Comparison of Claim Reserves Methods Using Insurance Portfolio Generators

Petr Vejmělka[1] | *Charles University, Prague, Czech Republic*
Tomáš Cipra[2] | *Charles University, Prague, Czech Republic*

Received 22.4.2024 (revision received 29.7.2024), **Accepted** (reviewed) 26.8.2024, **Published** 13.12.2024

Abstract

Different reserving methods can be used to predict claim values in non-life insurance. This article compares two different methodological approaches to reserving methods, namely, Chain-ladder (the traditional approach to reserving in non-life insurance) and state-space modeling (the modern approach based on recursive Kalman filtering). Moreover, the paper compares both methods with the involvement of clustering which divides claims into several groups according to their similarity and ensures greater homogeneity of data. To be able to compare the accuracy of reserve predictions numerically one suggests three types of generators of large insurance portfolios that represent well the behavior of the given methods in practice (one of them is derived directly from a real Czech non-life insurance claims portfolio). The obtained results may serve as a hint to improve the state-space methodology in order to give comparable results with classical approaches to reserving since in future the state-space modeling will be important for micro reserving where the "clustering" gains nearly a form of individual policy contracts.

Keywords

Chain-ladder, claims portfolio generators, clustering, loss reserving, non-life insurance, state-space model

DOI

https://doi.org/10.54694/stat.2024.22

JEL code

C32, C53, G22

INTRODUCTION

Loss reserving is crucial for each insurance company since it is used to estimate funds for future claim payments and obtained results serve to ensure the financial stability of the insurer. In this paper, we deal with the estimation of reserves in non-life insurance. The article aims to provide a description of loss reserving and focuses on the comparison of two reserving methods and their application to insurance

[1] Department of Probability and Mathematical Statistics, Charles University, Sokolovská 83, 186 75 Prague 8, Czech Republic. Corresponding author: e-mail: vejmelp@karlin.mff.cuni.cz. ORCID: <https://orcid.org/0000-0002-6561-3663>.

[2] Department of Probability and Mathematical Statistics, Charles University, Sokolovská 83, 186 75 Prague 8, Czech Republic. ORCID: <https://orcid.org/0000-0003-0604-324X>.

claims portfolios. The Chain-ladder method, usually considered as a benchmark method in loss reserving and known for its effectiveness across diverse insurance scenarios, is compared to the log-normal model, which is a representative of state-space modeling and modern reserving methods. These reserving methods are supplemented by clustering method CLARA, introduced in Kaufman and Rousseeuw (1990), to assess the benefit of clustering for reserve accuracy (according to our experience, the method CLARA seems to be the most suitable in the context of reserving).

In order to dispose of a sufficient number of portfolios to which we could apply the considered methods and consequently compare particular approaches, we propose three different types of non-life insurance claims portfolio generators. For each generator and an adequately large sample of generated portfolios, the accuracy of the claims and reserve estimates is compared using different techniques, namely, comparing boxplots and further verifying the improvement of prediction accuracy using the paired sign test for equality of medians of reserves deviations. The section concerning the construction of insurance portfolio generators can be also useful for actuarial practice generally (not only for reserving).

This paper is structured as follows. Section 1 serves as a brief literature review. Section 2 provides an overview of the importance of loss reserving in non-life insurance, discusses two different approaches to reserving, explains the concept of clustering in the context of loss reserving and surveys the references in literature. Section 3 provides an overview of two distinct regulatory frameworks used in the insurance industry, Solvency II and IFRS 17, and introduces the concept of Claims Development Results (CDR). In Section 4, we present the generators used to create the insurance claims portfolios used for the numerical study. Section 5 presents the results of the numerical study and discusses the accuracy of estimates. Finally, the last section summarizes the conclusions achieved in this paper and suggests further research possibility.

1 LITERATURE REVIEW

As mentioned above, actuarial science is a field that has undergone significant evolution. This also includes reserving as its important part. Thus, there are many publications dealing with this field. The development of the reserving methods has progressed from basic deterministic approaches to sophisticated statistical and machine-learning techniques.

As examples of the deterministic methods that are based on the extrapolation of historical data, we can mention the Chain-Ladder method, see, e.g., Wüthrich and Merz (2008) including its stochastic model by Mack (1993), and Bornhuetter-Ferguson method presented in Bornhuetter and Ferguson (1972).

Due to the gradually increasing computational power and increased data availability, researchers could use more complex statistical methods. One of the representatives from this group is the generalized linear model that allowed to robustly model the relationship between claims data and influencing factors, see, e.g., England and Verrall (2002).

A different approach, based on state-space modeling, enables dynamic modeling of claims processes by incorporating both observed data and latent variables. Verall (1989) was a pioneering article introducing state-space models for claims reserving, presenting their ability to capture the complexity of claims development over time. Later articles, such as De Jong (2005) or Atherino et al. (2010), and recent advancements by, e.g., Costa and Pizzinga (2020) or Hendrych and Cipra (2021) further enhance the flexibility and predictive power of this approach.

The advent of machine learning has further transformed reserving methodologies. Techniques such as random forests and neural networks, presented, for example, in Wüthrich (2018), provide non-linear modeling capabilities. Recent studies by, e.g., Duval and Pigeon (2019), De Felice and Moriconi (2019) or Delong et al. (2020) can be used as an additional source of knowledge in the area of reserving.

Since there is an enormous volume of literature including internet reports dealing with reserving, in addition to the publications we have presented so far, we only list several other references dealing with this topic from various points of view: Balona and Richman (2020), Cipra (2010), England et al. (2018), Kaas et al. (2004), Munroe et al. (2018), Wüthrich and Merz (2008), The Actuarial Community (2022), Zhang (2010). Some of them will be referred later in the text to explain specific problems of reserving.

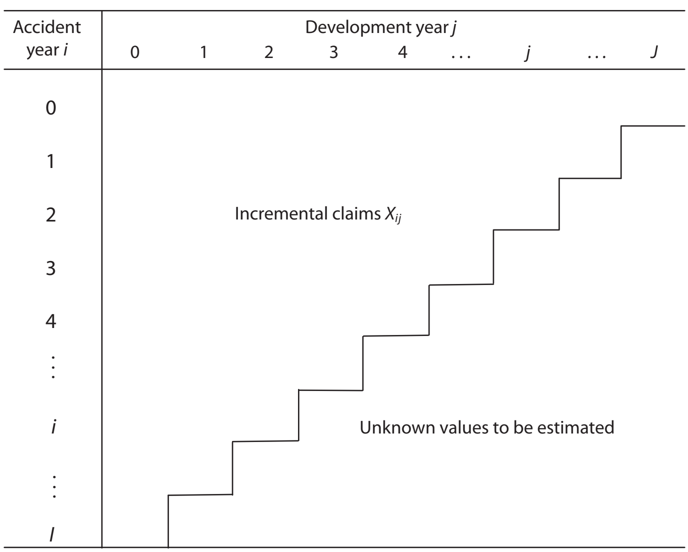
2 LOSS RESERVES IN NON-LIFE INSURANCE

Loss reserving is an important aspect of non-life insurance, which serves as a financial tool to estimate funds for future claim payments. Insurance companies may consider various methods to predict the ultimate value of claims that have been reported, but still not settled, and those that have not yet been reported. The main purpose of reserving is to ensure that insurance companies have appropriate funds to cover claim settlements, which is connected to the maintenance of financial stability and fulfilment of insurers' obligations to policyholders. One must steadily monitor and adjust the loss reserves because of the development of insurance portfolios caused by the reporting new information.

