Peer-reviewed version available at *Risks* **2018**, *6*, 108; doi:10.3390/risks6040108

Article

A Robust General Multivariate Chain Ladder Method

Kris Peremans [1,†], Stefan Van Aelst [1,†] and Tim Verdonck [1,†,*]

[1] Department of Mathematics, KU Leuven, Celestijnenlaan 200B, 3001 Leuven, Belgium
* Correspondence: tim.verdonck@kuleuven.be; Tel.: +32-16-32-27-89
† These authors contributed equally to this work.

Abstract: The chain ladder method is a popular technique to estimate the future reserves needed to handle claims that are not fully settled. Since the predictions of the aggregate portfolio (consisting of different subportfolios) in general differ from the sum of the predictions of the subportfolios, a general multivariate chain ladder (GMCL) method has already been proposed. However, the GMCL method is based on the seemingly unrelated regression (SUR) technique which makes it very sensitive to outliers. To address this issue a robust alternative is introduced which estimates the SUR parameters in a more outlier resistant way. With the robust methodology it is possible to detect which claims have an abnormally large influence on the reserve estimates. We introduce a simulation design to generate artificial multivariate run-off triangles based on the GMCL model and illustrate the importance of taking into account contemporaneous correlations and structural connections between the run-off triangles. By adding contamination to these artificial datasets, the sensitivity of the traditional GMCL method and the good performance of the robust GMCL method is shown. From the analysis of a portfolio from practice it is clear that the robust GMCL method can provide better insight in the structure of the data.

Keywords: Claims reserving; Contemporaneous correlations; Outliers; Robust MM-estimators; Seemingly unrelated regression

---

1. Introduction

Stochastic claims reserving in non-life insurance, also known as general insurance in the UK or property and casualty insurance in the US, is an important and challenging discipline for actuaries. Since the claims settlement in non-life insurance may last several years, e.g. due to long legal procedures or difficulties in determining the size of the claim, insurers have to build up reserves enabling them to handle the liabilities related to current insurance contracts. These outstanding claims reserves are often the largest position on the liability side of the balance sheet of a non-life insurance company.

With the introduction of new regulatory guidelines for the insurance business (e.g. Solvency II in Europe) there is a growing awareness that advanced statistical techniques should be used for forecasting the future claims payments. A comprehensive discussion on the Solvency II directive and its implications may be found in Dreksler et al. (2015).

A well-known and widely used technique to forecast future claims is the chain ladder method, a deterministic algorithm which estimates the future claims recursively using a set of development factors. To include a stochastic component, this simple technique can be embedded into the statistical framework of generalized linear models (GLM), introduced by Nelder and Wedderburn (1972). The relationship between the deterministic chain ladder method and various stochastic models based on GLMs is discussed in England and Verrall (2002) and Wüthrich and Merz (2008) for instance.

In practice, a non-life insurance company subdivides portfolios into several correlated subportfolios, such that each subportfolio, presented in the form of a run-off triangle, satisfies certain homogeneity properties. The chain ladder method is then typically applied to the different single run-off triangles, ignoring the contemporaneous correlations between these various subportfolios. It is well known that the chain ladder predictions for the aggregate portfolio, which consists of the sum of the different subportfolios, is in general different from the sum of the chain ladder predictions for each of the separate subportfolios (Ajne 1994). To address this issue the claims reserving problem is also studied in a multivariate context to cope with the problem of dependence between different subportfolios. Braun (2004) studied the bivariate model which takes into account the correlation between two subportfolios of an aggregate portfolio. Merz and Wüthrich (2007) consider claims reserving for a portfolio consisting of N correlated run-off triangles. Pröhl and Schmidt (2005) and Schmidt (2006) proposed a multivariate chain ladder (MCL) model where they deduced multivariate chain ladder predictors that take into account the dependency between the different subportfolios. These predictors are shown to satisfy a classical optimality criterion. Moreover, it is explained how multivariate methods solve the lack of additivity of the chain ladder predictions. Multivariate methods also have the advantage that we can learn something about the behavior of several subportfolios by observing another subportfolio. Merz and Wüthrich (2008) further discussed the conditional mean squared error of prediction (MSEP) for the MCL model.

Recently, Zhang (2010) proposed a general multivariate chain ladder (GMCL) model that further extends the MCL model by including intercepts to improve model adequacy. The parameters of this flexible model are estimated using the seemingly unrelated regression (SUR) framework. The SUR model (Zellner 1962) is a generalization of a linear regression model which consists of more than one equation and where the error terms of these equations are contemporaneously correlated. SUR models have found considerable use in many applications in econometrics, finance and insurance. Taking into account the contemporaneous correlations among different portfolios may lead to more accurate uncertainty assessments. Another advantage is that also structural relationships between triangles where the development of one triangle depends on past losses from other triangles can be included in the GMCL model. The GMCL model also allows joint development of the paid and incurred losses from multiple business lines. The similarity and difference between the GMCL model on bivariate data and the Munich chain ladder model (Quarg and Mack 2004) are discussed by Zhang (2010), who also shows that several existing multivariate claims reserving estimators can find their equivalent in the SUR estimator family.

To estimate the parameters in a SUR model, one typically uses the feasible generalized least squares (FGLS) estimator (Zellner 1962)), which takes into account the covariance structure of the errors. Since FGLS is based on the classical covariance matrix and ordinary least squares estimation, using FGLS makes the SUR estimates and thus in particular the GMCL estimates very sensitive to outliers. Outliers are observations that differ from the majority of the data and it is well known that these atypical observations can have a large impact on traditional statistical methods. On the other hand, robust methods provide estimates for the claim provisions which resemble the classical estimates that would have been obtained if there were no outliers in the data, while they do not model the outlier generating process. As a consequence of fitting the majority of the data well, robust methods also provide a reliable method to detect outliers. Observations which are flagged as outliers can then be examined in detail by experts to understand their origin. In Koenker and Portnoy (1990) a robust SUR estimator is proposed based on M-estimators. Since this procedure is not affine equivariant and does not take full account of the multivariate nature of the problem, a method based on S-estimators was introduced in Bilodeau and Duchesne (2000). This robust SUR estimator is regression and affine equivariant, but is computationally expensive. Therefore, Hubert et al. (2017) proposed the FastSUR algorithm, which implements the ideas of the FastS algorithm (Salibian-Barrera and Yohai 2006) for the SUR S-estimator. Recently, Peremans and Van Aelst (2018) developed robust inference for the SUR model based on MM-estimators.

This paper is structured as follows. A review of the GMCL model of Zhang (2010) is given in Section 2. In Section 3 the GMCL model is formulated in the SUR framework and the FGLS estimator is introduced. Section 4 describes robust MM-estimators for estimating the parameters in GMCL models and its numerical algorithm for computation. We then show the good performance of these estimators in an extensive simulation study in Section 5. In Section 6 the robust procedure is illustrated on a real dataset from a non-life business line. Some concluding remarks and potential directions for further research are given in Section 7. The Appendix contains the parameter estimates obtained from the GMCL models for the real dataset.

