Efficient capital management using an internal model: a case of non-life insurance

Ilze Zariņa-Cirule*, Gaida Pettere and Irina Voronova

Department of Entrepreneurship and Management, Faculty of Engineering Economics and Management and Department of Engineering Mathematics, Faculty of Computer Science and Information, Riga Technical University, Riga, Latvia

Received 2 December 2021, accepted 24 March 2022, available online 24 August 2022

© 2022 Authors. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International License CC BY 4.0 (http://creativecommons.org/licenses/by/4.0).

Abstract. The main goal of insurance company management is to increase shareholders’ value and implement a strategy that promotes sustainable growth of the company. Well-known possible measures intended to achieve that goal are as follows: share price, economic value, market capitalisation, gross premiums earned and solvency ratio. These measures include efficient capital management as capital expenses could be a major cost position depending on risk appetite and the extent of capital needed to support it. This research focuses on non-life insurers for reserve risk modelling. In the current study, a more accurate risk quantification model has been developed than the standard model provided by the EU regulator under the Solvency II framework. The proposed model provides capital cost gains as well. A case study based on non-life real data set with underwriting in the Baltic countries is discussed with inclusion of pandemic trends that had an impact on economies and customer behaviours. The study considers different non-life reserve distributions for each insurance business line, risk aggregation and the way of choosing the most appropriate type of copula model for non-life reserve risk. Adequate capital is calculated by applying value at risk at 99.5%, which is mandatory in the EU market. The study considers which selected tests have to be implemented in order to choose the most appropriate copula model for reserve risk.

Keywords: capital management, copula, financial stability, internal model, non-life insurance, stochastic reserving, value at risk.

1. INTRODUCTION

The EU insurance industry has fully recognized and acknowledged the role of capital in risk monitoring and control since the introduction of Solvency II framework in 2016. It is the first risk-based insurance regulation after financial crises and financial market pressure from the banking sector where such regulations have been in force since 2010. The previous solvency margin requirements were established in 1973 under the first council directive for non-life insurers (73/239/EEC), which had no direct link between riskiness of portfolio and required capital needs.

The new regulation has extended performance measures such as return on risk-adjusted capital and return on equity, all of which are reflected in the concept of the risk underwritten by insurers. Today’s mindset of culture and management means not only the measure of profit or whether the expected net cash flow is positive, but it also implies whether the return obtained within a given group of contracts is proportionate to the risk incurred. Dacorogna (2018) has published an overview on the new solvency rules, their imple-

*Corresponding author, ilzezarina@inbox.lv
mentation in the organisational structure and their adoption procedure. The author describes how a progressive insurer should operate from the enterprise risk management perspective, taking into account direct tasks to be performed with regard to different management roles (e.g., chief executive, chief risk, chief underwriting, chief financial officers, etc.).

The object of the paper is non-life insurance, the subject – efficient capital management using an internal model.

The present section considers the concept of capital, providing the distinction between available capital $C_a$ and total risk-adjusted capital for all risks $C_{rT}$ (or solvency capital requirement, SCR) based on probability distributions. We also provide literature review of the main internal model concepts and pertinent mathematical approaches – copula, stochastic and deterministic reserving, and hypothesis tests for the copula model. In Sections 2 and 3, we introduce theoretical concepts and the algorithm of the model proposed by the authors of the article. Section 4 describes the results of the case study where the Baltic countries’ non-life insurance data are used and identifies the difference between the required capital of a standard model and that of the internal model. In the last section, we conclude the paper by summarising the features marking the evolution of the internal model in terms of future perspectives and challenges.

Capital can be seen as a guarantee to every customer that the company will meet all its obligations up to a certain level of probability. The EU regulator of insurance and pensions industry (EIOPA) requires a level of 99.5% within one year. In their turn, customer’s obligations are claims (e.g., costs of repairing hail damage to car, fire to household real estate). Insurance provides a fundamental social function and, therefore, a regulator imposes minimum levels to the amount of capital that it must hold. This minimum amount in various papers is called risk-adjusted capital, regulatory capital, or (SCR). Capital actually held by the insurer is called economical capital or available capital $C_a$ which is higher than the regulatory capital $C_{rT}$ and driven by many considerations, such as to protect the company from insolvency, to maintain the rating given by major rating agencies (e.g., S&P), to appeal to the investors or to increase the number of customers, especially corporate customers. The company’s solvency ratio, $SR$, is then defined as follows:

$$SR = \frac{C_a}{C_{rT}} > 1.$$  \hfill (1)

Available capital is provided by investors of the insurance company, who ask for a certain return on the capital, which is above the level of almost risk-free return that they could get from government bonds. The required return depends on the level of riskiness. The next performance measure is return on required capital (RORC), which should be maximised by the management to achieve the highest return with a given risk expressed as the required capital for all risks $C_{rT}$ and annual profit $Profit_a$. It is defined as follows:

$$RORC = \frac{Profit_a}{C_{rT}},$$  \hfill (2)

where the aim of management is to maximise function

$$f(Profit_a, C_{rT}) = \frac{Profit_a}{C_{rT}},$$  \hfill (3)

where $C_{rT} > 0$. Formula (3) explains a well-known principle of the efficient frontier in the modern portfolio theory, which was first formulated by Markowitz (1952). The aim of the present paper is to provide the detailed algorithm of the model for required capital $C_{rT}$, which is called an internal or partial model under Solvency II framework. The proposed model reflects reserve risk assessment. In non-life insurance companies claim reserving is the main process which determines what is held on the balance sheet for claims that are not yet settled, affects the level of risk premium, influences the capital that is held to support the solvency position, as well as affects dividend distribution, its frequency and stability. Thus, the amount of capital that must be held for reserve risk is crucial for both society and investors of the company.