In practice, one can encounter different reserving methods, but these methods can be divided into two main groups of the so-called micro and macro reserving. The difference between these two types of reserving lies in the form of data, one is working with. The micro approach involves a detailed investigation of individual claims, with detailed information about each claim. On the other hand, the macro approach takes a broader perspective and individual claims are considered in an aggregated form. Nevertheless, both approaches are complementary and can be mutually combined.

In the case of macro reserving, one of the most used data representations is a run-off triangle, also called a development triangle. This triangle is a tabular representation of historical claims data over multiple periods. Typically, the claims are organized by accident years, which are represented by rows ($i = 0, ..., I$), and delays in reporting or payments captured in columns ($j = 0, ..., J$). Finally, the diagonals of the triangle represent particular calendar years. Based on the values appearing in the triangle, we can divide these triangles into incremental and cumulative ones. An illustration of a run-off triangle can be found in Figure 1. The actuaries can use diverse statistical techniques and mathematical models to analyze the patterns within the run-off triangle in order to complete the triangle into the rectangle by estimating the unknown future payments, the ultimate losses and the appropriate reserves.

Figure 1 Illustration of a run-off triangle



Source: Own construction based on Wüthrich and Merz (2008)

2.1 Classical approaches to loss reserving

In this article, we apply two different approaches to reserving. One of them, called Chain-ladder, can be classified as a classical approach to loss reserving, since it is a time-tested methodology proving its effectiveness across diverse insurance scenarios. Such classical methods are popular for their balance between simplicity and efficiency.

The Chain-ladder method, see e.g. Mack (1993), is one of the fundamental deterministic approaches utilized in non-life insurance reserving. It gained its success mainly due to its simplicity and easy applicability. It operates on the assumption that historical development patterns will persist into the future. Let $C_{i,j}$ represent the cumulative amount of claims occurred during accident year and paid till development year delay j. The chain-ladder estimates the unknown values with the use of so-called development factors $f_0, \ldots, f_{J-1}$, that connect cumulative values in the j-th and j + 1-th column and which are estimated as the ratios $\hat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}$. By extrapolating these development factors, one can estimate ultimate losses $\hat{C}_{i,J}$ for each accident year i as $C_{i,J-i} \cdot \hat{f}_{J-i} \cdot \ldots \cdot \hat{f}_{J-1}$. While the Chain-ladder method provides a practical and intuitive framework, its simplicity may cause inaccuracies of estimates in more complex insurance cases. For further details about the Chain-ladder, see Wüthrich and Merz (2008) or Cipra (2010).

The univariate Chain-ladder can be generalized to its multivariate version, where several run-off triangles are considered at once. Considering claims from more triangles allows for a detailed analysis of the impact of various covariates on the development patterns of claims. This approach can lead to improving the precision of loss reserving.

The multivariate Chain-ladder can be expressed within the SUR (Seemingly Unrelated Regression) framework, which is beneficial for the robustness of parameter estimates in situations when there are correlations in the error terms across equations. This model was introduced in Zhang (2010). When considering N development triangles, one works with vectors of cumulative claims $\boldsymbol{C}_{i,k} = \left(C_{i,k}^{(1)}, \ldots, C_{i,k}^{(N)}\right)'$ in the model $\boldsymbol{C}_{i,k+1} = \boldsymbol{B}_k \cdot \boldsymbol{C}_{i,k} + \boldsymbol{\varepsilon}_{i,k}$, where $\boldsymbol{B}_k$ is a development matrix of type $N \times N$ in development period k and $\boldsymbol{\varepsilon}_{i,k}$ is an N-dimensional random vector with several assumptions imposed on it. For these assumptions and more details on the multivariate Chain-ladder see Zhang (2010).

In addition to the models based on the Chain-ladder approach, there is also a number of similar models that work with development triangles, e.g., Bornhuetter–Ferguson, Benktander–Hovinen or Cape–Cod model (see, e.g., Wüthrich and Merz, 2008).

2.2 State-space models in loss reserving

A different approach to loss reserving that is considered in this article is so-called state-space modeling, where a linear state-space model plays a significant role in non-life insurance loss reserving. In this framework, one still works with aggregated claims ordered primarily to run-off triangles, but these triangles are transformed into the time series that are then modeled. The linear state-space models assume that the unobservable states and their observations follow linear relationships over time. Generally, such a model is given by the following system of equations:

$$\boldsymbol{y}_t = \boldsymbol{Z}_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \tag{1}$$

$$\boldsymbol{\alpha}_{t+1} = \boldsymbol{T}_t \boldsymbol{\alpha}_t + \boldsymbol{R}_t \boldsymbol{\eta}_t, \tag{2}$$

where $\boldsymbol{y}_t$ is a p-dimensional observation vector at time t, $\boldsymbol{\alpha}_t$ is an m-dimensional state vector at time t, $\boldsymbol{Z}_t$, $\boldsymbol{T}_t$, and $\boldsymbol{R}_t$ are matrices of parameters of types ($p \times m$), ($m \times m$) and ($m \times k$), respectively. Random vectors $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ are assumed to be normally distributed, where $\boldsymbol{\varepsilon}_t \sim N(\boldsymbol{0}, \boldsymbol{H}_t)$ is a p-dimensional random vector and $\boldsymbol{\eta}_t \sim N(\mathbf{0},\boldsymbol{Q}_t)$ is a k-dimensional random vector with covariance matrices $\boldsymbol{H}_t$ and $\boldsymbol{Q}_t$, respectively. Moreover, $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ are assumed to be independent. Since the state vectors are estimated by a recursive algorithm in time, it is also necessary to set an initial state $\boldsymbol{\alpha}_1$, fulfilling $\boldsymbol{\alpha}_1 \sim N(\boldsymbol{a}_1,\boldsymbol{P}_1)$, which is independent to $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ ($\boldsymbol{a}_1$ and $\boldsymbol{P}_1$ are some initial estimates).

This linear framework offers mathematical tractability and flexibility, allowing for efficient estimation and prediction using the Kalman filter (see Brockwell and Davis, 1991). The practical implementation can be realized by means of a selected software, e.g., using *KFAS* package available in software R and introduced in Helske (2017).

As mentioned above, run-off triangles need to be transformed to the form of time series. Then, they can enter the model as the observation vector $\boldsymbol{y}_t$**.** There are several ways how to order claims occurring in the development triangle. A row-wise ordering, resulting in time series with missing observations, was introduced in Atherino et al. (2010). Supposing normality of data is in this particular task vastly simplistic, thus, in order to meet the assumption of normality in the state-space model, Hendrych and Cipra (2021) introduced a log-normal model, assuming log-normal distribution of incremental claims. Hence, when working with logarithmically transformed incremental claims, one obtains the normal data.