2. General Multivariate Chain Ladder Model

We assume that the non-life insurance company needs to handle $M \geq 1$ subportfolios. Let I and K denote the final accident and development period respectively. For $1 \leqslant i \leqslant I$, $1 \leqslant k \leqslant K$ and $1 \leqslant m \leqslant M$ denote $C_{i,k}^{(m)}$ as the cumulative claims amount of accident period i and development period k of subportfolio m. Depending on the size of K, one refers to long or short tail business and for simplicity we take $K = I$.

At time I we have observed the claims $C_{i,k}^{(m)}$ with $i + k - 1 \leqslant I$ for every subportfolio m. Typically, a subportfolio m is then presented in the form of a run-off triangle as illustrated in Table 1. This triangle

Table 1. Typical representation of subportfolio m as a run-off triangle.

accident	**development period** k						
period i	1	2	…	k	…	$I-1$	I
1	$C_{1,1}^{(m)}$	$C_{1,2}^{(m)}$	…	$C_{1,k}^{(m)}$	…	$C_{1,I-1}^{(m)}$	$C_{1,I}^{(m)}$
2	$C_{2,1}^{(m)}$	$C_{2,2}^{(m)}$	…	$C_{2,k}^{(m)}$	…	$C_{2,I-1}^{(m)}$	
⋮	…	…	…	…	…		
i	$C_{i,1}^{(m)}$	$C_{i,2}^{(m)}$	…	$C_{i,k}^{(m)}$			
⋮	…	…	…				
$I-1$	$C_{I-1,1}^{(m)}$	$C_{I-1,2}^{(m)}$					
I	$C_{I,1}^{(m)}$						

structure shows the development of claims for each accident period. Usually yearly, quarterly or monthly periods are used. The columns represent the development periods whereas the diagonals present payments in the same calendar period. The overall outstanding reserve R that will need to be paid in future, is defined as

$$R = \sum_{m=1}^{M} \sum_{i=2}^{I} \left(C_{i,I}^{(m)} - C_{i,I-i+1}^{(m)} \right),$$

and depends on the ultimate claim values $C_{i,I}^{(m)}$. The aim of claims reserving is then to complete the run-off triangles into squares, i.e. forecasting the future claims in the bottom right corner of the run-off triangles in order to estimate the overall outstanding reserves.

Let $\boldsymbol{C}_{i,k} = (C_{i,k}^{(1)}, \ldots, C_{i,k}^{(M)})'$ denote the vector of cumulative claims of accident period i and development period k. Consider the following model structure from development period k to $k+1$:

$$\boldsymbol{C}_{i,k+1} = \boldsymbol{A}_k + \boldsymbol{B}_k \boldsymbol{C}_{i,k} + \boldsymbol{\epsilon}_{i,k}, \tag{1}$$

for $i = 1, \ldots, I-k$, where $\boldsymbol{A}_k$ is the M vector containing intercepts $\beta_{10}, \ldots, \beta_{M0}$, and where $\boldsymbol{B}_k$ is the corresponding $M \times M$ development matrix that contains the development parameters $\beta_{m1}, \ldots, \beta_{mM}$ for run-off triangle m in row m. Moreover, $\boldsymbol{\epsilon}_{i,k} = (\epsilon_{i,k}^{(1)}, \ldots, \epsilon_{i,k}^{(M)})'$ are independent and symmetrically distributed errors. For a non-diagonal development matrix $\boldsymbol{B}_k$, the model allows the development of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0556.v1

Peer-reviewed version available at Risks 2018, 6, 108; doi:10.3390/risks6040108

one run-off triangle in development period k to depend on the claims in the other run-off triangles at development period k. Moreover, it is assumed that the errors $\boldsymbol{\epsilon}_{i,k}$ satisfy

$$\mathrm{E}(\boldsymbol{\epsilon}_{i,k}|\mathcal{D}_{i,k}) = 0 \tag{2}$$

$$\mathrm{Cov}(\boldsymbol{\epsilon}_{i,k}|\mathcal{D}_{i,k}) = \mathrm{diag}(\boldsymbol{C}_{i,k})^{1/2}\boldsymbol{\Sigma}_k\,\mathrm{diag}(\boldsymbol{C}_{i,k})^{1/2}, \tag{3}$$

where $\mathcal{D}_{i,k} = \{\boldsymbol{C}_{i,j}|j \leqslant k\}$, the set of cumulative claims for accident period i up to and including development period k, $\boldsymbol{\Sigma}_k$ is a symmetric positive definite $M \times M$ matrix, and diag is the operator that turns its argument(s) into a diagonal matrix. Consequently, for a non-diagonal matrix $\boldsymbol{\Sigma}_k$ the components of the error terms $\boldsymbol{\epsilon}_{i,k}$ are allowed to be correlated. Equations (1), (2), and (3) for $k = 1, \ldots, I-1$ constitute the general multivariate chain ladder model as proposed in Zhang (2010). A separate chain ladder (SCL) model can be obtained as a special case by taking $\boldsymbol{A}_k$ the zero vector, and by imposing that $\boldsymbol{B}_k$ and $\boldsymbol{\Sigma}_k$ are diagonal matrices. The advantages of the GMCL model over already existing models like SCL are evident (Zhang 2010). The parameters $\boldsymbol{A}_k$, $\boldsymbol{B}_k$ and $\boldsymbol{\Sigma}_k$ are unknown model parameters and need to be estimated from historic claims.

3. Seemingly Unrelated Regression

In Zhang (2010) the model structure from development period k to $k+1$, given in equation (1) for $i = 1, \ldots, I-k$, has been rewritten as a multiple linear regression model. Omitting the dependence on k, the following system of equations is obtained:

$$\begin{pmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_M \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \boldsymbol{X}_M \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_M \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \vdots \\ \boldsymbol{\varepsilon}_M \end{pmatrix}, \tag{4}$$

where for $m = 1, \ldots, M$ and $n = I-k$ it holds that

- $\boldsymbol{y}_m = (C^{(m)}_{1,k+1}, \ldots, C^{(m)}_{n,k+1})'$ is the n vector of all observed losses at development period $k+1$ from triangle m;
- $\boldsymbol{X}_m = ((1, \boldsymbol{C}'_{1,k})', \ldots, (1, \boldsymbol{C}'_{n,k})')'$ is the $n \times (M+1)$ matrix of the first n observations at development period k from each triangle, including the constant 1 for the intercept. Hence, $\boldsymbol{X}_1 = \ldots = \boldsymbol{X}_M$;
- $\boldsymbol{\beta}_m = (\beta_{m0}, \ldots, \beta_{mM})'$ is the $(M+1)$ vector of development parameters, including the intercept;
- $\boldsymbol{\varepsilon}_m = (\epsilon^{(m)}_{1,k+1}, \ldots, \epsilon^{(m)}_{n,k+1})'$ is the n vector of error terms.