Although insurance claims data are usually modelled by skewed, heavy-tailed distributions, a regulator applies many simplifications in a standard formula, where risk aggregation for reserve risk is done using the Pearson linear correlation matrix (EIOPA 2014). The proposed model for required capital is modelled using
risk aggregation with copula model which addresses insurance data specifics by aggregating risks with an adequate modelling approach. Such an approach will avoid insolvency due to inaccurate capital assessment for dependent losses between lines of business (LoB) and by ignoring non-linear dynamics, especially in tails. Risk aggregation technique applied in a standard formula is mentioned as the main weakness based on the authors’ previous literature review (Zarina et al. 2019).

By searching for the terms ‘insurance & copula’ in the Scopus database, we would like to highlight the most cited and the most relevant papers and books that have also affected the proposed model described in subsequent sections of the paper. McNeil et al. (2005) provide a comprehensive examination of the theoretical concepts and modelling techniques of quantitative risk management, including a copula approach for the financial sector. Patton (2006) tests asymmetry in a model of the dependence between the Deutsche mark and the yen, and takes into account correlation via a copula approach. Aas et al. (2009) show how financial multivariate data, which exhibit complex patterns of dependence in the tails, can be modelled using a cascade of bivariate-copula. Cherubini et al. (2004) apply copulas to major topics in derivative pricing and counter-party default risk analysis. In turn, Genest et al. (2009) provide goodness-of-fit testing of copula models with application to insurance data.

Claim reserve calculation between lines of business is the next major topic. Although different techniques are used, classical chain ladder method (CLM) is usually the basis according to the 2016 report on reserving practices by the International Actuarial Association. We will apply a stochastic method with underlying deterministic chain ladder approach. The concept of the method was introduced by Tarbell (1934) and it became well-known in the early 1970s. England and Verrall (2002) have summarised literature on the topic pointing out notable papers such as Kremer (1982), Taylor and Ashe (1983), Renshaw (1989). We refer to the most cited papers in the Scopus database after searching the terms ‘chain ladder & reserve’ and ‘chainladder & reserve & copula’. Mack (1993a, 1993b, 1994) describes the standard error formula for a distribution-free reserving CLM, which helps assess reserve risk. Ashe et al. (1986) introduced a stochastic approach for reserve risk assessment with bootstrap estimates of prediction errors in claim reserving based on a bootstrapping statistical technique first proposed by Efron (1979). The method produces an estimate of the distribution of future cash outflows. Shi and Frees (2011) apply copula approach and parametric bootstrap for multiple multiyear run-off triangles for personal and motor business with data of the majority of U.S. insurers. Fersini and Melisi (2016) applied a stochastic model to evaluate the fair value of motor third-party liability and quantification of the capital requirement from the Solvency II perspective using the data of Italy. We have not identified any paper addressing the whole process consisting of a proposed algorithm for an internal model with application to real data sets (underwriting in the Baltic market) where different distributions for lines of business for the product exist, data is joined with a correlation matrix, and copula family tests are performed in the context of Solvency II framework with one-year time horizon.

2. THEORETICAL APPROACH OF INTERNAL MODEL TECHNIQUES

2.1. A case for non-internal default model or standard formula set by the EU regulators

Calculation steps and underlying assumptions are described in Solvency II Directive (138/2009/EC) and Commission Delegated Regulation (EU) 2015/35 supplementing the directive. Firstly, required capital $C_r$ is set aside with a defined capital standard based on a 0.995 quantile $\alpha$, called value at risk ($\text{VaR}_\alpha$), where time horizon is one year, and it is calculated by considering risk mitigation such as reinsurance protection. Confidence level must be also used for an internal model. Secondly, it is assumed in standard formula that reserve distribution for each and every product and line of business has log-normal distribution. Next, 68–95–99.7 rule or empirical rule is used. Capital for reserve risk $C_r$ in case of product (or line of business) $e$ in insurer’s portfolio is as follows:

$$C_r = 3 \cdot \sigma_e \cdot CBE_e,$$

(4)
where $\sigma_e$ denotes volatility measure, standard deviation for $e$ product reserve risk and $CBE_e$ is volume measure or the best estimate of the claim reserve in economical balance sheet for the product $e$. Generally casualty insurers’ portfolios consist of different lines of business. Correlation and diversification effect then is reflected by calculating a standard deviation coefficient $\sigma_{total}$ for the whole portfolio as follows:

$$
\sigma_{total} = \frac{1}{CBE_{total}} \sqrt{\sum_{e,p} CorrS_{(e,p)} \cdot \sigma_e \cdot \sigma_p \cdot CBE_e \cdot CBE_p},
$$  \hspace{1cm} (5)

where $CBE_{total}$ is the sum of claim reserves best estimate after reinsurance for all the lines of business, with the sum covering all possible combinations ($e$, $p$) of the lines of business $e$ to $p$; $CorrS_{(e,p)}$ signifies the correlation coefficient between lines of business $e$ and $p$ set out by the Solvency II Directive.