The considered log-normal model is given by the following equations:

$$y_t(n) - y_t^0(n) = \alpha_t(n) + \varepsilon_t(n), \tag{3}$$

$$\alpha_{t+1}(n) = \alpha_{t-s+1}(n) + \eta_t(n), \tag{4}$$

where $y_t(n)$ and $y_t^0(n)$ represent the logarithmized incremental claims with appropriately adjusted index after row-wise ordering in the n-th run-off triangle for $n = 1, \ldots, N$. The values $y_t^0(n)$ correspond to the logarithmized values from the first column of the n-th run-off triangle. Their subtraction in (3) is considered as a setting of initial levels in the observation equation. For random variables $\varepsilon_t(n)$ and $\eta_t(n)$, we suppose $\varepsilon_t(n) \sim N(0, \sigma_\varepsilon(n,n))$ and $\eta_t(n) \sim N(0, \sigma_\eta(n,n))$, respectively.

Formulas (3) and (4) need to be transformed into their matrix forms that correspond to Formulas (1) and (2). Then it is possible to use appropriate software to estimate the model and subsequently forecast the missing observations in vector .

Based on Hendrych and Cipra (2021), the matrix form of the log-normal model is:

$$\boldsymbol{y}_t - \boldsymbol{y}_t^0 = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & 0 & \cdots & 0 \end{pmatrix} \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \tag{5}$$

$$\boldsymbol{\alpha}_{t+1} = \begin{pmatrix} \begin{matrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{matrix} & & \\ & \ddots & \\ & & \begin{matrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{matrix} \end{pmatrix} \boldsymbol{\alpha}_t + \begin{pmatrix} \begin{matrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{matrix} & & \\ & \ddots & \\ & & \begin{matrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{matrix} \end{pmatrix} \boldsymbol{\eta}_t, \tag{6}$$

where $\boldsymbol{y}_t = \left(y_t(1),\ldots,y_t(N)\right)'$, $\boldsymbol{y}_t^0 = \left(y_t^0(1),\ldots,y_t^0(1)\right)'$, $\boldsymbol{\varepsilon}_t = \left(\varepsilon_t(1),\ldots,\varepsilon_t(N)\right)'$,

$$\boldsymbol{\eta}_t = \left(\eta_t(1),0,\ldots,0,\ldots,\eta_t(N),0,\ldots,0\right)',$$

$$\boldsymbol{\alpha}_t = \left(\alpha_t(1),\ldots,\alpha_{t-s+1}(1),\ldots,\alpha_t(N),\ldots,\alpha_{t-s+1}(N)\right)'.$$

Moreover, residual vectors $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ have covariance matrices $Var(\boldsymbol{\varepsilon}_t) = \boldsymbol{H}_t = \left(\sigma_\varepsilon(m,n)\right)_{m,n=1,\ldots,N}$ and $Var(\boldsymbol{\eta}_t) = \boldsymbol{Q}_t = \left(\sigma_\eta(m,n)\right)_{m,n=1,\ldots,N}$, respectively.

After using Kalman smoothing and obtaining smoothed time series, one can transform the obtained values back to their original scale and proceed to the reserve calculation.

2.3 Preliminary clustering of loss data

Since in many cases it can be advantageous to work with more run-off triangles, rather than with single one, as mentioned above, one can deal with the problem of optimal distribution of claims to several groups that are then represented by individual development triangles. In some situations, the run-off triangles are naturally separated but, in many cases, there is only one portfolio of insurance claims with a potential to group it into more than one subportfolios. For this purpose, one can use various clustering methods.

Clustering corresponds to the grouping of similar elements based on their characteristics and identifies patterns and relationships within the data. In the context of loss reserving, one can assume that policies with comparable risk factors and exposure may have similar claim development behavior. By categorizing these policies into clusters with similar risk characteristics, insurers can take such a clustering into account in their loss reserving models to better capture the unique dynamics within each group.

In literature, there are numerous methods of unsupervised clustering that can be used in the context of insurance claims. In Vejmělka (2023), several methods, that are implemented in software R, have been compared. Namely, the function called *balanced_clustering* from the package *anticlust*, see Papenberg and Klau (2021), the function *Kmeans* from the *stats* package, proposed in Hartigan and Wong (1979), the function *Mclust* in the package *mclust*, see Scrucca et al. (2016), and the function *Clara_Medoids* in package *ClusterR,* introduced in Kaufman and Rousseeuw (1990).

Since in Vejmělka (2023) the CLARA (Clustering Large Applications) method provided the best results among the clustering methods involved in the comparison study, it is also preferred in the numerical study which is a part of this article. CLARA is an algorithm that extends the PAM (Partitioning Around Medoids) algorithm. PAM itself is an effective clustering method, however, it can be numerically complex in the case of large datasets due to its quadratic time complexity. This can be considerably problematic in the context of insurance data that can be quite extensive. The aim of CLARA is to overcome this limitation by performing the clustering on a subset of the data, which results in more computationally acceptable situation for larger datasets.

CLARA can be described as a three-step process. In the first step, multiple randomly chosen subsamples of the dataset are selected. The PAM algorithm is then applied to each subsample, medoids are calculated and observations from the subsamples are assigned to the appropriate clusters. Secondly, the obtained clusters are evaluated based on their overall stability. Finally, the most stable medoids and clusters are selected.

A completely different approach involves machine learning algorithms using neural networks. They can identify complex patterns within large datasets and due to the automatic process of identifying clusters, the need of manual interventions is reduced which can speed up the analysis. Nevertheless, complex deep learning models usually operate as black boxes, which results in very difficult or even impossible interpretation of the underlying decision-making processes. This lack of transparency can be a significant problem, especially in insurance, where interpretability is crucial for regulatory compliance. Several examples and references may be found, e. g., in Du (2010) or Kauffmann et al. (2022).

3 REGULATORY FRAMEWORK: CLAIMS DEVELOPMENT RESULTS

Solvency II and IFRS 17 represent two distinct regulatory frameworks used in the insurance industry, each with its own set of objectives and requirements. Solvency II, which has been established by the European Union, focuses primarily on ensuring the financial stability and solvency of insurance companies operating within the EU. It requires thorough evaluation of risks and sufficient capitalization to protect policyholders and keep market credible. In contrast, IFRS 17, an accounting standard developed by the International Accounting Standards Board, deals mainly with financial reporting and accounting standards for insurance contracts. It aims to enhance transparency and comparability of financial statements by requiring insurers to provide more detailed information about their insurance contracts.

The fundamental difference between Solvency II and IFRS 17 from a computational point of view subsists mainly in the fact that Solvency II considers risk over a one-year time horizon, whereas IFRS 17 is based on the fulfilment cash flows over their lifetime, see England et al. (2018). Reserves estimated using appropriate reserving methods can be used as one of the inputs to estimate future claims liabilities when determining the Best Estimate under IFRS 17. However, the Best Estimate may also incorporate additional modifying considerations and extensions.