From (2) and (3) it follows that

$$\mathrm{Cov}(\boldsymbol{\varepsilon}) = \mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \mathrm{diag}(\boldsymbol{V}_k)^{1/2}(\boldsymbol{\Sigma}_k \otimes \boldsymbol{I}_n)\,\mathrm{diag}(\boldsymbol{V}_k)^{1/2},$$

where $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}'_1, \ldots, \boldsymbol{\varepsilon}'_M)'$, and $\boldsymbol{V}_k = (\boldsymbol{V}^{(1)\prime}_k, \ldots, \boldsymbol{V}^{(M)\prime}_k)'$ with $\boldsymbol{V}^{(m)}_k = (C^{(m)}_{1,k}, \ldots, C^{(m)}_{n,k})'$ for $m = 1, \ldots, M$. Moreover, $\boldsymbol{I}_n$ is the identity matrix of size n and $\otimes$ represents the Kronecker product.

Pre-multiplying both sides of equation (4) by $\mathrm{diag}(\boldsymbol{V}_k)^{-1/2}$ leads to the following linear regression model

$$\begin{pmatrix} \boldsymbol{y}^*_1 \\ \vdots \\ \boldsymbol{y}^*_M \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}^*_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \boldsymbol{X}^*_M \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_M \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}^*_1 \\ \vdots \\ \boldsymbol{\varepsilon}^*_M \end{pmatrix}, \tag{5}$$

where $\boldsymbol{y}^*_m = \mathrm{diag}(\boldsymbol{V}^{(m)}_k)^{-1/2}\boldsymbol{y}_m$, $\boldsymbol{X}^*_m = \mathrm{bdiag}(\boldsymbol{V}^{(m)}_k)^{-1/2}\boldsymbol{X}_m$, and $\boldsymbol{\varepsilon}^*_m = \mathrm{diag}(\boldsymbol{V}^{(m)}_k)^{-1/2}\boldsymbol{\varepsilon}_m$. Note that now the $n \times (M+1)$ matrices $\boldsymbol{X}^*_m$ are different for each equation, i.e. $\boldsymbol{X}^*_m \neq \boldsymbol{X}^*_{m'}$ for $m \neq m'$. Moreover, denote $\boldsymbol{\varepsilon}^* = (\boldsymbol{\varepsilon}^{*\prime}_1, \ldots, \boldsymbol{\varepsilon}^{*\prime}_M)'$, then for the representation of the GMCL model given in (5)

Peer-reviewed version available at *Risks* **2018**, *6*, 108; doi:10.3390/risks6040108

the error covariance matrix $\text{Cov}(\boldsymbol{\varepsilon}^*)$ is consistent with the SUR assumption of contemporaneous correlation (Zellner 1962):

$$\text{Cov}(\boldsymbol{\varepsilon}^*) = \text{diag}(\boldsymbol{V}_k)^{-1/2}\,\text{Cov}(\boldsymbol{\varepsilon})\,\text{diag}(\boldsymbol{V}_k)^{-1/2} = \boldsymbol{\Sigma}_k \otimes \boldsymbol{I}_n.$$

Hence, it is straightforward to estimate the development parameters by using estimators for SUR models on the transformed data.

Consider the estimation of the unknown development parameters $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_M')'$ under the SUR model given in (5). The equations in this model can be considered as M separate linear regression models of the form

$$\boldsymbol{y}_m^* = \boldsymbol{X}_m^*\boldsymbol{\beta}_m + \boldsymbol{\varepsilon}_m^*, \tag{6}$$

for $m = 1, \ldots, M$. Then each linear regression model can be estimated separately by least squares (LS). However, this method may yield inefficient estimates since it ignores the correlation structure in the error terms. Generalized least squares (GLS) is a modification of least squares that can deal with any type of correlation. In this context, the GLS estimator for the model in (5) becomes

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^{*\prime}(\boldsymbol{\Sigma}_k^{-1} \otimes \boldsymbol{I}_n)\boldsymbol{X}^*)^{-1}\boldsymbol{X}^{*\prime}(\boldsymbol{\Sigma}_k^{-1} \otimes \boldsymbol{I}_n)\boldsymbol{y}^*, \tag{7}$$

where $\boldsymbol{X}^* = \text{diag}(\boldsymbol{X}_1^*, \ldots, \boldsymbol{X}_M^*)$, a block diagonal matrix of size $nM \times M(M+1)$, and $\boldsymbol{y}^* = (\boldsymbol{y}_1^{*\prime}, \ldots, \boldsymbol{y}_M^{*\prime})'$. GLS produces efficient estimators (Zellner 1962). However, since $\boldsymbol{\Sigma}_k$ is unknown a feasible GLS (FGLS) estimator is usually introduced. FGLS replaces the unknown matrix $\boldsymbol{\Sigma}_k$ in (7) with $\hat{\boldsymbol{\Sigma}}_k = (\hat{\boldsymbol{\varepsilon}}_1^*, \ldots, \hat{\boldsymbol{\varepsilon}}_M^*)'(\hat{\boldsymbol{\varepsilon}}_1^*, \ldots, \hat{\boldsymbol{\varepsilon}}_M^*)/n$, where $\hat{\boldsymbol{\varepsilon}}_m^*$ are the residuals obtained from estimating (6) by least squares. The efficiency of FGLS is in general smaller than for GLS, although the asymptotic efficiency of both methods is identical. Note that this two-step procedure can be iterated until convergence of the development parameter estimates. After estimating the development parameters $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_M')'$ or equivalently the development matrix $(\boldsymbol{A}_k, \boldsymbol{B}_k) = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_M)'$ using the LS or the FGLS estimation procedure consecutively for all development periods $k = 1, \ldots, I-1$, the bottom right corner of the run-off triangles can be predicted and the overall reserve estimate $\hat{R}$ can be obtained (for all M triangles simultaneously).

4. Robust GMCL Method

In the univariate setting ($M = 1$) Verdonck and Debruyne (2011) have demonstrated that the chain ladder method is very sensitive to outliers. Several robust alternatives have already been developed in the univariate claims reserving framework (see e.g. Brazauskas et al. (2009), Brazauskas (2009), Verdonck et al. (2009), Verdonck and Van Wouwe (2011), Pitselis et al. (2015) and Peremans et al. (2017)). Even one outlier can lead to a huge over- or underestimation of the overall reserve estimate. Moreover, Hubert et al. (2017) have shown that FGLS estimators in the GMCL model are also not robust and that an outlier in one of the run-off triangles may also affect the estimates of future claims in the other run-off triangles. Note that the multivariate aspect makes the task of outlier detection more challenging because outliers can be univariate or multivariate. Multivariate outliers are observations that deviate from the multivariate pattern indicated by the majority of the observations, i.e. inconsistent with the covariance structure of the dataset, but in contrast to univariate outliers are not necessarily extreme along a single coordinate (a single run-off triangle). Therefore, univariate outlier detection methods may fail to find these outliers and it is important to rely on robust multivariate alternatives. When we combine robust SUR methods with the GMCL model, we obtain robust reserve estimates and diagnostics for outlier detection.