2.2. Internal model approach using copula

2.2.1. General principles

The authors propose, instead of formula (5) and correlation coefficients set out by the EIOPA, applying Spearman rank correlations, reserve distributions and another risk aggregation technique – copula. The risk aggregation procedure is the same as in the standard model and market practice if an alternative model is not accepted by regulators. As mentioned in the section above, alternative capital requirement for reserve risk should be calculated by using formula:

$$
C_e = VaR_{99.5\%}^e - CBE_e,
$$  \hspace{1cm} (6)

where $VaR_{99.5\%}^e$ denotes value at risk ($VaR$) at 99.5% confidence level for line of business $e$ and $CBE_e$ is the best estimate of claim reserve for line of business $e$ or $VaR$ at 50% confidence level which represents fair value of liabilities in economical balance sheet. The same principle works for aggregated reserve risk (of different business lines), which is the difference between value at risk at 99.5% and the mean or the best estimate booked in economical balance sheet.

2.2.2. Types of uncertainty in models and bootstrap chain ladder method for claims best estimate

Deterministic CLM is one of the key methods that have been developed for use in non-life insurance. This method is used to derive reserve estimates and provide a single estimate of reserves to be booked without uncertainty and potential shift assessment around the estimate. Real data sets are organized in a triangle format (e.g., incurred claims) where past development is used as a guide for estimation claims development in future. The concept was introduced by Tarbell in 1934 and it became well known in the early 1970s. The basis of the method is as follows:

$$
\{IC_{ij}: i = 1, \ldots, n; j = 1, \ldots, n - i + 1\},
$$

$$
D_{ij} = \sum_{k=1}^{i} IC_{ik},
$$

$$
\tilde{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} D_{ij}}{\sum_{i=1}^{n-j+1} D_{ij}}, \quad j \in \{2, \ldots, n\},
$$

$$
\tilde{D}_{ln-i+2} = D_{ln-i+1} \tilde{\lambda}_{n-i+2}, \quad i \in \{2, \ldots, n\},
$$

$$
\tilde{D}_{lk} = \tilde{D}_{l,k-1} \tilde{\lambda}_k, \quad k \in \{n - i + 3, n - i + 4, \ldots, n\}, i \in \{3, \ldots, n\},
$$

where $IC$ represents incremental claims data; the suffix $i$ refers to the row indicating accident year; the suffix $j$ refers to the column and indicates the delay, here assumed to be measured in years. $D_{ij}$ denotes assumed cumulative claims. The development factors of the CLM are denoted by $\{\tilde{\lambda}_j: j = 2, \ldots, n\}$ and the estimates
of the development factors $\lambda_j$ are then applied to the latest cumulative claims in each row $D_{i,n-i+1}$ to produce forecasts of future values of cumulative claims.

However, the calculated estimates can be reliable if there is sufficient historical data and historical uncertainty can also be assumed as future uncertainty. Internal capital modelling team must take into account types of uncertainty errors in the model that will improve reality (see Fig. 1).

Figure 1 demonstrates how important it is to use expert judgment during the reserving process. The proposed internal model excludes the procedure of expert judgement as each entity product design, and local legal requirements differ from each other. Stochastic method application is crucial for determination of capital requirements. However, it is an essential tool during the business planning process, merger and acquisition transactions, reinsurance pricing, and the approach is not restricted by the modelling distributions. Mathematical representation of prediction uncertainty is measured with the mean squared error of prediction, which can be divided into two components for process and parameter estimation variance, and overall practice changes depending on which stochastic reserving method or procedure is used:

$$MSEP(\hat{X}_{i,j}) \approx \text{Var}(X_{i,j}) + \text{Var}(\hat{X}_{i,j}),$$

where $X$ denotes an unknown future value or claims best estimate and $\hat{X}$ represents the estimators. This formula has been explored in the context of stochastic reserving by such authors as Taylor and Ashe (1983), Renshaw (1994), and based on conditional probabilities by Mack (1993a), Merz and Wüthrich (2008).

Fig. 1. Types of uncertainty for reserve setting and its capital requirements (created by the authors based on Marshall et al. 2008 and Hindley 2017).
Deterministic CLM is an underlying method for a stochastic method used in further study – bootstrap chain ladder. The method is simulation based and therefore produces an estimate of the full distribution of future claims and operates within one-year time horizon. Claim distribution finding is the key reason why a stochastic reserving technique is used in the model and specific distributions for each line of business will be later used for risk aggregation process in the copula approach and for finding aggregated distribution.

Bootstrapping (Efron & Tibshirani 1993) is a powerful yet simple simulation technique; the methodology is based on sampling with replacement from the observed data sample to create a large number of pseudo-samples, which are consistent with the underlying distribution (England and Verrall 1999).