In the Solvency II framework, it is fundamental to assess the insurer's ability to meet its obligations over a one-year horizon, as outlined in Pillar 1 (Quantitative Requirements). The methodology denoted as Claims Development Results (CDR) plays an important role in this process. It offers an insight into the development of insurance claims over time, specifically enabling insurers to project claims liabilities for the forthcoming year. It can be used for estimation of Solvency Capital Requirement (SCR), which is a key component of Solvency II. To estimate the SCR, a log-normal distribution with the mean equal to the expected ultimate loss and the standard deviation corresponding to the standard deviation of the CDR, is often applied. Then the -th percentile of this distribution can be used as the estimate of the SCR. One can find a more detailed information, e.g., in England et al. (2018) or Munroe et al. (2018).

The observable CDR for accident year i and accounting year $I + 1$ has been defined in England et al. (2018) in the following way:

$$\widehat{CDR}_i\left(I+1\right)=\hat{R}_i^{D_I}-\left(X_{i,I-i+1}+\hat{R}_i^{D_{I+1}}\right)=\hat{C}_{i,J}^{I}-\hat{C}_{i,J}^{I+1}, \tag{7}$$

where $\hat{R}_i^{D_I}$ and $\hat{R}_i^{D_{I+1}}$ are the reserves for the claims occurred at year i estimated at time I and $I + 1$, respectively. Value $X_{i,I-i+1}$ is an appropriate incremental claim. Finally, $\hat{C}_{i,J}^{I}$ and $\hat{C}_{i,J}^{I+1}$ are the estimates of ultimate claims calculated at time I and $I + 1$, respectively (all for accident year i). The aggregated CDR is then defined as:

$$\widehat{CDR}\left(I+1\right)=\sum_{i=1}^{I}\widehat{CDR}_i\left(I+1\right). \tag{8}$$

Since one of the goals of the numerical study is to compare the considered reserving methods how accurately they can estimate the actual values of the reserves and CDRs, we have to use some aggregate metric for the CDR comparison. For this purpose, we use a little modified form of the so-called CDR score introduced in Balona and Richman (2020):

$$\widehat{CDR}_{SCORE}\left(I+1\right)=\sqrt{\frac{\sum_{i=1}^{I}\left(\left|X_{i,I-i+1}\right|\cdot\left|\widehat{CDR}_i\left(I+1\right)\right|\right)}{\sum_{i=1}^{I}\left|X_{i,I-i+1}\right|}}. \tag{9}$$

where the difference lies in the choice of the absolute value of $CDR_i(I + 1)$ instead of the square function considered in the mentioned article. The aim is to minimize the CDR score which means that one has stable reserves with a reasonable change in reserves between calendar years I and $I + 1$. Therefore, one should calculate this metric in addition to the reserve estimation.

4 CLAIMS PORTFOLIO GENERATORS

In order to effectively evaluate considered methods in practice, it is necessary to apply them across a large number of portfolios. Additionally, having the data about the future development of such claims, in the language of run-off triangles to know the values in the lower triangle as well, is crucial for comparison of the reserves against actual outcomes. Therefore, an ideal solution consists in application of an insurance claims generator, which addresses both of these requirements. However, it is important to construct and calibrate these generators consistently to reflect the characteristics of real portfolios as closely as possible. In the numerical study presented in this paper, we apply the aforementioned methods to portfolios generated by three distinct generators, which are described in the following three subsections.

4.1 Generator based on real Czech data

The first generator used for insurance claims portfolios creation is the one created by the authors that is based on a real insurance claims portfolio delivered by the Czech Insurers' Bureau. This particular portfolio consists of claims paid by the guarantee fund administered by the Czech Insurers' Bureau, which covers expenses arising from incidents caused by vehicles without third-party liability insurance or unidentified vehicles where the responsible party is unknown. Our data concerned exclusively the claims that occurred between years 2001 and 2010. Furthermore, we considered several adjustments of the underlying portfolio which, however, do not affect the credibility of the generated portfolios.

The generator has been constructed applying maximum likelihood and kernel density estimation. First, the corresponding volume of claims is generated and for each claim its type is determined. Furthermore, the year of the claim occurrence and, based on the type of claim, the number of payments, the time until settlement and the size of claim are simulated. Since the number of payments and their delays often influence the value of these payments, this is also considered during the simulation. A portfolio, generated in this way, should have similar properties as the original portfolio.

This generator was used to generate 500 portfolios consisting of approximately 65 000 claims. An illustration of these data can be found in Table 1.

Table 1 Illustration of data in Czech Insurers' Bureau portfolios

ID	Type	AM	PM	Payment	Ultimate
1	1	64	69	23 706	23 706
2	1	109	124	40 813	40 813
3	2	94	101	9 629	36 704
3	2	94	112	4 587	36 704
3	2	94	122	22 488	36 704

Source: Own construction

For each claim information on *ID*, which serves to distinguish different claims, its *Type* (two possible types), accident month *AM* and *Ultimate* value, is known. Additionally, each row corresponds to one payment of amount *Payment* paid in month *PM*.

4.2 Generator Gabrielli and Wüthrich (2018)

The second generator proposed in Gabrielli and Wüthrich (2018) is based on neural networks to incorporate individual claims feature information. This individual claims history simulation machine was constructed applying a neural network architecture which was calibrated to real insurance data that have occurred between 1994 and 2005. For each of these claims, there is available additional information, such as the line of business or age of the injured, together with 12 years of claims development.

The architecture of this individual claims simulation machine consists of eight steps. Firstly, reporting delays that correspond to the differences between accident and reporting years are simulated. This is followed by payment indicator simulation, whether there are any payments or not. In the third step a number of payments is simulated, followed by a total claim size simulation. The last four steps then serve for cash flows modeling. For a detailed description of the simulation machine see Gabrielli and Wüthrich (2018).

After some adjustments that are necessary before clustering, which mainly consisted in excluding redundant information from the data, 500 portfolios comprising approximately 320 000 claims were generated. A few examples of generated data can be found in Table 2.

Table 2 Illustration of data in Gabrielli and Wüthrich portfolios

ID	LoB	AY	AQ	Age	RepDel	Payment	PayDel
1	4	1994	4	25	0	97	1
2	1	1994	2	39	0	1 476	0
2	1	1994	2	39	0	705	1
3	1	1994	1	38	0	5 709	0
3	1	1994	1	38	0	2 358	2

Source: Own construction

In addition to *ID*, for each claim it is also known its accident year *AY*, accident quarter *AQ*, reporting delay *RepDel*, age of the injured *Age* and line of business *LoB* (four possible types). In this case, each row corresponds to one payment of amount *Payment* paid with delay *PayDel*.