We now introduce MM-estimators for the SUR model in (5) as studied by Peremans and Van Aelst (2018). The system of equations in (5) can be rewritten as another linear regression model by reordering

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0556.v1
Peer-reviewed version available at Risks 2018, 6, 108; doi:10.3390/risks6040108

the equations. Let $\mathcal{Y}_i^*$, $\mathcal{X}_i^*$ and $\boldsymbol{e}_i^*$ be the subvector or submatrix of $\boldsymbol{y}^*$, $\boldsymbol{X}^*$ and $\boldsymbol{\varepsilon}^*$ respectively by extracting rows $i, i+n, \ldots, i+n(M-1)$. Then the system of equations in (5) is equivalent to

$$\mathcal{Y}_i^* = \mathcal{X}_i^* \boldsymbol{\beta} + \boldsymbol{e}_i^*, \tag{8}$$

for $i = 1, \ldots, n$. In this case we easily obtain that $\text{Cov}(\boldsymbol{e}_i^*) = \boldsymbol{\Sigma}_k$. Decompose the covariance matrix $\boldsymbol{\Sigma}_k$ into a shape component $\boldsymbol{\Gamma}_k$ and a scale parameter σ_k such that $\boldsymbol{\Sigma}_k = \sigma_k^2 \boldsymbol{\Gamma}_k$ with $|\boldsymbol{\Gamma}_k| = 1$. Here $|\boldsymbol{A}|$ denotes the determinant of the matrix $\boldsymbol{A}$. Since we assume that $\boldsymbol{\Sigma}_k$ is positive definite, such a decomposition always exists. Let $\boldsymbol{e}_i^*(\boldsymbol{b})$ be equal to $\mathcal{Y}_i^* - \mathcal{X}_i^* \boldsymbol{b}$ for any $M(M+1)$ vector $\boldsymbol{b}$ according to the SUR representation in (8). Then, given an initial estimator of the scale $\hat{\sigma}_k$, the MM-estimators $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Gamma}}_k)$ minimize

$$\frac{1}{n} \sum_{i=1}^{n} \rho \left(\frac{\sqrt{\boldsymbol{e}_i^*(\boldsymbol{b})' \boldsymbol{G}^{-1} \boldsymbol{e}_i^*(\boldsymbol{b})}}{\hat{\sigma}_k} \right),$$

over all $M(M+1)$ vectors $\boldsymbol{b}$ and positive definite symmetric $M \times M$ matrices $\boldsymbol{G}$ with $|\boldsymbol{G}| = 1$. The MM-estimator for covariance is defined as $\hat{\boldsymbol{\Sigma}}_k = \hat{\sigma}_k^2 \hat{\boldsymbol{\Gamma}}_k$. The function ρ should satisfy the following conditions:

- ρ is symmetric, twice continuously differentiable and satisfies $\rho(0) = 0$;
- ρ is strictly increasing on $[0, c]$ and constant on $[c, \infty[$ for some $c > 0$.

Evidently, taking $\rho(x) = x^2$ fulfills the conditions and yields the iterated FGLS estimator. To be robust against outliers, it is necessary to consider bounded ρ functions. The most popular family of ρ functions for MM-estimators is the class of Tukey bisquare ρ functions given by $\rho(x) = \min(x^2/2 - x^4/2c^2 + x^6/6c^4, c^2/6)$. The tuning parameter $c > 0$ is usually chosen to obtain a certain level of asymptotic efficiency under the SUR model with normally distributed errors.

MM-estimators require an initial estimator of scale $\hat{\sigma}_k$. In order for MM-estimators to be robust, also this scale estimator should be robust. Therefore, highly robust S-estimators are computed to obtain a highly robust scale estimator. S-estimators have been introduced for SUR models in Bilodeau and Duchesne (2000), and a computational efficient algorithm has been proposed in Hubert et al. (2017). Robustness can be measured by the breakdown point of an estimator, which is roughly equal to the maximal fraction of contaminated observations that an estimator can tolerate before its bias becomes unbounded. For MM-estimators the breakdown point can be up to 50%. In this paper we have tuned the MM-estimators to have a 25% breakdown point and 95% normal efficiency, which is commonly considered to be a good compromise between robustness and precision of the estimator.

MM-estimators do not have explicit solutions, although they satisfy a similar set of equations as the FGLS estimators given in (7). Indeed, the MM-estimators $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_k)$ satisfy the following set of equations

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^{*\prime}(\hat{\boldsymbol{\Sigma}}_k^{-1} \otimes \boldsymbol{D}_k)\boldsymbol{X}^*)^{-1} \boldsymbol{X}^{*\prime}(\hat{\boldsymbol{\Sigma}}_k^{-1} \otimes \boldsymbol{D}_k)\boldsymbol{y}^*$$

$$\hat{\boldsymbol{\Sigma}}_k = M(\boldsymbol{e}_1^*(\hat{\boldsymbol{\beta}}), \ldots, \boldsymbol{e}_n^*(\hat{\boldsymbol{\beta}}))\boldsymbol{D}_k(\boldsymbol{e}_1^*(\hat{\boldsymbol{\beta}}), \ldots, \boldsymbol{e}_n^*(\hat{\boldsymbol{\beta}}))' \left(\sum_{i=1}^{n} \rho'(d_i) d_i \right)^{-1}$$

with $\boldsymbol{D}_k = \text{diag}(w(d_1), \ldots, w(d_n))$ where $w(x) = \rho'(x)/x$, $d_i^2 = \boldsymbol{e}_i^*(\hat{\boldsymbol{\beta}})' \hat{\boldsymbol{\Sigma}}_k^{-1} \boldsymbol{e}_i^*(\hat{\boldsymbol{\beta}})$, and $\boldsymbol{e}_i^*(\hat{\boldsymbol{\beta}}) = \mathcal{Y}_i^* - \mathcal{X}_i^* \hat{\boldsymbol{\beta}}$ are the residuals derived from the representation in (8). Starting from the initial S-estimates, MM-estimates are calculated easily by iterating these estimating equations until convergence. If w is bounded and non-increasing, the convergence of this iterative procedure to a local minimum is guaranteed (Maronna et al. 2006). The function w can be interpreted as a weight function that can be used to identify outliers. Indeed, a small value of $w(d_i)$ corresponds with a large residual distance d_i and indicates that the observation corresponding to accident period i is an outlier. For more details on the properties of S and MM-estimators, we refer to Peremans and Van Aelst (2018). We now explore

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0556.v1

Peer-reviewed version available at Risks 2018, 6, 108; doi:10.3390/risks6040108

the use of these robust estimators in the GMCL model to obtain robust reserve estimates and identify outliers in the run-off triangles.

5. Simulation Study

First, we introduce a simulation design according to the GMCL model to generate multivariate run-off triangles. Then, we investigate the prediction performance of the classical and robust estimators for GMCL models by simulation.