In a standard application of bootstrapping, where data are assumed to be independent and identically distributed, resampling with replacement is performed with source data. In regression type problems, the data are usually assumed to be independent, but not identically distributed, since the means (and possibly the variances) depend on covariates. Therefore, in regression type problems, it is common to use bootstrap residuals, rather than the raw data, since the residuals are approximately independent and identically distributed, or can be made so. For generalised linear models (GLM), there is a range of extended definitions of residuals; the precise form being dictated by the underlying modelling distribution (see McCullagh & Nelder 1989). For the over-dispersed Poisson chain ladder model, we use the Pearson residuals for bootstrapping. After discarding the suffices that indicate the origin and development year, the Pearson residuals \( r_p \) are defined as

\[
r_p = \frac{IC - \hat{m}}{\sqrt{m}},
\]

where \( \hat{m} \) is the fitted incremental claim and IC denotes incremental claim amount given by the over-dispersed Poisson chain ladder model (see England and Verrall 2002, p. 21). The bootstrap process involves resampling with replacement from the residuals. A bootstrap data sample is then created by inverting Eq. (13) and using the resampled residuals together with the fitted values. Given a resampled Pearson residual \( r_p^* \) together with the fitted value \( m \), the associated bootstrap incremental claims amount \( IC^* \) is given by

\[
IC^* = r_p^* \sqrt{\hat{m}} + \hat{m}.
\]

Resampling the residuals (with replacement) gives rise to a new triangle of past claims payments. Having obtained the bootstrap sample, the model is refitted and the statistic of interest calculated. Strictly, we ought to fit an over-dispersed Poisson GLM to the bootstrap sample to obtain a bootstrap reserve estimate. However, we can obtain identical reserve estimates using standard chain ladder (CL) methodology. At this point the usefulness of the bootstrap process becomes apparent, we do not need sophisticated software to fit the model, a spreadsheet will suffice. Having fitted the CL model to the bootstrap sample and obtained forecast incremental claims payments, we invoke the second stage of the procedure which replicates the process variance. This is achieved by simulating an observed claims payment for each future cell in the run-off triangle, using the bootstrap value as the mean and using the process distribution assumed in the underlying model, which, in this case, is over-dispersed Poisson model. The procedure is repeated a large number of times, each time providing a new bootstrap value and simulated forecast payment. For each iteration, the reserves are calculated by summing the simulated forecast payments. The set of reserves obtained in this way forms the predictive distribution, from which summary statistics, such as the prediction error, can be obtained (which is simply the standard deviation of the distribution of reserve estimates). More detailed description of the bootstrap procedure is given in England and Verrall (2002) and Hindley (2017).

General procedure of a non-parametric residual resampling bootstrap with regard to claims best estimate is as follows (Hindley 2017):
1. Define a statistical model that is appropriate for modelling the claims development process. This model will produce estimates of the future claims payments.
2. Fit this model to an observed data triangle.
3. Determine appropriately defined residuals between the fitted statistical model and the observed data.
4. Use Monte Carlo simulation to produce random selections of the residuals (with replacement).
5. Use the randomly generated residuals to generate new ‘pseudo data’ analogues to the observed data sample.
6. Re-fit statistical model to each version of the pseudo data and predict forecasts of the future claims payments, ensuring that the process error is incorporated in a suitable way.
7. Finally, examine the distribution of the forecasts to produce estimates of the prediction error – related to uncertainty caused by both parameter and process error.

2.2.3. One-year bootstrap approximation for vector – one-year run-off shifts for the best estimate

In the context of capital requirement setting in internal modelling, we are interested in one-year time horizon and, therefore, with regard to the reserving area it is one-year claim development and its distributions. Merz and Wüthrich (2008, 2014) have published the way how claim development for one year can be derived using the bootstrap CLM. The main advantages of the bootstrap methodology are summarised by Boumezoued et al. (2011) and Diers (2008) and illustrated in Fig. 2.

2.2.4. Distribution fitting method for one-year run-off vector

Claim distributions usually are skewed. Several distributions such as gamma, Weibull, normal, log-normal, exponential were applied to data and tested by goodness-of-fit tests. One-sample Kolmogorov–Smirnov test, AIC test and Q-Q plots were used to find out the best fit.

2.2.5. Spearman’s rho rank correlation

Natural catastrophes or pandemic events (or both) have occurred in past years affecting different lines of business (property insurance, motor own damage) resulting in high correlation between claim developments. Correlation matrix is created from real data and correlations between various underlying risks are assessed. Correlation matrix is calculated by using Spearman’s rho rank correlation (Spearman 1904). The Spearman correlation coefficient is defined as the Pearson correlation coefficient between the ranks. Ranks in the reserving context are calculated from incurred claims in each accident year for each line of business.

2.2.6. Risk aggregation via copula approach

Actuary or reserve risk holder needs to know the volatility best estimate of the company’s portfolio and estimate the value at risk at certain confidence level. In order to obtain a multivariate distribution of an aggregate risk level considering all lines of business, copula approach is used. Next, diversification effect can be calculated as the difference between sums of all risks and aggregated risk from the multivariate distribution.