4.3 Generator Wang and Wüthrich (2022)

The third individual claims generator introduced in Wang and Wüthrich (2022) is based on the R package *SynthETIC* of Avanzi et al. (2021). The *SynthETIC* simulator specifically allows for desirable data features typically occurring in practice. It has been structured in such a way that the generated portfolio of claims should resemble an auto liability portfolio. Moreover, its code has a modular form. The generator consists of eight modules such as, e.g., claim occurrence date, number of partial payments, sizes of partial payments without allowance for inflation, distribution of payments over time or claim inflation. Such an independent coding allows adjustments in each module, their possible replacements or removal according to a particular purpose.

This simulator has been modified by Wang and Wüthrich. They have complemented this simulation environment with additional claim features resulting with the enhanced generator. For more details see Wang and Wüthrich (2022).

Similarly, as in the previous generator, several adjustments in the generated portfolios have been considered. Again, 500 portfolios were generated, in this case each with approximately 50 000 claims. An illustration of the data follows (see Table 3).

Table 3 Illustration of data in Wang and Wüthrich portfolios

ID	Type	AY	Payment	PayDel	Ultimate
1	1	1	2 434	1	2 434
2	4	1	11 017	0	34 110
2	4	1	11 815	1	34 110
2	4	1	11 278	2	34 110
3	3	1	2 428	1	2 428

Source: Own construction

Variable *ID* has the same meaning as before. Each claim is described by its accident year *AY* and its *Type* (six possible classes). Again, each row corresponds to one payment of amount *Payment* paid with delay *PayDel*.

The clustering method CLARA described in Section 2.3 requires the feature matrix entering the clustering process in a numerical form, i.e., it is necessary to transform the categorical variables, such as *Type*, *LoB* or *AQ* into dummy variables. Moreover, variable *Age* is further considered in the form of three dummy variables , (≤ 30, ≥ 51 and the interval between).

5 COMPARISON OF NUMERICAL STUDY RESULTS

This section deals with the numerical study using data created by means of generators from Section 4. The generators create not only the data in the upper run-off triangle in Figure 1, that are used for claim reserves prediction, but also for the predicted lower triangle in Figure 1. Hence, one can evaluate for each generated portfolio the accuracy of particular reserve methods.

Given that the log-normal model assumes the log-normality of incremental claims, it is also important to verify that such an assumption holds for the considered data. There are several tests that can be used, both those that test directly the log-normality of incremental claims and also those that test the normality of the logarithmized values. In our case we have chosen the well-known Kolmogorov-Smirnov test that was used for random samples of claims of size 500, since the total number of generated incremental claims is too large. Repeated testing for various random samples has shown that the null hypothesis of log-normality of the data cannot be rejected.

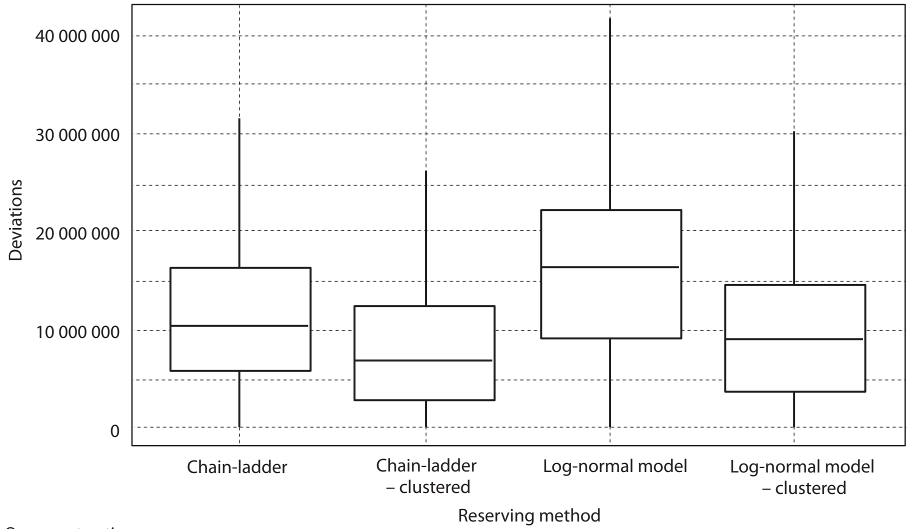
5.1 Results of numerical study

For each portfolio, introduced in Subsections 4.1–4.3, the main interest concerns comparing deviations of reserves from the actual values and modified CDR scores presented in Section 3 for the considered reserving method described in Section 2. This is achieved using graphs with boxplots. Results for the Czech portfolio presented in Section 4.1 follow.

Figure 2 presents deviations among estimated and actual reserves depicted graphically by means of boxplots over 500 portfolios generated by the Czech data generator from Section 4.1 for particular reserve methods (Chain-ladder, Chain-ladder clustered, log-normal model and log-normal model clustered). Figure 3 is analogous for modified CDR score.

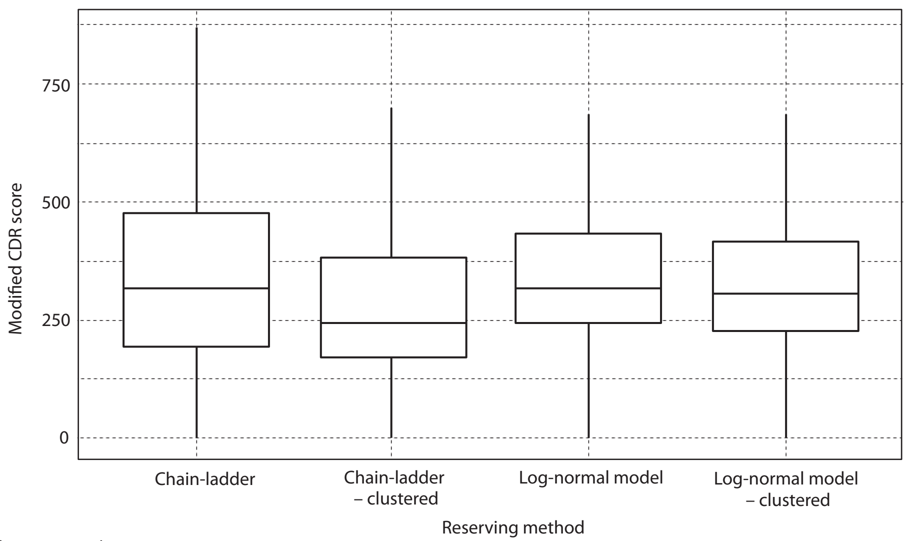
In addition to the boxplots presented in Figures 2 and 3, one can be interested in the accuracy of individual estimates, not only the aggregated reserve. For this purpose, we consider two approaches how to measure and compare the accuracy of considered reserving methods concerning individual estimates.