We consider the case where two run-off triangles are available ($M = 2$), but the results can easily be generalized to more triangles ($M > 2$). To generate two run-off triangles under the GMCL model in (1), we first generate $C_{i,1}^{(m)}$ for $i = 1, \ldots, I$ and $m = 1, 2$ independently from a uniform distribution on the interval $[10^4, 2 \times 10^4]$. These numbers represent the losses observed in the first development period. Then, let

$$\boldsymbol{A}_k = \begin{pmatrix} 10^4 s_k \\ 10^4 s_k \end{pmatrix}, \quad \boldsymbol{B}_k = \begin{pmatrix} 1 & 0.1 s_k \\ 0.1 s_k & 1 \end{pmatrix},$$

for $k = 1, \ldots, I-1$ with $s_k = 0.9^{(k-1)}$. The entries of the first (second) rows determine the increase of the cumulative claims of the first (second) triangle. Note that the structural connections among triangles, i.e. the non-diagonal entries of $\boldsymbol{B}_k$, decrease towards zero for $k \to I-1$ to ensure that the cumulative claims stabilize at a certain point in time. Furthermore, assume that the error terms $\boldsymbol{e}_1^*, \ldots, \boldsymbol{e}_n^*$ from the representation in (8) are independently and normally distributed with mean zero and covariance $\boldsymbol{\Sigma}_k$. The covariance matrices $\boldsymbol{\Sigma}_k$ are defined by multiplying the equicorrelation matrix with correlation 0.5 by the scalar $10^2 s_k$ for $k = 1, \ldots, I-1$. This choice of $\boldsymbol{\Sigma}_k$ leads to error terms that become smaller for $k \to I-1$. If no shrinkage would be applied on the covariance matrices, then the error terms would grow on average because they are linearly related to the cumulative claims of the previous period which increase over time. Finally, the cumulative claims $C_{i,k}^{(m)}$ for $i = 1, \ldots, I$, $k = 2, \ldots, I$ and $m = 1, 2$ can be computed according to the GMCL model in (1) by generating independent error terms from the aforementioned error distribution. We have chosen the parameters $\boldsymbol{A}_k$, $\boldsymbol{B}_k$ and $\boldsymbol{\Sigma}_k$ such that the resulting run-off triangles resemble real data. The cumulative and incremental claims of two run-off triangles simulated according to this data generating process are shown in Figure 1. Note that the patterns in these run-off triangles behave similar for every accident period.

Consider the prediction of a single cell $\mathrm{E}(C_{i,k}^{(m)})$ of subportfolio m for $i + k > I + 1$, i.e. the prediction of a future loss. Given historic claims of M subportfolios, the development parameters $\boldsymbol{A}_k$ and $\boldsymbol{B}_k$ of the GMCL model can be estimated for $k = 1, \ldots, I-1$. Following the GMCL model these parameter estimators yield a corresponding prediction estimator $\hat{C}_{i,k}^{(m)}$ for $\mathrm{E}(C_{i,k}^{(m)})$. In order to measure the prediction accuracy of the estimator $\hat{C}_{i,k}^{(m)}$, we consider its *mean squared error of prediction* (MSEP), given by

$$\mathrm{MSEP}(\hat{C}_{i,k}^{(m)}) = \mathrm{E}(\hat{C}_{i,k}^{(m)} - \mathrm{E}(C_{i,k}^{(m)}))^2.$$

Since in general it is not possible to derive a simple expression for the MSEP, we adopt a Monte-Carlo simulation strategy to estimate this quantity. By repeatedly generating M run-off triangles as described before, fitting the GMCL model and predicting $\mathrm{E}(C_{i,k}^{(m)})$ through the computation of the estimator $\hat{C}_{i,k}^{(m)}$, we obtain J prediction estimators denoted by $(\hat{C}_{i,k}^{(m)})_1, \ldots, (\hat{C}_{i,k}^{(m)})_J$. Then, an estimator of the MSEP of $\hat{C}_{i,k}^{(m)}$ is given by

$$\widehat{\mathrm{MSEP}}(\hat{C}_{i,k}^{(m)}) = \frac{1}{J} \sum_{j=1}^{J} ((\hat{C}_{i,k}^{(m)})_j - \mathrm{E}(C_{i,k}^{(m)}))^2.$$

Smaller values of MSEP indicate a better prediction performance. In our simulation results we will report the square root of the MSEP denoted by RMSEP.

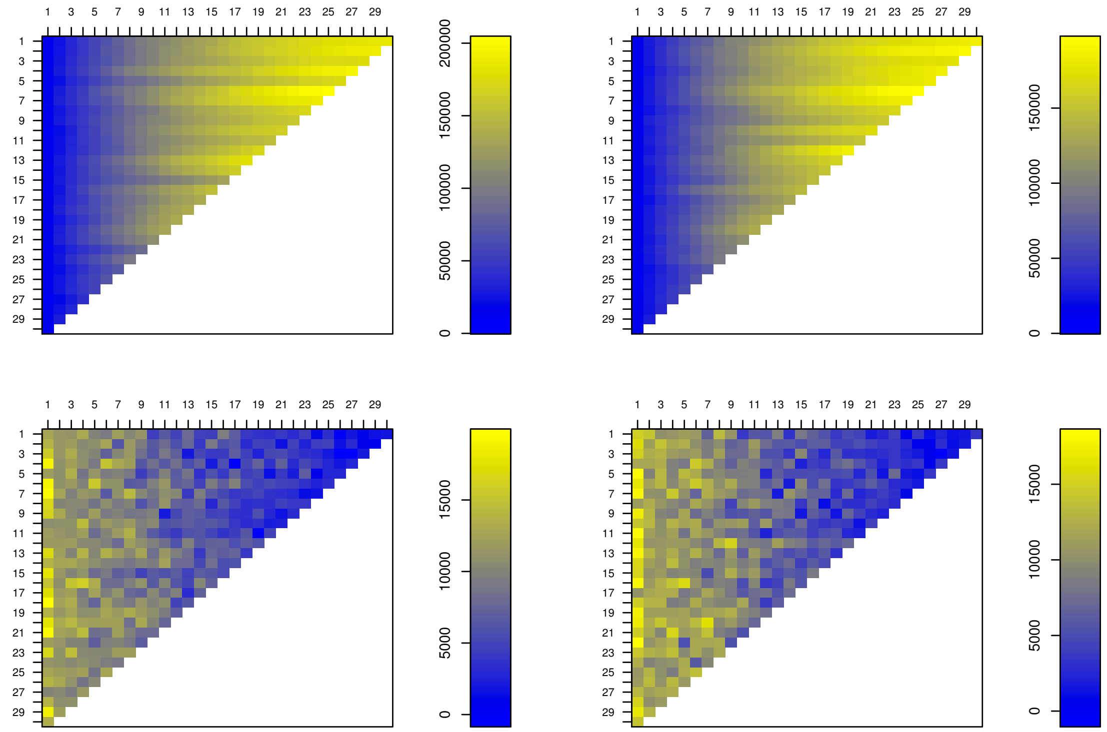

Figure 1. Cumulative and incremental claims for a pair of dependent run-off triangles. The top figures show the cumulative claims of both triangles, whereas the bottom figures show the incremental claims. Development periods are on the horizontal axis, accident periods are on the vertical axis. The bar plot represents a color code indicating the magnitude of the numbers.

For data simulated as described before we consider three procedures: the SCL model in combination with LS (in short SCL-LS) and the GMCL model in combination with FGLS and robust MM-estimators (in short GMCL-FGLS and GMCL-MM respectively). As noted by Zhang (2010, pp. 595-596) it is difficult to fit the SUR models for the upper right part of the triangles because the data is scarce. To avoid numerical instabilities, it is recommended to use SCL for the development in the tail. Naturally, we advice to combine the robust procedure based on MM-estimators with a robust SCL method such as proposed in Verdonck and Debruyne (2011) for the tail development. Since the focus of this paper is on the multivariate model, we present all results without the tail development part, i.e. the final 10 development periods using traditional or robust SCL.