Fig. 2. One-year best estimate (based on Boumezoued et al. 2011).
Copulas are certain distribution function of a random $d$-vector. Let us recall that the distribution function $H$ of a $d$-dimensional random vector $X = (X_1, \ldots, X_d)'$ is the function defined by

$$H(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X_1 \leq x_1, \ldots, X_d \leq x_d), \ x = (x_1, \ldots, x_d)' \in \mathbb{R}^d.$$ \hspace{1cm} (15)

The distribution function $F_k$ of $X_k$, $k \in \{1, \ldots, d\}$ can be recovered from the distribution function of a random $d$-vector $H$ by $F_k(x_k) = H(\infty, \ldots, \infty, x_k, \infty, \ldots, \infty), x_k \in \mathbb{R}$. This is why $F_1, \ldots, F_d$ are called the univariate margins of $H$ or the marginal distribution functions of $X$. Sklar’s theorem can be used to create copula families from existing families of distribution function of a random $d$-vector. It is a central theorem of copula theory. Proof can be found in Sklar (1996); a probabilistic one in Rüschendorf (2009). For the copula families from existing families of distribution function of a random $d$-vector

The copula $C$ of a random $d$-vector defined by Sklar’s theorem from the multivariate normal distribution $N_d(\mathbf{0}, \mathbf{P})$, where $\mathbf{P}$ is correlation matrix of $X \sim N_d(\mathbf{0}, \mathbf{P})$. If $\Phi_d$ denotes the distribution function of the latter, $C_d^\circ(\mathbf{u})$ is given, for any $\mathbf{u} \in [0, 1]^d$ by

$$C_d^\circ(\mathbf{u}) = \Phi_d(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)) = \int_{-\infty}^{\Phi^{-1}(u_1)} \ldots \int_{-\infty}^{\Phi^{-1}(u_d)} \frac{e^{-(\frac{1}{2} x^T \mathbf{P}^{-1} x)}}{(2\pi)^{\frac{d}{2}} \sqrt{\det \mathbf{P}}} \, dx_1 \ldots dx_d,$$ \hspace{1cm} (16)

where $\Phi^{-1}$ denotes the quantile function of $N(0,1)$ (Hofert et al. 2018).

$t$-copula

The $t$-copula $C_{d,v}^t$ is the distribution function of a random $d$-vector defined by Sklar’s theorem from the multivariate $t$ distribution with location vector $\mathbf{0}$, correlation matrix $\mathbf{P}$ and $v > 0$ degrees of freedom. If $t_{d,v}$ denotes the distribution function of the latter, $C_{d,v}^t(\mathbf{u})$ is given, for any $\mathbf{u} \in [0, 1]^d$, by

$$C_{d,v}^t(\mathbf{u}) = t_{d,v}(t_v^{-1}(u_1), \ldots, t_v^{-1}(u_d)) = \frac{\Gamma\left(\frac{v+d}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{\det \mathbf{P}}} \left(1 + \frac{x^T \mathbf{P}^{-1} x}{v}\right)^{-\frac{v+d}{2}} dx_1 \ldots dx_d,$$ \hspace{1cm} (17)

where $t_v^{-1}$ denotes the quantile function of the univariate Student $t$ distribution with $v$ degrees of freedom (Hofert et al. 2018).

Assume now that the copula $C$ has been selected. We are interested in the value at risk ($VaR$) of a position by using the Monte Carlo method which generates a number $N$ of such scenarios and the sample $\alpha$-quantile is then the one period value at risk at the confidence $\alpha$ defined by:

$$VaR_\alpha = F_{C}^-(\alpha),$$ \hspace{1cm} (18)
where \( L^+ \) is aggregate loss, \( F_{X_k} \) is known as loss distribution and \( \alpha \) denotes the confidence level \( \alpha \in (0, 1) \). The capital needed is the difference between reserve in economical balance sheet \( BE_{\text{total}} \) and the value at risk at the confidence level \( \alpha = 0.995 \):

\[
c_r = |BE_{\text{total}} - VaR_{0.995}|.
\] (21)

2.2.7. Hypothesis tests for selection of copula model

We have used hypothesis tests in order to validate various copula models. Basic graphical diagnostics can be sufficient in practice for finding risk assessment approximation. However, it is not a sufficient argument of internal capital model methodology, documentation package for national regulators, and financial market authorities. Formal statistical tests, which compute \( p \)-values that can help as to guide the choice of the hypothesized copula family, play an important role. We assume this goodness-of-fit issue for adequate parametric copula family amounts formally by testing

\[
H_0: C \in \mathcal{C} \text{ versus } H_1: C \notin \mathcal{C},
\] (22)

where under \( H_0 \) the choice of the hypothesised copula family \( \mathcal{C} \) cannot be rejected and \( H_1 \) states that the choice of the hypothesised copula family \( \mathcal{C} \) can be rejected.