Figure 2 Deviations of reserves – Generator based on real Czech data



Source: Own construction

Figure 3 Modified CDR score – Generator based on real Czech data



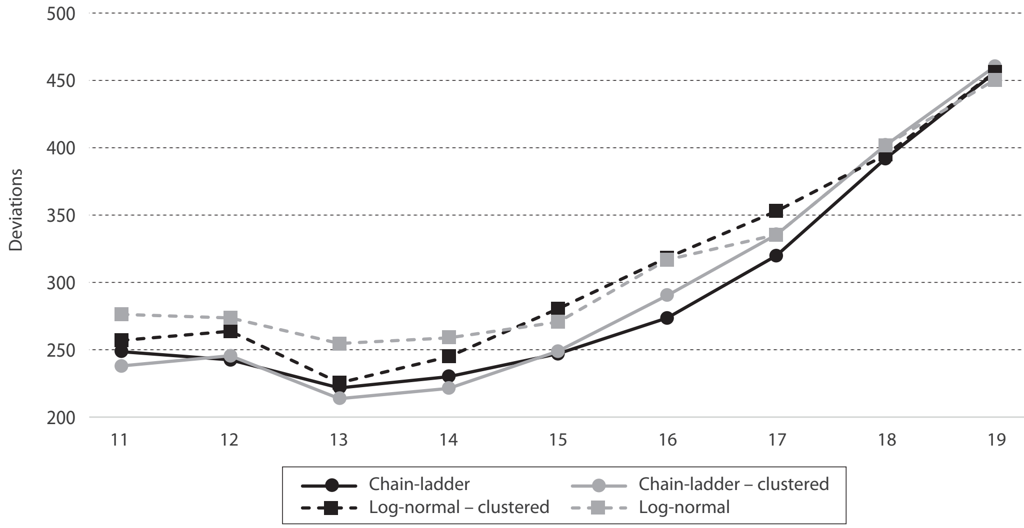
Source: Own construction

In the first case, the accuracy of estimates in individual calendar years following the last accounting year for which the claims are known is compared. Graphical results are presented only for the first portfolio to save space. For each subdiagonal in the lower run-off triangle, we calculate the following value:

$$D_k = \sqrt{\frac{\sum_{i=k-J}^{I}\left(\left|X_{i,k-i}\right| \cdot \left|X_{i,k-i} - \hat{X}_{i,k-i}\right|\right)}{\sum_{i=k-J}^{I}\left|X_{i,k-i}\right|}}, \tag{10}$$

where $k > I$. Values D_k for the considered reserving methods are shown by means of the graph in Figure 4.

Figure 4 Deviations development – Generator based on real Czech data

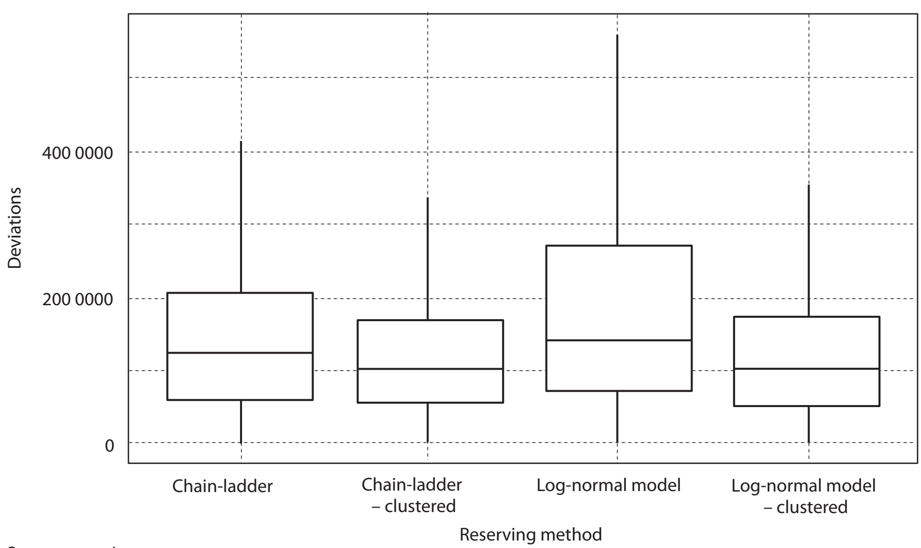


Source: Own construction

Since the second approach to the accuracy of individual estimates is applied to all three portfolios, it will be presented at the end of this subsection.

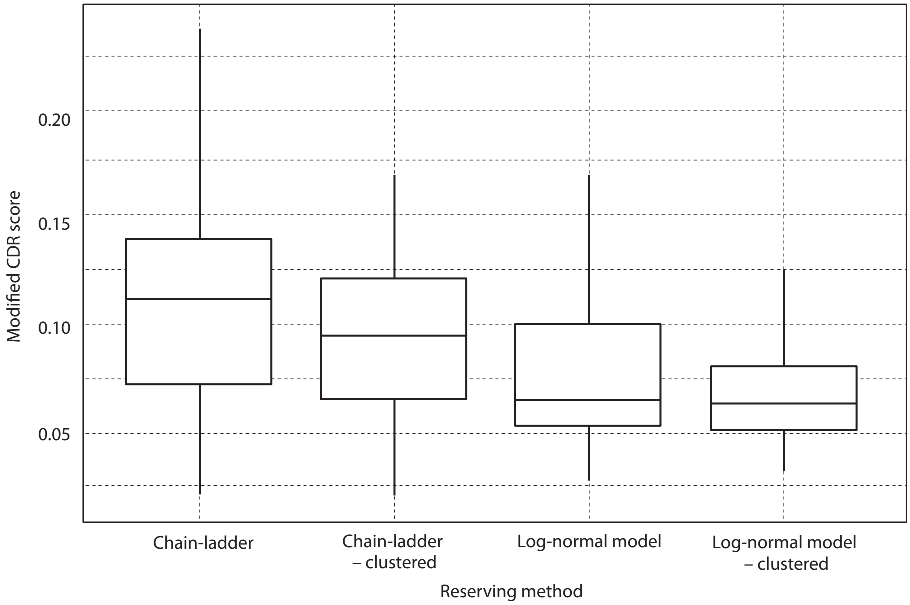
For the second portfolio from Section 4.2, we present the results corresponding to the deviations of reserves and modified CDR score in the same form as for the first portfolio, see Figures 5 and 6.

Figure 5 Deviations of reserves – Generator Gabrielli and Wüthrich (2018)



Source: Own construction

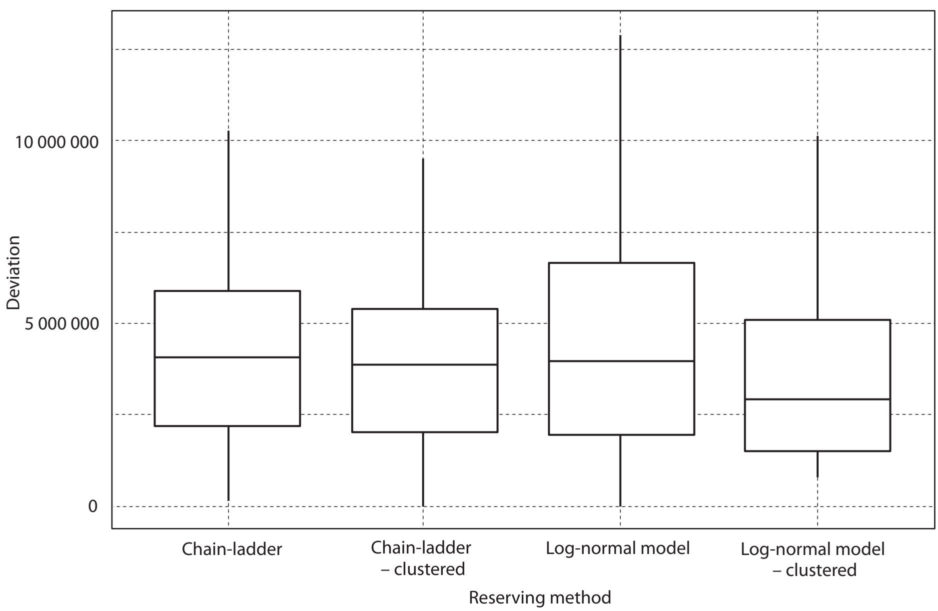
Figure 6 Modified CDR score – Generator Gabrielli and Wüthrich (2018)