Consider the prediction of the expected claim size $\mathrm{E}(C_{I,2}^{(m)})$ for $m = 1, 2$. The top right panel of Figure 2 shows the estimated RMSEP of $\hat{C}_{I,2}^{(1)}$ for SCL-LS, GMCL-FGLS and GMCL-MM as a function of the total number of accident periods I ranging from 25 to 50 for $J = 1000$ simulations. We can see that the RMSEP estimates are larger for SCL-LS. This is expected because SCL does not take structural connections among run-off triangles into account and contemporaneous correlations between the error terms of the run-off triangles are ignored. Note that GMCL-FGLS and GMCL-MM perform similarly in this setting where the triangles contain only regular measurements. Moreover, similar performance was obtained for $\hat{C}_{I,2}^{(2)}$ and hence, these results are omitted.

We now change the parameters $\boldsymbol{A}_k$, $\boldsymbol{B}_k$ and $\boldsymbol{\Sigma}_k$ in the simulation design in such a way that it matches the SCL structure. For $k = 1, \ldots, I-1$ take

$$\boldsymbol{A}_k = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{B}_k = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

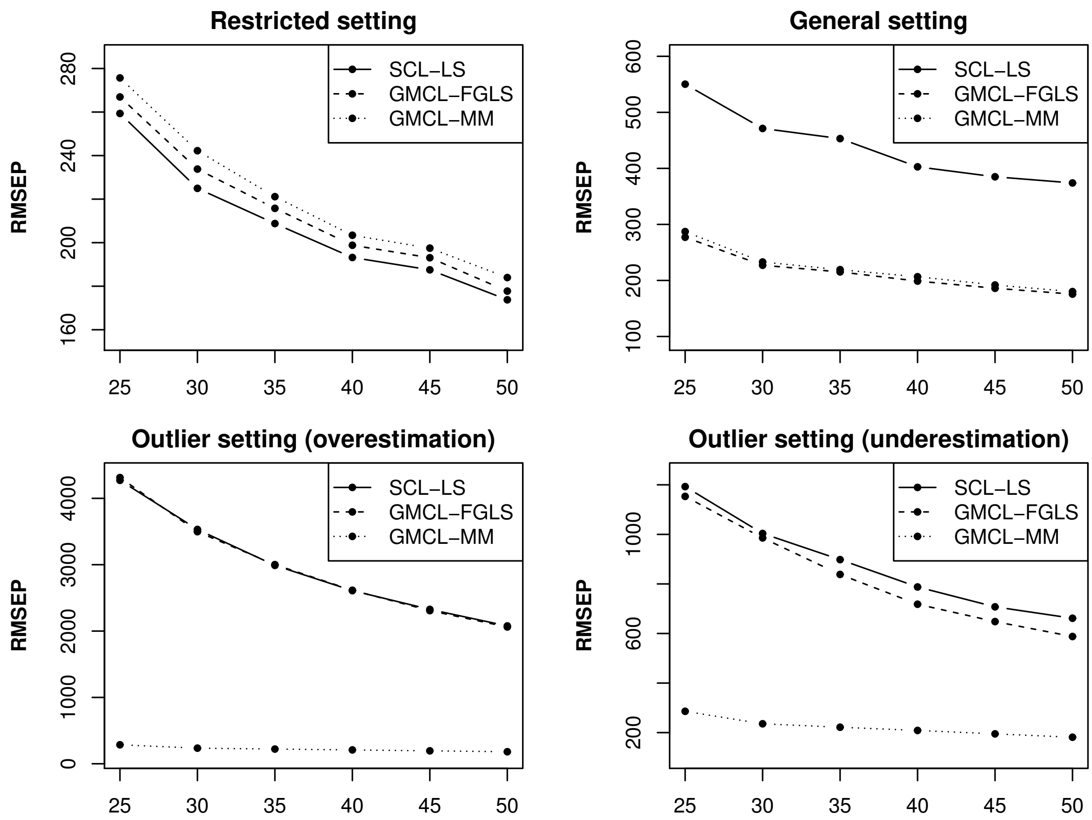

Figure 2. RMSEP estimates of $\hat{C}^{(1)}_{I,2}$ obtained from SCL-LS, GMCL-FGLS and GMCL-MM as a function of I for the restricted, general and outlier settings.

and let $\boldsymbol{\Sigma}_k$ be the identity matrix multiplied with the scalar $10^2 s_k$. In this setting SCL is optimal, whereas the GMCL model uses too many parameters. Intercepts, slopes measuring the effects of the other triangles and correlation parameters are unnecessary in this case. When we compare the results of both estimation procedures, presented in the top left window of Figure 2, we observe that the RMSEP is only slightly larger for GMCL models.

To illustrate the sensitivity of the classical procedures and the robustness of MM-estimators, we now consider the following outlier setting: for each pair of run-off triangles we replace the simulated error term $\boldsymbol{e}_2$ to generate $\boldsymbol{C}_{2,2}$ with $(10^5, 10^5)'$. Based on $J = 1000$ generated pairs of triangles of this kind, we obtained the results in the bottom left panel of Figure 2. Clearly, both classical estimates break down because they largely overestimate $\mathrm{E}(C^{(m)}_{I,2})$, while the robust estimates are not influenced by the outliers. The robust results are similar to the classical results that were obtained when no outliers were present in the data. We also show the effect of small losses in run-off triangles. Therefore, we consider a second outlier setting: for each pair of run-off triangles we replace $\boldsymbol{C}_{2,2}$ with $(0,0)'$. The bottom right plot of Figure 2 shows the RMSEP estimates for this outlier setting. Now, both classical estimators underestimate $\mathrm{E}(C^{(m)}_{I,2})$ due to a small loss observed in accident period two, leading to large RMSEP values. On the other hand, the robust method resists the effect of the outlier and still performs well. In both outlier settings the robust method can also detect the outlier because the weight of the corresponding accident period is zero as can be seen in Figure 3 for the first outlier setting. For the second outlier setting the plot of weights is nearly identical.

To illustrate the impact of the outlier's distance to the regular data, we also consider a third outlier setting: for each pair of run-off triangles we replace the simulated error term $\boldsymbol{e}_2$ to generate $\boldsymbol{C}_{2,2}$ with

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0556.v1
Peer-reviewed version available at Risks 2018, 6, 108; doi:10.3390/risks6040108

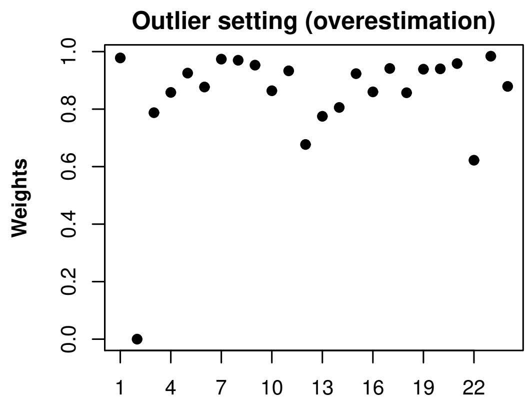

Figure 3. Weights obtained from GMCL-MM for a pair of dependent run-off triangles with one outlier.