**Parametric bootstrap**

As suggested in Fermanian (2005), Quessy (2005), and Genest and Rémiillard (2008), a natural goodness-of-fit test consists of comparing \( C_n \) with an estimate \( C_{\hat{\theta}_n} \) of \( C \) obtained under the assumption that \( C \in \mathcal{C} \) holds. The estimated margins are used to form the sample

\[
U_{i,n}^j = \left( F_{n,1}(X_{i1}), \ldots, F_{n,d}(X_{id}) \right), i \in \{1, \ldots, n\},
\] (23)

where for any \( j \in \{1, \ldots, d\}, F_j \) is estimated by using component samples of \( X_1, \ldots, X_n \),

\[
F_{n,j}(x) = \frac{1}{n+1} \sum_{i=1}^n 1(X_{ij} < x), x \in \mathbb{R}.
\] (24)

In the previous statement, \( \theta_0 \) is an estimate (parameter vector) of \( \theta \) computed from the pseudo-observations \( U_{1,1}, \ldots, U_{n,n} \) such as the maximum pseudo-likelihood estimator.

We use an approach that appears to perform particularly well according to the large-scale simulations carried out in Genest et al. (2009), where Cramer–von Mises statistic is used for the test fitting:

\[
S_n^{gof} = \int_{[0,1]^d} n \left( C_n(u) - C_{\theta_0}(u) \right)^2 \, dC_n(u) = \sum_{i=1}^n \left( C_n(U_{i,1}) - C_{\theta_0}(U_{i,1}) \right)^2.
\] (25)

An approximate \( p \)-value for the test based on \( S_n^{gof} \) can be obtained by means of a parametric bootstrap whose asymptotic validity is investigated in Genest and Remillard (2008). Advantage of the method is its conceptual simplicity.

**Parametric Bootstrap** algorithm is summarised by Hofert et al. (2018):

1. Compute the pseudo-observations \( U_{1,1}, \ldots, U_{n,n} \).
2. Compute an estimate \( \theta_n \) of \( \theta \) from the pseudo-observations \( U_{1,1}, \ldots, U_{n,n} \).
3. Compute the test statistic \( S_n^{gof} \).
4. For some large integer \( N \), repeat the following steps for every \( k \in \{1, \ldots, N\}: \)
   4.1. Generate a pseudo-random sample \( U_{1,k}^{(k)}, \ldots, U_{n,k}^{(k)} \) from the fitted copula \( C_{\theta_0} \) and compute the corresponding pseudo-observations \( U_{1,k}, \ldots, U_{n,k} \).
   4.2. Compute an estimate \( \theta_n^{(k)} \) of \( \theta \) from the pseudo-observations \( U_{1,k}, \ldots, U_{n,k} \) using the same (rank-based) estimator as in Step 2.
4.3. Compute the corresponding value $S_{n}^{gof,(k)}$ of $S_{n}^{gof}$ as:

$$S_{n}^{gof,(k)} = \sum_{i=1}^{n} (c_{n}^{(k)}(u_{i,n}^{(k)}) - c_{\theta_{n}^{k}}(U_{i,n}))^2,$$

where

$$c_{n}^{(k)}(u) = \frac{1}{n} \sum_{i=1}^{n} 1(u_{i,n}^{(k)} \leq u), \ u \in [0,1]^d.$$

Under $H_0$, $S_{n}^{gof,(k)}$ are approximately independent copies of $S_{n}^{gof}$.  

5. An approximate $p$-value for the test is given by

$$\left(\frac{1}{2} + \sum_{k=1}^{N} 1(S_{n}^{gof,(k)} \geq S_{n}^{gof})\right) / (N + 1).$$

Cross-validation criterion

There is a possibility that all candidate parametric copula families are rejected when the sample size is large or none of the families is rejected when the sample size is small. The test that uses *Akaike information criterion* (AIC) and performs the selection of the best ranked family can be justified by using formula

$$AIC = 2(l_{n,max} - m),$$

where $l_{n,max}$ is the maximized likelihood function and $m$ is the total number of marginal and copula parameters. Grønneberg and Hjort (2014) have defined cross-validation copula information criterion up to a multiplicative constant, the first-order equivalent of the cross validation criterion:

$$\widehat{x}_{n} = \frac{1}{n} \sum_{i=1}^{n} \log \theta_{n-1}(F_{n-1}(X_i)).$$

where $\theta_{n-1}$ is the maximum pseudo-likelihood estimate computed from the sample $X_1, \ldots, X_{n-1}, X_{n+1}, \ldots, X_n$ and

$$F_{n-1}(x) = (F_{n,1-1}(x_1), \ldots, F_{n,d-1}(x_d)), \ x \in \mathbb{R}^d,$$

with

$$F_{n,j-1}(x) = \begin{cases} \frac{1}{n} \sum_{k=1, k \neq j}^{n} 1(X_{k,j} \leq x), & \text{if } x \geq \min X_{k,j}, k \in \{1, \ldots, n\}\setminus\{j\} \\ \frac{1}{n}, & \text{otherwise}. \end{cases}$$

This test leaves out and penalises copula families with too many parameters that tend to overfit. Several authors have produced papers with the aim of improving the AIC formula approach and historical development of the copula theory in a more detailed way, for instance Claeskens and Hjort (2011), Grønneberg and Hjort (2014), Jordanger and Tjøstheim (2014), McNeil et al. (2015).

