire and non-fire claims by including a covariate in the hierarchical reserving model. However, this latter approach would result in an unfair comparison between individual models, which can use this covariate, and traditional reserving methods for aggregate data, which can not use this covariate.

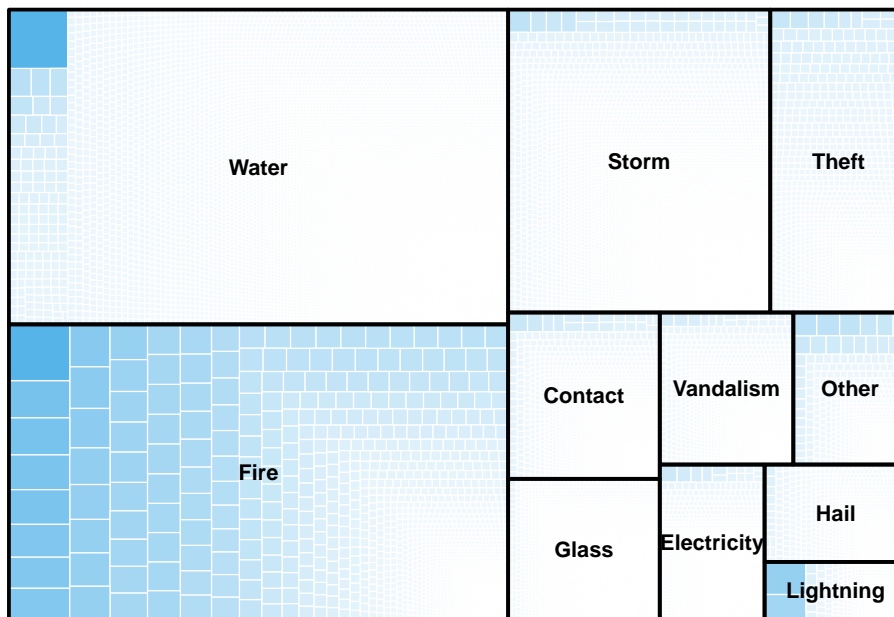


Figure 2: Treemap of individual claims observed on 31 December, 2017 grouped into the 12 risk categories present in the portfolio as coded in the covariate `catnat`. Each claim is represented by a rectangle, where the size of this rectangle visualizes the amount paid for that claim by the end of December, 2017.

4.2 Hierarchical reserving models for fire and non-fire claims

We analyse the performance of hierarchical reserving models based on GLMs, GBMs and the chain ladder method on 365 evaluation dates between January 1, 2015 and December 31, 2015. Instead of a single out-of-time evaluation (as e.g. in [Antonio and Plat \(2014\)](#); [Wüthrich \(2018\)](#)) the moving window evaluation enables a more thorough assessment of the sensitivity and general applicability of the model. On each evaluation date τ we train the models on the observed data (January, 2011 until τ) and compare the out-of-sample reserve estimate with the actual claim development over the next two development years.

4.2.1 Hierarchical reserving models

Hierarchical GLM The hierarchical GLM follows the three layer structure close, payment and size defined in assumption (A3) and models each of these layers with a Generalized Linear Model (GLM). Actuaries are familiar with GLMs, given the long tradition of using GLMs in insurance pricing and reserving. Therefore, GLMs are the most likely candidate for supporting the transition from aggregate to individual reserving in practice. As is common in insurance

pricing, we bin the continuous variables `age.insured`, `construction.year`, `property.value` and `rep.delay`. Table 2 shows the chosen bins for each covariate. We do not include the continuous development covariates `size.last.year` and `total.amount.paid` in the hierarchical GLM, since these covariates are highly correlated with the development year. On the first evaluation date, January 1, 2015, we select the optimal set of covariates for each of the three GLMs (`close`, `payment` and `size`) using forward selection with 5-fold cross validation, i.e. we iteratively add the covariate that results in the largest increase in the weighted likelihood (2) over all hold-out folds. In the moving window evaluation, we do not reselect the covariates on the other 364 evaluation dates, but recalibrate the parameters on each evaluation date using the most recent data.