Source: Own construction

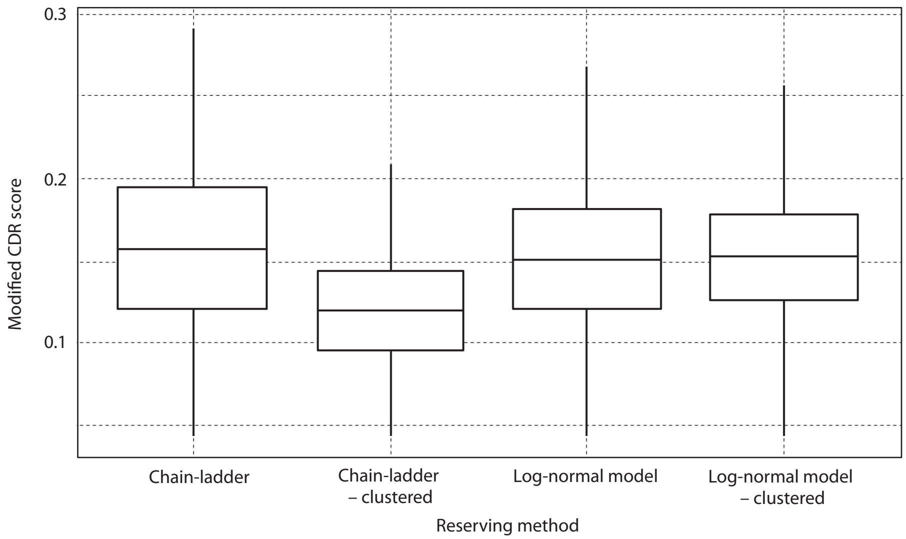
Finally, boxplots corresponding to the results obtained for the third portfolio from Section 4.3 are given in Figures 7 and 8.

Figure 7 Deviations of reserves – Generator Wang and Wüthrich (2022)



Source: Own construction

Figure 8 Modified CDR score – Generator Wang and Wüthrich (2022)



Source: Own construction

The second approach mentioned above compares all the estimated values in the lower run-off triangle (see Figure 1). The corresponding metric is calculated as the sum of squared differences of estimated and actual values (we call it the modified Frobenius norm). Since for each generated portfolio one obtains one value for each reserving method, we present medians of these values in Table 4.

Table 4 Medians of modified Frobenius norm of estimated run-off triangles

Portfolio	Chain-ladder	Chain-ladder – clustered	Log-normal model	Log-normal model – clustered
Gen. based on real Czech data	6 832 916	6 828 086	7 497 782	7 269 261
Gen. Gabrielli and Wüthrich	130 800	126 495	144 577	126 337
Gen. Wang and Wüthrich	1 732 335	1 728 779	2 193 398	2 179 933

Source: Own construction

5.2 Discussion of results

The results obtained in Section 5.1 allow us to compare the considered reserving methods including the impact of clustering. Figure 2 shows that Chain-ladder achieves significantly better results when compared to the log-normal model. The Chain-ladder boxplot values are lower than in the case of the log-normal model and the median of deviations is notably lower as well. One can see that clustering considerably improves the estimates, since for both reserving methods the boxplots narrowed down. Applying the paired sign test for equality of medians, the null hypothesis is strongly rejected (with p-value less than 0.001) in both cases in favor of one-sided alternative. This confirms the significant improvement of the accuracy of reserve estimates after clustering. Without clustering, the Chain-ladder dominates the log-normal model, but after incorporating clustering, the difference almost disappears.

In Figure 3, where the modified CDR scores are compared, a slight improvement after clustering can be seen as well, however, it is nearly negligible for the log-normal model. When comparing the methods by means of a graph showing the deviations development over individual calendar years (Figure 4), one can see that the Chain-ladder achieves almost always lower values than the log-normal model. In the first half, the most accurate variant was the clustered version, in the second half the non-clustered one.

Similar results as in Figure 2 can be found in the remaining figures corresponding to the other considered portfolios. Again, the pairwise sign tests confirm clustering improvement. However, significantly lower values of the modified CDR score can be observed for the log-normal model.

All reserving methods are also compared with respect to the modified Frobenius norm introduced in Section 5.1. Table 4 contains medians of the calculated norm values for each generated portfolio and one can see that also according to this table, the clustering improves the estimates. However, the dominance of the Chain-ladder persists.

CONCLUSION

This article discussed the importance of loss reserving in non-life insurance and compared two different reserving methodological methods – the Chain-ladder method and the state-space modeling. It also confirmed the significance of clustering in the context of loss reserving. The concept of run-off triangles and how actuaries can use statistical techniques and mathematical models to analyze the patterns within the run-off triangles to estimate future payments and appropriate reserves were explained. Additionally, we provided an overview of two distinct regulatory frameworks used in insurance, Solvency II and IFRS 17, and introduced the CDR approach. In this context, the paper can be useful for actuaries dealing with reserve estimation in non-life insurance practice.

In the numerical study, we presented the comparison of deviations of reserves from actual values and modified CDR scores for different reserving methods using boxplots. We discussed the accuracy of estimates in individual calendar years and explored the overall accuracy of the estimates. The numerical study demonstrated the benefit of clustering when considered in loss reserving. The results show that the Chain-ladder achieved better results when compared to the log-normal model, and that clustering considerably improved the estimates for both reserving methods. We also discussed the accuracy of individual estimates and presented graphical results for the first portfolio.

The obtained results may serve as a hint to improve the state-space methodology in order to give comparable results with classical approaches to reserving. The reason is obvious: in future the so-called micro reserving will play a key role in non-life insurance reserving based on neural networks and deep learning where the classical methods of the type of Chain-ladder will be quite insufficient for the corresponding computational procedures (see, e.g., Avanzi et al., 2021; Balona and Richman, 2020; Gabrielli and Wűthrich, 2018; Wang and Wűthrich, 2022; Wűthrich and Merz, 2008). One can also try to extend the one step ahead CDR predictions to the ones for more steps ahead in multivariate run-off triangles including analytical formulas for prediction errors.

ACKNOWLEDGEMENT

The authors thank to Dr. Radek Hendrych for his efficient help when preparing this paper and to anonymous reviewers for helpful comments and improvements.