3. PRACTICAL APPROACH AND ALGORITHM OF INTERNAL MODEL

This section describes the basis of each simulation and analytical techniques, as well as provides relevant, primary reference papers. The algorithm of calculation is demonstrated in Fig. 3.
Step 1 – Data collection

We have collected incurred Baltic claim data sets from past nine years (2012–2020), including accident years and development years. The data set also includes pandemic trends, which had an impact on economics and customer behaviours. The model has been applied to data sets of three lines of business – motor third party liability insurance (MTPL), general third-party liability insurance (GTPL) and credit and suretyship insurance (C&S). The data set is in line with CL using formula (7) for general third party liability business line. An example can be seen in Fig. 4.

Step 2 – Calculation of correlation matrix

We have assessed Spearman correlations in occurred claims. Calculations are performed based on incurred claims for each accident year as end of 2020 and the average rank of each year (Annex 2). Figure 5 demonstrates the differences in correlations between lines of business according to the internal and the non-internal or standard formula model.
Step 3 – Assessment if claims follow distributions

We have used R package ChainLadder (Gesmann et al. 2015) and its key functions CDR (calculates standard deviation of the claims development result after one year), BootChainLadder for real non-life data sets, for results derived in Table 1. Then the obtained one-year potential best estimate is tested to examine whether it follows certain distribution by using R package MASS (Venables and Ripley 2002). Probability distribution that real data follows, it’s histograms, theoretical densities and numerical results of hypothesis tests, and Q-Q plots can be seen in Annex 1. We have chosen three different distributions as combination of numerical results and Q-Q plots (Annex 1), and applied such distributions and parameters as seen in Table 1.

Step 4 – Copula simulation and choice of model by applying hypothesis tests

We have used R package copula (Hofert et al. 2020, Yan 2007, Kojadinovic and Yan 2010, Hofert and Maechler 2011) and its functions rCopula (for normal copula) and tcopula. And we have used R package gofCopula (Zhang et al. 2016, Genest et al. 2009) with its key functions xvCopula and gofcopula for goodness-of-fit tests.

### Table 1. Unique distributions, parameters and claims best estimate

<table>
<thead>
<tr>
<th>Distribution</th>
<th>MTPL Log-normal</th>
<th>GTPL Normal</th>
<th>C&amp;S Weibull</th>
</tr>
</thead>
<tbody>
<tr>
<td>meanlog/scale</td>
<td>15.93</td>
<td>2 859 768</td>
<td>3 115 480</td>
</tr>
<tr>
<td>sddl/log/shape</td>
<td>0.16</td>
<td>613 643.50</td>
<td>4.60</td>
</tr>
<tr>
<td>CBE, Best estimate in EUR</td>
<td>8 352 978</td>
<td>2 859 768</td>
<td>1 180 261</td>
</tr>
<tr>
<td>σ standard deviation for standard formula’s model provided by regulator</td>
<td>0.16</td>
<td>0.22</td>
<td>0.91</td>
</tr>
</tbody>
</table>
Table 2. Capital requirement results and capital gains in euros

<table>
<thead>
<tr>
<th>Model/Approach</th>
<th>VaR 99.5%</th>
<th>Reserve risk, $C_r$</th>
<th>Capital savings</th>
<th>Capital gains in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard model provided by regulator</td>
<td>19 565 636</td>
<td>7 172 629</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal model using normal copula</td>
<td>16 665 854</td>
<td>4 272 847</td>
<td>2 899 781</td>
<td>40.43</td>
</tr>
<tr>
<td>Internal model using t-copula, df = 4</td>
<td>16 734 212</td>
<td>4 341 205</td>
<td>2 831 424</td>
<td>39.48</td>
</tr>
</tbody>
</table>

Table 3. Goodness of fit, numerical results of hypothesis testing

<table>
<thead>
<tr>
<th>Model/Approach</th>
<th>AIC and p-value</th>
<th>Parametric bootstrap</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal model using normal copula</td>
<td>0.02 and 0.8027</td>
<td>5.47</td>
<td>Plausible, $H_0$ cannot be rejected</td>
</tr>
<tr>
<td>Internal model using t-copula, df = 4</td>
<td>0.21 and 0.0005</td>
<td>-623</td>
<td>$H_0$ is rejected</td>
</tr>
</tbody>
</table>

4. CASE STUDY RESULTS AND CAPITAL GAINS

Capital requirement for reserve risk with standard model is 7 172 629 euros, with internal model using normal copula 4 272 847 euros; and using t-copula it is 4 341 205 euros. Capital savings by comparing the proposed internal models versus the standard model are 40.43% with normal copula and 39.48% with t-copula (Table 2). However, t-copula (number of degrees of freedom (df) is 4) as a copula family must be rejected based on hypothesis tests (Table 3).

5. CONCLUSIONS AND FURTHER STUDY

A more accurate risk quantification model has been developed than the standard model provided by the EU regulator under the Solvency II framework. The proposed model provides capital cost gains as well. The case study has shown that capital needs in case of non-internal model application can vary up to forty percent. The proposed model can be used not only as an internal model but also as an individual tool by each and every insurer for their own risk assessment and financial planning process. Real data set provided capital gains, but for others it could indicate capital shortage. The provided full algorithm can be easily adjusted depending on entity data specifics; more dimensions (LoB) can be added and other claim reserving methods applied; for instance, methods based on neural network architecture or machine learning. Next, these new claim reserving methods can be used for measuring digitalization and automatic claim handling impact on required capital for reserve risk. These are also further research steps.