Variable	Bins
<code>age.insured</code>	[0, 39], [40, 49], [50, 64], 65+, NA
<code>construction.year</code>	1950–, [1950, 1969], [1970, 1984], 1985+, NA
<code>property.value</code>	150,000–, (150,000, 200,000], (200,000, 250,000], 250,000+, NA
<code>rep.delay</code>	5–, [5, 21], 21+

Table 2: List of chosen bins for the continuous covariates in the hierarchical GLM.

Figure 3a shows the selected covariates in each GLM as well as a measure of the importance of each selected covariate. We compute covariate importance as the increase in the weighted likelihood (2) over all hold-out folds when sequentially adding covariates using forward selection. These increases are rescaled per GLM and sum to 100. For non-fire claims the set of selected covariates changes only slightly when we omit extreme weather events. This is in line with the low importance assigned to the covariate `extreme.weather` when these claims are included. The interaction `dev.year * rep.month` allows a more accurate determination of a claim’s age, while still reserving in a yearly framework. This is by far the most important determinant for the settlement and payment process of non-fire claims. The importance of development year as a covariate for individual reserving is a strong validation for aggregate reserving models, which cannot use other covariates. Surprisingly, `dev.year` and `dev.year * rep.month` have little effect on the size of non-fire claims. The most important determinants for the payment size are the claim type as coded in `catnat` and the `initial.reserve` set by the expert. The data set contains less fire claims and as a result fewer covariates are selected in the corresponding GLMs. Although these GLMs might be less predictive, the few selected covariates obtain high importance scores, since scores are always scaled to 100. In particular, `valuables` has an importance of 100, since it is the only covariate selected in the GLM for the settlement of fire claims. As with non-fire claims the `initial.reserve` is an important predictor for the payment size.

Hierarchical GBM The hierarchical GBM follows the same three layer structure as the hierarchical GLM, but models each layer with a tree based Gradient Boosting Model (GBM). GBMs, as introduced by Friedman (2001), model the data with a sequence of shallow decision trees, in which each tree improves the fit of the previous trees. The GBM has three major advantages. First, through a sequence of trees a non-linear effect can be estimated for continuous covariates, thus removing the need to bin continuous variables. Second, automatic feature selection is integrated in the calibration process. Third, simple interaction effects between the covariates are automatically modelled. As a result of these advantages, the covariates `age.insured`, `construction.year`, `property.value`, `rep.delay` and `initial.reserve` can be included as continuous covariates. Furthermore, we do not include the interaction `dev.year`