$10^4(d,d)'$ where d ranges from -1 to 1. Non-contaminated error terms take values between -3000 and 3000 for the first development period. Therefore, the situations when $|d| > 0.3$ are cases with outliers. Again $J = 1000$ bivariate run-off triangles are generated and the prediction accuracy of the expected claim $E(C_{I,2}^{(m)})$ is measured by MSEP. As opposed to the previous simulations we now fix the number of accident periods I to 25. Figure 4 contains the RMSEP results for different outlier distances d. When

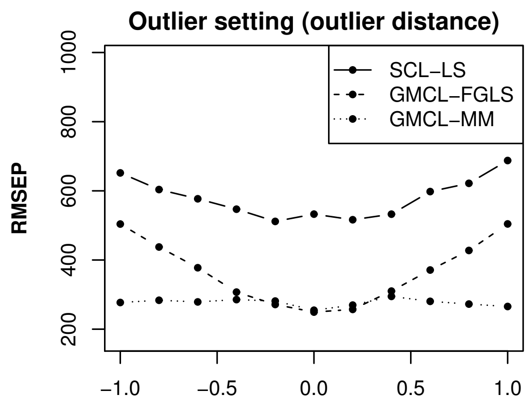

Figure 4. RMSEP estimates of $\hat{C}_{I,2}^{(1)}$ obtained from SCL-LS, GMCL-FGLS and GMCL-MM as a function of the outlier distance d.

$|d| \leq 0.3$ no outliers are generated and the prediction performance of the procedures GMCL-FGLS and GMCL-MM are identical, as we have seen before. For situations with outliers the classical methods yield large RMSEP values because their predictions under- or overestimate the target claim due to the presence of the outliers. The larger the outlier distance d, the worse the prediction accuracy is for non-robust methods. On the other hand, the prediction estimates obtained from the robust method remain stable for all situations.

A more general case is to consider the prediction of $E(C_{I,k}^{(m)})$ for $m = 1,2$ with $k > 2$. In particular, we consider $k = 15$. We repeat the same procedure of squaring $J = 500$ pairs of dependent triangles and measure the prediction accuracy of $\hat{C}_{I,15}^{(m)}$ by means of RMSEP. The results for the general setting are shown in Figure 5. The performance of the different methods is comparable to their performance in the previous setting when predicting $E(C_{I,2}^{(m)})$. However, since $k = 15$ the prediction of $E(C_{I,15}^{(m)})$ depends on 14 model fits, and consequently, the MSEP estimates of $\hat{C}_{I,15}^{(m)}$ become much larger. The prediction performance in the restricted setting and outlier settings (not shown) are also similar as before.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0556.v1
Peer-reviewed version available at Risks 2018, 6, 108; doi:10.3390/risks6040108

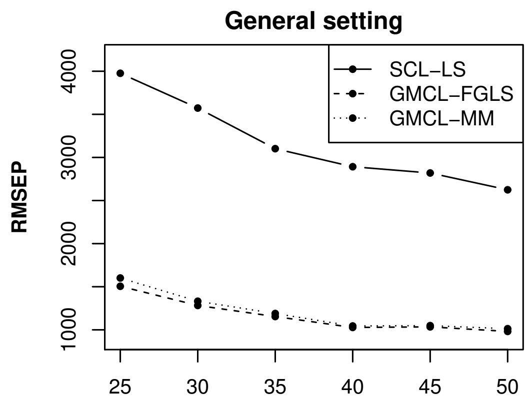

Figure 5. RMSEP estimates of $\hat{C}^{(1)}_{I,15}$ obtained from SCL-LS, GMCL-FGLS and GMCL-MM as a function of I for the general setting.

We have also investigated how the position of the outlier influences the prediction performance. Here the outlier's position refers to the development period in which it has occurred because the effect is similar for all accident years. If the outlier occurs after the target claim, then both the classical and robust methods yield reliable prediction results for the target claim. However, when the outlier occurs before the target claim, then the classical methods yield prediction estimates that are affected by the outlier, while the robust method remains reliable. Only when the outlier appears in the upper right tail of a run-off triangle, it will affect any method, whether it is robust or not, because there is not enough data available in this tail to be able to identify an outlier. Since the position of outliers is unknown in practice, this illustrates the importance of robust procedures which offer protection against outliers in almost any position of the run-off triangles.

6. Real Data

To illustrate the new methodology, we consider an example with paid and incurred data from a motor third party liability (MTPL) and a general third party liability (GTPL) insurance portfolio from a non-life insurance company operating in Belgium. The data have been recorded between March 2008 and December 2015. Quarterly data are available leading to run-off triangles of dimension 31×31 shown in Figure 6. Observe that from accident trimester 15 onwards the cumulative claim amounts for MTPL become much smaller. This effect is due to a decrease in total premium volume, and hence, also in total number of claims. For the GTPL data, accident trimester 1 seems suspicious. The claim amounts are much larger in comparison to any other period. Finally, notice that for the first 15 accident trimesters the losses in the subportfolios are almost fully developed, i.e. the changes in consecutive cumulative claims are minuscule in the last development years.

We model these run-off triangles separately with SCL and jointly with GMCL. The joint model is given by equation (1) with $M = 3$. The separate model simplifies the joint model by excluding intercepts, structural connections and contemporaneous correlations. We have applied SCL-LS, GMCL-FGLS and GMCL-MM to square the run-off triangles up until period 21. As explained before, we exclude the tail development part in order to focus on the multivariate models.

Table A1 in the Appendix contains the estimates of the development parameters and the sample correlations between the resulting residuals obtained by SCL-LS for all development periods. While the run-off triangles have been modeled separately, for some development periods there are substantial correlations between the residuals which indicates that the independence assumption might be violated for these data.

The parameter estimates obtained from GMCL-FGLS are summarized in Table A2 in the Appendix. The slope estimates $\hat{\beta}_{21}, \hat{\beta}_{31}, \hat{\beta}_{12}, \hat{\beta}_{32}, \hat{\beta}_{13}$ and $\hat{\beta}_{23}$ measure the contribution of the other two triangles