ACKNOWLEDGEMENT

The publication costs of this article were partially covered by the Estonian Academy of Sciences.
ANNEX 1.

Claim distributions, histogram and theoretical densities, Q-Q plots for MTPL LoB (up), GTPL LoB (middle), C&S LoB (down).

<table>
<thead>
<tr>
<th>ANNEX 1.</th>
<th>MTPL Test: Two-sample nonparametric multivariate test of means</th>
<th>MTPL Test: One-sample Kolmogorov–Smirnov test</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAMMA</td>
<td>309 900</td>
<td>1,00E+00</td>
</tr>
<tr>
<td>WEIBULL</td>
<td>311 348</td>
<td>1,00E+00</td>
</tr>
<tr>
<td>NORMAL</td>
<td>310 124</td>
<td>1,57E-09</td>
</tr>
<tr>
<td>LOGNORMAL</td>
<td>309 938</td>
<td>4,56E-01</td>
</tr>
<tr>
<td>EXPONENTIAL</td>
<td>311 348</td>
<td>1,00E+00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANNEX 1.</th>
<th>GTPL Test: Two-sample nonparametric multivariate test of means</th>
<th>GTPL Test: One-sample Kolmogorov–Smirnov test</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAMMA</td>
<td>294 593</td>
<td>2,20E-16</td>
</tr>
<tr>
<td>WEIBULL</td>
<td>295 827</td>
<td>2,20E-16</td>
</tr>
<tr>
<td>NORMAL</td>
<td>294 675</td>
<td>2,20E-16</td>
</tr>
<tr>
<td>LOGNORMAL</td>
<td>295 101</td>
<td>3,00E-14</td>
</tr>
<tr>
<td>EXPONENTIAL</td>
<td>295 827</td>
<td>2,20E-16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANNEX 1.</th>
<th>SUREITY Test: Two-sample nonparametric multivariate test of means</th>
<th>SUREITY Test: One-sample Kolmogorov–Smirnov test</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAMMA</td>
<td>291 093</td>
<td>2,20E-16</td>
</tr>
<tr>
<td>WEIBULL</td>
<td>291 252</td>
<td>2,20E-16</td>
</tr>
<tr>
<td>NORMAL</td>
<td>296 848</td>
<td>2,20E-16</td>
</tr>
<tr>
<td>LOGNORMAL</td>
<td>291 585</td>
<td>2,20E-16</td>
</tr>
<tr>
<td>EXPONENTIAL</td>
<td>291 252</td>
<td>2,20E-16</td>
</tr>
</tbody>
</table>

Continued on the next page
MTPL: Log-normal (chosen based on numerical results of hypothesis tests and Q-Q plot)

GTPL: Gamma

MTPL: Normal

MTPL: Gamma

GTPL: Normal (chosen based on numerical results of hypothesis tests)

GTPL: Log-normal

Surety: Gamma

MTPL: Log-normal (chosen based on numerical results of hypothesis tests and Q-Q plot)

Surety: Weibull (chosen based on numerical results of hypothesis test and Q-Q plot)

Surety: Log-normal

ANNEX 1. Continued.
REFERENCES


ANNEX 2.

Incurred claim and the calculated average ranks 2012–2020.

<table>
<thead>
<tr>
<th>Incurred claim for each year</th>
<th>The average rank for each year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MTPL</td>
</tr>
<tr>
<td>2012</td>
<td>22 880 869</td>
</tr>
<tr>
<td>2013</td>
<td>25 643 495</td>
</tr>
<tr>
<td>2014</td>
<td>28 824 391</td>
</tr>
<tr>
<td>2015</td>
<td>39 491 919</td>
</tr>
<tr>
<td>2016</td>
<td>39 879 874</td>
</tr>
<tr>
<td>2017</td>
<td>36 985 472</td>
</tr>
<tr>
<td>2018</td>
<td>43 412 120</td>
</tr>
<tr>
<td>2019</td>
<td>47 139 227</td>
</tr>
<tr>
<td>2020</td>
<td>40 975 775</td>
</tr>
</tbody>
</table>
Efektiivne kapitali juhtimine sisemudeli abil: kahjukindlustuse juhtum

Ilze Zariņa-Cīrule, Gaida Pettere ja Irina Voronova

Artiklis on välja töötatud koopulate teoorial põhinev sisemudel kindlustusfirma reservide hindamiseks lähtudes Solventsus II regulatsioonides määratud nõuetest. Mudelit on rakendatud kindlustusfirma kahjuandmetele, mis on esitatud ahel-redel meetodit järgides kolme erineva kahjude klassi korral. Leitud on parimad lähendavad jaotused iga kahjude klassi kirjeldamiseks ja need on ühendatud kolmemõõtmeliseks jaotuseks Gaussi koopula abil. Viimane andis kahjudele adekvaatema kirjelduse kui t-koopulal põhinev mudel. Saadud mudeli abil hinnatud vajalike reservide suurus on tunduvalt väiksem kui standardse metoodika abil leitud reservide kogusumma.