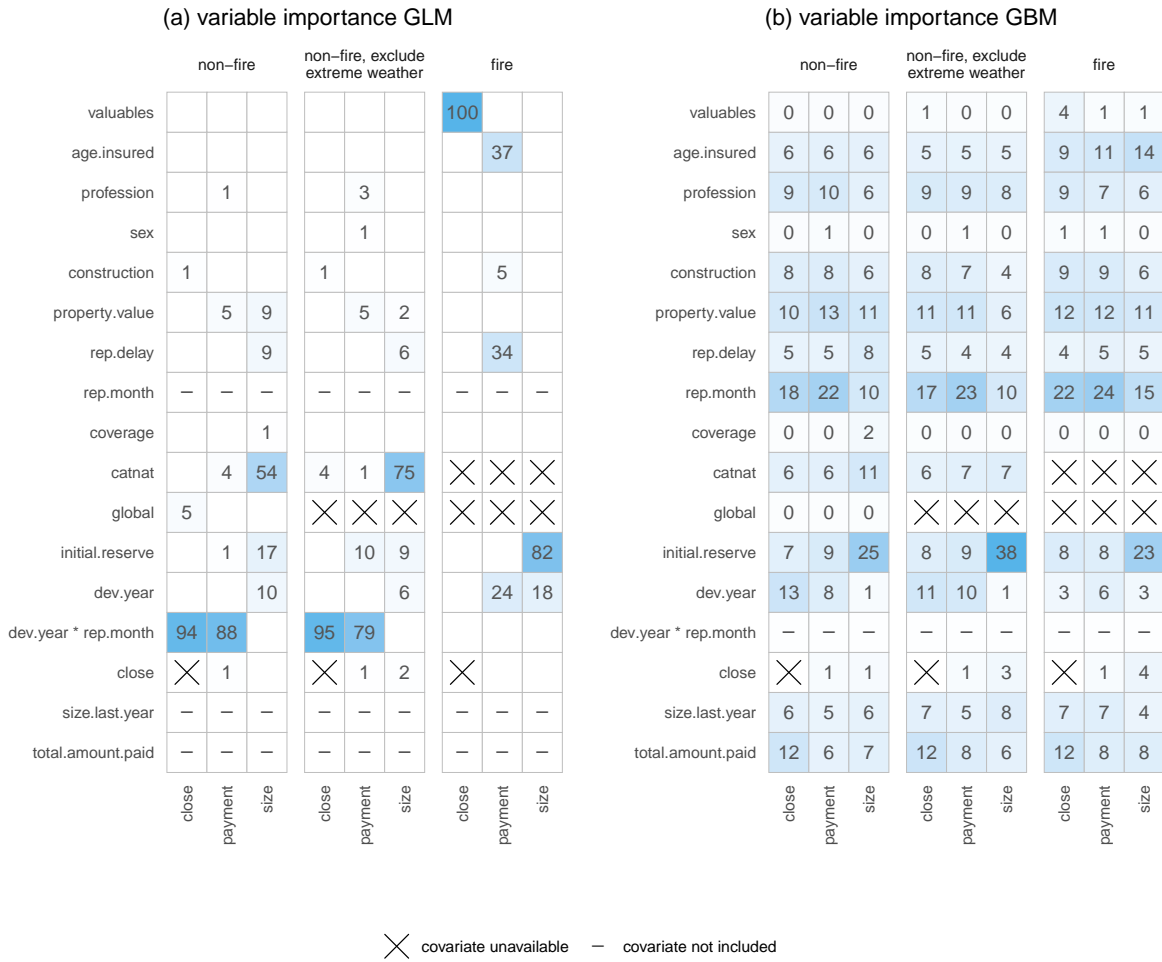


Figure 3: Relative importance of the selected covariates in (a) the hierarchical GLM and (b) the hierarchical GBM. Relative importance is computed as the increase in likelihood attributed to a single covariate relative to the total increase in likelihood caused by all covariates.

* `rep.month` as the model will automatically construct the relevant interactions. In return, a number of tuning parameters such as the number of trees and the depth of each tree have to be tuned. We tune these parameters on January 1, 2015 using the cross validation strategy of Section 2.2. Once tuned, these parameters remain fixed throughout the 364 remaining evaluation dates.

Figure 3b shows the relative importance of the covariates in the various GBMs. The importance of a specific covariate is expressed as the total improvement of the loss function over all splits including that covariate averaged over the 365 evaluation dates and scaled to 100. Since there is no explicit variable selection, importance is distributed over all covariates, which complicates the interpretation. For recent claims, `rep.month` allows for a more granular expression of the time elapsed since reporting the claim, which is important when modelling the target variables close, payment and size. `initial.reserve` is the most important covariate when predicting the size of payments. This shows that claim experts base their reserve estimate on information of the claim beyond the covariates available in model building. Similarly, the importance of other covariates shows that the practice of determining an initial reserve can be further improved

by using a statistical model. The claim type `catnat`, which was important in the hierarchical GLM, is less important in the GBM.

Chain ladder method We compare the previous individual hierarchical reserving models with the classical chain ladder method based on yearly aggregated data. As indicated in Section 3.1, the chain ladder method can be rephrased as a hierarchical reserving model with a single layer, i.e. the payment size. On each evaluation date, we compute the RBNS reserve by applying the classical chain ladder method to a runoff triangle of payment sizes aggregated by reporting and development year. The choice for aggregating by reporting year results in an estimate for the RBNS reserve, as motivated in Section 3. Confidence bounds for the reserve estimate will be derived from a normal assumption combined with the standard errors under the Mack model (Mack, 1999).