References

ATHERINO, R., PIZZINGA, A., FERNANDES, C. (2010). A row-wise stacking of the runoff triangle: State space alternatives for IBNR reserve prediction [online]. *ASTIN Bulletin*, 40(2): 917–946. <http://doi.org/10.2143/AST.40.2.2061141>.

AVANZI, B., TAYLOR, G., WANG, M., WONG, B. (2021). SynthETIC: an individual insurance claim simulator with feature control [online]. *Insurance: Mathematics and Economics*, 100: 296–308. <https://doi.org/10.1016/j.insmatheco.2021.06.004>.

BALONA, C., RICHMAN, R. (2020). The actuary and IBNR techniques: A machine learning approach [online]. *SSRN Electronic Journal*. <http://doi.org/10.2139/ssrn.3697256>.

BORNHUETTER, R. L., FERGUSON, R. E. (1972). The actuary and IBNR. *Proceedings of the Casualty Actuarial Society*, 59: 181–195.

BROCKWELL, P. J., DAVIS, R. A. (1991). *Time Series: Theory and Methods*. 2nd Ed. New York: Springer-Verlag. ISBN 978-0-387-97429-3

CIPRA, T. (2010). *Financial and Insurance Formulas.* Berlin, Heidelberg: Springer-Verlag. ISBN 978-3-7908-2592-3

COSTA, L., PIZZINGA, A. (2020). State-space models for predicting IBNR reserve in row-wise ordered runoff triangles: Calendar year IBNR reserves & tail effects [online]. *Journal of Forecasting*, 39: 438–48. <https://doi.org/10.1002/for.2638>.

DE FELICE, M., MORICONI, F. (2019). Claim Watching and Individual Claims Reserving Using Classification and Regression Trees [online]. *Risks*, 7(4): 1–36. <https://doi.org/10.3390/risks7040102>.

DE JONG, P. (2005). State Space Models in Actuarial Science. *Proceedings of the Second Brazilian Conference on Statistic*: 16–31.

DELONG, Ł., LINDHOLM, M., WÜTHRICH, M. V. (2021). Collective reserving using individual claims data [online]. *Scandinavian Actuarial Journal*, 1: 1–28. <https://doi.org/10.1080/03461238.2021.1921836>.

DU, K. L. (2010). Clustering: A neural network approach [online]. *Neural Networks*, 23: 89–107. <http://doi.org/10.1016/j.neunet.2009.08.007>.

DUVAL, F., PIGEON, M. (2019). Individual Loss Reserving Using a Gradient Boosting-Based Approach [online]. *Risks*, 7(3). <https://doi.org/10.3390/risks7030079>.

ENGLAND, P., VERRALL, R. J., (2002). Stochastic Claims Reserving in General Insurance [online]. *British Actuarial Journal*, 8(3): 443–518. <https://doi.org/10.1017/S1357321700003809>.

ENGLAND, P., VERRALL, R. J., WÜTHRICH, M. V. (2018). On the lifetime and one-year views of reserve risk, with application to IFRS 17 and Solvency II risk margins [online]. *SSRN Electronic Journal*. <http://doi.org/10.2139/ssrn.3141239>.

GABRIELLI, A. V., WÜTHRICH, M. (2018). An individual claims history simulation machine [online]. *Risks*, 6(2). <https://doi.org/10.3390/risks6020029>.

HARTIGAN, J. A., WONG, M. A. (1979). Algorithm AS 136: A K-means clustering algorithm [online]. *Applied Statistics*, 28: 100–108. <https://doi.org/10.2307/2346830>.

HELSKE, J. (2017). KFAS: Exponential family state space models in R [online]. *Journal of Statistical Software*, 78(10): 1–38. <https://doi.org/10.18637/jss.v078.i10>.

HENDRYCH, R., CIPRA, T. (2021). Applying state space models to stochastic claims reserving [online]. *ASTIN Bulletin*, 51(1): 267–301. <https://doi.org/10.1017/asb.2020.38>.

KAAS, R., GOOVAERTS, M., DHAENE, J., DENUIT, M. (2004). *Modern Actuarial Risk Theory*. Berlin, Heidelberg: Springer-Verlag. ISBN 978-3-540-70998-5

KAUFMAN, L., ROUSSEEUW, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: Wiley. ISBN 0-471-87876-6

KAUFFMANN, J., ESDERS, M., RUFF, L., MONTAVON, G., SAMEK, W., MÜLLER, K. R. (2022). From clustering to cluster explanations via neural networks [online]. *IEEE Transactions on Neural Networks and Learning Systems,* 35(2): 1926–1940. <https://doi.org/10.1109/tnnls.2022.3185901>.

MACK, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates [online]. *ASTIN Bulletin*, 23(2): 213–225. <https://doi.org/10.2143/AST.23.2.2005092>.

MUNROE, D., ZEHNWIRTH, B., GOLDENBERG, I. (2018). Solvency capital requirement and the claims development result [online]. *British Actuarial Journal*, 23(15). <https://doi.org/10.1017/S1357321718000041>.

PAPENBERG, M., KLAU, G. W. (2021). Using anticlustering to partition data sets into equivalent parts [online]. *Psychological Methods*, 26(2): 161–174. <https://doi.org/10.1037/met0000301>.

SCRUCCA, L., FOP, M., MURPHY, B., RAFTERY, A. E. (2016). Mclust 5: clustering, classification and density estimation using Gaussian finite mixture models [online]. *The R Journal*, 8(1): 289–317. <https://doi.org/10.32614/RJ-2016-021>.

THE ACTUARIAL COMMUNITY (2022). *Loss Data Analytics* [online]. [cit. 22.7.2024]. <https://openacttexts.github.io/Loss-Data-Analytics/ChapLossReserves.html#S:Data>.

VEJMĚLKA, P. (2023). Clustering methods usable in loss reserving in non-life insurance and their comparison. *Proceedings of the 41st International Conference on Mathematical Methods in Economics:* 421–426. ISBN 978-80-11-04132-8

VERRALL, R. J. (1989). A State Space Representation of the Chain Ladder Linear Model [online]. *Journal of the Institute of Actuaries*, 116(3): 589–610. <https://doi.org/10.1017/S0020268100036714>.

WANG, M., WÜTHRICH, M. (2022). Individual claims generator for claims reserving studies: Data Simulation R [online]. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4127073>.

WÜTHRICH, M. V. (2018). Machine learning in individual claims reserving [online]. *Scandinavian Actuarial Journal*, 2018(6): 465–480. <https://doi.org/10.1080/03461238.2018.1428681>.

WÜTHRICH, M. V., MERZ, M. (2008). *Stochastic Claims Reserving Methods in Insurance.* West Sussex: Wiley & Sons. ISBN 978-0-470-77272-0

ZHANG, Y. (2010). A general multivariate chain ladder model [online]. *Insurance: Mathematics and Economics*, 46(3): 588–599. <https://doi.org/10.1016/j.insmatheco.2010.03.002>.