4.2.2 Evaluation of the RBNS reserve

On each evaluation date we predict the expected RBNS reserve for the open claims over the next two years. We measure model performance via the percentage error of the predicted reserve compared to the actual reserve, that is

$$\text{percentage error} = \frac{\text{predicted} - \text{actual}}{\text{actual}} \cdot 100\%.$$

Figure 4 shows the evolution of the percentage error between January 2015 and December 2015 as obtained with the three hierarchical reserving models. The percentage error is capped at 100% for improved readability of the figures. Table 3 summarizes the daily errors by calculating the average percentage error and the average absolute percentage error over the 365 evaluation dates.

The reserve for non-fire claims (Figure 4a) combines the outstanding amounts on many small claims, which provides a sufficiently rich data set for accurately training the individual hierarchical reserving models. This results in a similar performance for the hierarchical GLM and GBM. Extreme weather events in past years, produce outliers in the cells in the runoff triangle. This has a large impact on the chain ladder method, which fails to provide reasonable reserve estimates. This is a well known weakness of the chain ladder method and it is interesting that the individual models do not have this weakness, since they scale the reserve estimate automatically with the number of claims. Furthermore, Figure 3 shows that the covariate `extreme.weather` is rarely selected in the hierarchical models, which indicates that the development of these extreme weather claims does not fundamentally differ from regular claims. Not having to separate these extreme weather events from the other claims is a major advantage of individual reserving. When we remove extreme weather events (Figure 4b), performance across all three models becomes relatively competitive. Both the chain ladder method and the individual hierarchical models benefit from a data set with a large number of claims. Table 3 shows that performance is slightly better for the individual models, which surprisingly perform even better when we would not exclude extreme weather events from the data set. We observe higher prediction errors for all three models when predicting the reserve for fire claims (Figure 4c). The combination of a low claim frequency and potentially high costs makes the reserve for fire claims difficult to predict.

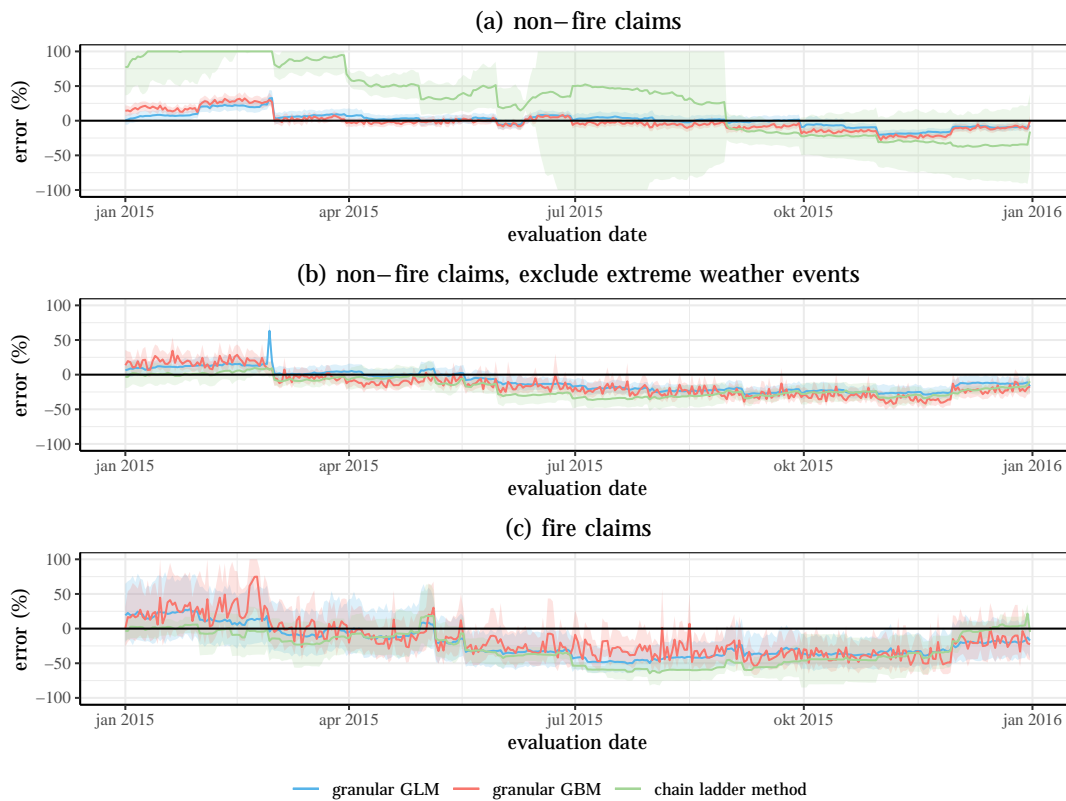


Figure 4: Percentage error in the prediction of the RBNS reserve on evaluation dates between January 1, 2015 and December 31, 2015 under the hierarchical GLM, hierarchical GBM and the chain ladder method. Errors are limited to 100%. (a) shows the reserve for non-fire claims, (b) the reserve for non-fire claims, when extreme weather events are excluded and (c) the reserve for fire claims.

5 Conclusion

We propose the hierarchical reserving model as a general framework for RBNS reserving in discrete time. By adding layers and choosing predictive models this framework can be tailored to any insurance portfolio. At the same time, our approach enables the development of best practices for calibration (see Section 2.2) and offers statistical tools for comparing hierarchical models. Model comparison extends to many existing reserving models, which can be restated as hierarchical reserving models. Moreover, Section 3 presents a connection with aggregate reserving models, allowing a data driven choice between aggregate and individual reserving. We illustrate our framework on a detailed case study with a home insurance data set. The flexibility of the framework is demonstrated by calibrating the same three layer structure with generalized linear models and gradient boosting models. As a best practice, we minimize the effect of day-to-day volatility when comparing our reserving models, by evaluating the performance over 365 evaluation days. The individual hierarchical models consistently outperform the classical chain ladder method on aggregated data and have the additional benefit that extreme weather events do not have to be removed prior to reserving.

Portfolio	hierarchical GLM		hierarchical GBM		chain ladder	
	$\mu(PE)$	$\mu(PE)$	$\mu(PE)$	$\mu(PE)$	$\mu(PE)$	$\mu(PE)$
non-fire claims	0.92	7.32	-1.80	10.23	33.89	51.31
non-fire claims, exclude extreme weather	-9.76	14.90	-14.28	20.18	-18.10	19.07
fire-claims	-20.82	26.44	-16.42	26.50	-28.41	29.76

Table 3: Evaluation of the average performance of the hierarchical GLM, hierarchical GBM and chain ladder method over 365 evaluation dates between January 1, 2015 and December 31, 2015. Average performance is expressed as the mean percentage error and the mean absolute percentage error.

6 Acknowledgements

The authors thank the anonymous referees for useful comments which led to significant improvements of the paper. This work was supported by the Argenta Research chair at KU Leuven; KU Leuven’s research council [project COMPACT C24/15/001]; and Research Foundation Flanders (FWO) [grant number 11G4619N]. We thank Roel Verbelen for his helpful comments at the start of this project.

References

- Katrien Antonio and Richard Plat. Micro-level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*, 7:649–669, 2014. URL <https://doi.org/10.1080/03461238.2012.755938>.
- Jonas Crevecoeur. *hirem: Hierarchical Reserving Models*, 2020. URL <https://github.com/jonascreevecoeur/hirem>. R package version 1.0.
- Michel Denuit and Julien Trufin. Beyond the Tweedie reserving model: The collective approach to loss development. *North American Actuarial Journal*, 21(4):611–619, 2017. URL <https://doi.org/10.1080/10920277.2017.1353428>.
- Michel Denuit and Julien Trufin. Collective loss reserving with two types of claims in motor third party liability insurance. *Journal of Computational and Applied Mathematics*, 335:168 – 184, 2018. URL <https://doi.org/10.1016/j.cam.2017.11.044>.
- Edward W. Frees and Emiliano A. Valdez. Hierarchical insurance claims modeling. *Journal of the American Statistical Association*, 103(484):1457–1469, 2008. URL <https://doi.org/10.1198/016214508000000823>.
- Jerome H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 29(5):1189–1232, 2001. URL <https://doi.org/10.1214/aos/1013203451>.
- Brandon Greenwell, Bradley Boehmke, Jay Cunningham, and GBM Developers. *gbm: Generalized Boosted Regression Models*, 2018. URL <https://github.com/gbm-developers/gbm>. R package version 2.1.4.9000.
- Roel Henckaerts, Katrien Antonio, Maxime Clijsters, and Roel Verbelen. A data driven binning strategy for the construction of insurance tariff classes. *Scandinavian Actuarial Journal*, 2018 (8):681–705, 2018. URL <https://doi.org/10.1080/03461238.2018.1429300>.
- Roel Henckaerts, Marie-Pier Côté, Katrien Antonio, and Roel Verbelen. Boosting insights in insurance tariff plans with tree-based machine learning methods. *North American Actuarial Journal*, 2020. In press.

- C. R. Larsen. An individual claims reserving model. *ASTIN Bulletin*, 37(1):113 – 132, 2007.
- Olivier Lopez, Xavier Milhaud, and Pierre-E. Thérond. Tree-based censored regression with applications in insurance. *Electron. J. Statist.*, 10(2):2685–2716, 2016. URL <https://doi.org/10.1214/16-EJS1189>.
- Olivier Lopez, Xavier Milhaud, and Pierre-E. Thron. A tree-based algorithm adapted to microlevel reserving and long development claims. *ASTIN Bulletin*, 49(3):741762, 2019. URL <https://doi.org/10.1017/asb.2019.1>.
- Thomas Mack. Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, 23:213–225, 1993. URL <https://doi.org/10.2143/AST.23.2.2005092>.
- Thomas Mack. The standard error of chain ladder reserve estimates: Recursive calculation and inclusion of a tail factor. *ASTIN Bulletin*, 29:361–366, 1999. URL <https://doi.org/10.2143/AST.29.2.504622>.
- María Dolores Martínez Miranda, Jens Perch Nielsen, and Richard J. Verrall. Double Chain Ladder. *ASTIN Bulletin*, 42:59–76, 2012. URL <https://doi.org/10.2143/AST.42.1.2160712>.
- Ragnar Norberg. Prediction of outstanding liabilities in non-life insurance. *ASTIN Bulletin*, 23(1):95–115, 1993. URL <https://doi.org/10.2143/AST.23.1.2005103>.
- Ragnar Norberg. Prediction of outstanding liabilities II. Model variations and extensions. *ASTIN Bulletin*, 29(1):5–27, 1999. URL <https://doi.org/10.2143/AST.29.1.504603>.
- Harry Southworth. *gbm: Generalized Boosted Regression Models*, 2015. URL <https://github.com/harrysouthworth/gbm>. R package version 2.1.
- Masashi Sugiyama, Matthias Krauledat, and Klaus-Robert Müller. Covariate shift adaptation by importance weighted cross validation. *J. Mach. Learn. Res.*, 8:9851005, December 2007a. ISSN 1532-4435.
- Masashi Sugiyama, Shinichi Nakajima, Hisashi Kashima, Paul von Büna, and Motoaki Kawanabe. Direct importance estimation with model selection and its application to covariate shift adaptation. In *Proceedings of the 20th International Conference on Neural Information Processing Systems, NIPS’07*, pages 1433–1440, USA, 2007b. Curran Associates Inc. ISBN 978-1-60560-352-0. URL <http://dl.acm.org/citation.cfm?id=2981562.2981742>.
- Richard Verrall, Jens Perch Nielsen, and Anders Hedegaard Jessen. Prediction of RBNS and IBNR claims using claim amounts and claim counts. *ASTIN Bulletin*, 40(2):871887, 2010. URL <https://doi.org/10.2143/AST.40.2.2061139>.
- Felix Wahl, Mathias Lindholm, and Richard Verrall. The collective reserving model. *Insurance: Mathematics and Economics*, 87:34 – 50, 2019. URL <https://doi.org/10.1016/j.insmatheco.2019.04.003>.
- Mario V. Wüthrich. Machine learning in individual claims reserving. *Scandinavian Actuarial Journal*, 2018(6):465–480, 2018.
- Mario V. Wüthrich. Neural networks applied to chain–ladder reserving. *European Actuarial Journal*, 8(2):407–436, 2018. URL <https://doi.org/10.1007/s13385-018-0184-4>.