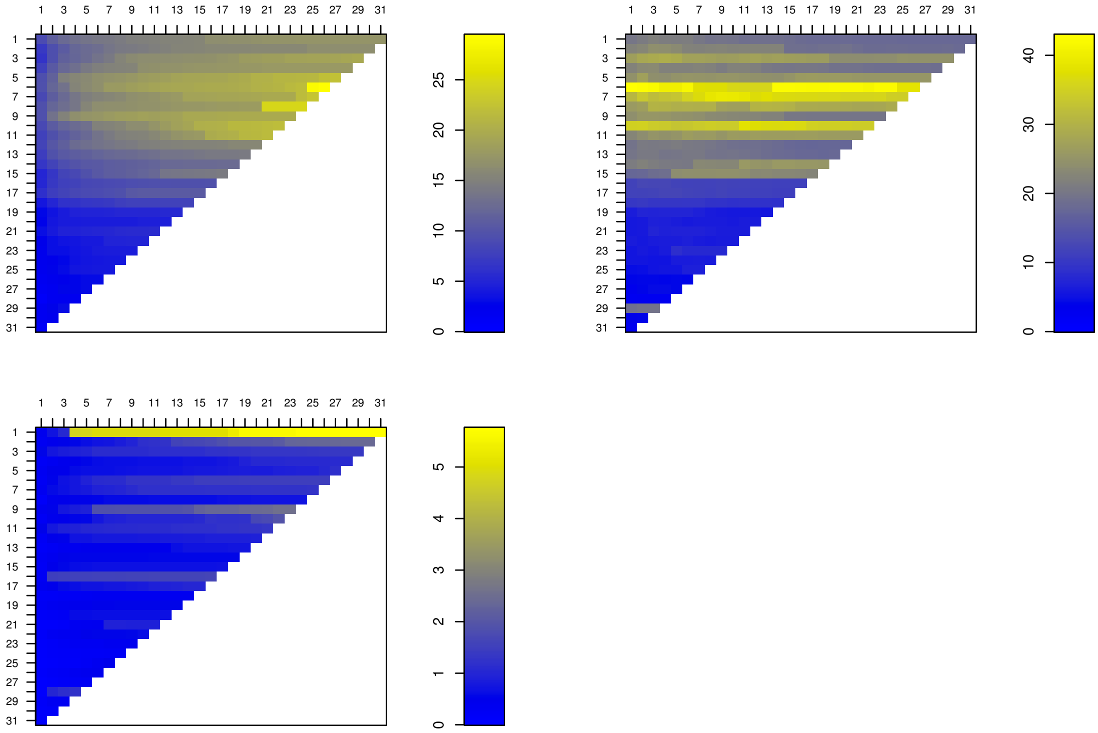

Figure 6. Cumulative run-off triangles (divided by 100000) of a real insurance portfolio. Top left: paid data of MTPL. Top right: incurred data of MTPL. Bottom left: paid data of GTPL. Development periods are on the horizontal axis, accident periods are on the vertical axis. The bar plot represents a color code indicating the magnitude of the numbers.

when predicting future losses in a triangle. From Table A2 it can be seen that for some development periods these estimates are substantially different from zero. They improve the model fit and the prediction performance. The last three columns of Table A2 contain the sample correlations between the residuals of the three run-off triangles, which have been obtained as

$$\hat{\rho}_{mm'} = \frac{\hat{\sigma}_{mm'}}{\sqrt{\hat{\sigma}_{mm}\hat{\sigma}_{m'm'}}},$$

for $m, m' = 1, 2, 3$, where $\hat{\sigma}_{mm'}$ are the entries of the covariance matrix $\hat{\boldsymbol{\Sigma}}_k$. Several moderate to large correlations have been obtained which again supports the joint GMCL model for these data.

We now apply the robust method GMCL-MM which yields the development parameter estimates shown in Table A3 in the Appendix. Based on this robust procedure we can now detect possible outliers. The weights assigned to each observation in the SUR models are shown in Figure 7. The smaller the weight, the more outlying is an observation with respect to the bulk of the data. For example, from Figure 7 we can observe that in the first development period there are two major outliers corresponding to accident trimesters 16 and 28 respectively.

The outliers identified by the GMCL-MM method may have affected the classical estimators, and hence, also the prediction of future losses. Hence, in Table 2 we compare the total reserve estimates for all methods. Let us first focus on the paid losses of the MTPL portfolio. The non-robust SCL-LS and GMCL-FGLS methods both yield a total reserve estimate that is larger than for the robust GMCL-MM. A close inspection of the predicted run-off triangles revealed that the transition from development trimester 20 to 21 is highly responsible for these large differences. For development trimester 21 one can observe in Figure 6 a large incremental increase of the losses that occurred in accident trimester

 doi:10.20944/preprints201807.0556.v1

Peer-reviewed version available at Risks 2018, 6, 108; doi:10.3390/risks6040108

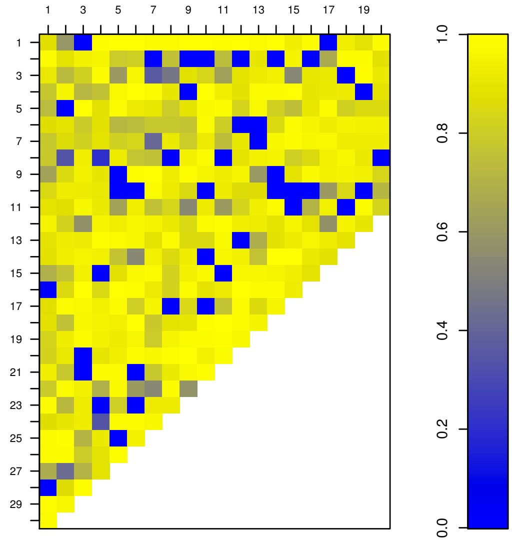

Figure 7. Weights obtained from GMCL-MM for a real insurance portfolio. Each row corresponds to an accident trimester used in the fitting procedure. Each columns represents a SUR model.

Table 2. Total reserve estimates for all run-off triangles of a real insurance portfolio obtained from SCL-LS, GMCL-FGLS and GMCL-MM.

Method	**Run-off Triangle**		
	MTPL paid	MTPL incurred	GTPL paid
SCL-LS	1924001	-654695	386949
GMCL-FGLS	12198112	-1175336	-670116
GMCL-MM	167221	1043591	-128463

8. The SCL-LS and GMCL-FGLS fits for this transition period are both largely influenced by this particular observation. Consequently, the predicted future losses from this development trimester onward are much larger. On the other hand, the robust GMCL-MM method is much less influenced by this observation and is able to flag this observation as an outlier.

Let us now consider the reserve estimates of the incurred losses. The two non-robust approaches agree quite well. The difference is mainly caused by accident trimester 29 for which unexpectedly small paid losses have been observed but at the same time large incurred losses were recorded. In the joint GMCL model the development factor β_{12} for model 7 differs from zero and thus influences the incurred losses obtained by GMCL-FGLS which is not the case for SCL-LS. Moreover, remark that these reserve estimates are negative. Negative reserve estimates are often observed for incurred run-off triangles due to overestimation of the losses. The robust total reserve estimate obtained by GMCL-MM is much larger than for the non-robust methods. This indicates that the presence of outliers has again affected the classical results. More specifically, in this case the classical procedures yield smaller prediction estimates as compared to the robust procedure. For example, one can verify that for the transition from development trimester 18 to 19 the prediction estimates obtained by GMCL-MM are much larger than those obtained by GMCL-FGLS.

Finally, we also consider the estimated reserve for the GTPL portfolio. The unusual data in the first accident trimester affect the total reserve estimates of both non-robust methods. On the other hand, the robust GMCL-MM detected the deviating pattern in the first accident trimester as well as other moderate outliers and yields a robust total reserve estimate that is not driven by atypical behavior in the available data. Note that the GMCL based methods yield negative reserve estimates for these data. While negative reserve estimates are not uncommon for incurred losses, they are rather unusual for run-off triangles with paid losses. However, the real data have been obtained from a small company and the company informed us that for some claims there has been substantial recovery of initially paid losses. These recoveries have an impact on the cumulative claims data which may explain the negative reserve estimates in this case.

To further investigate the performance of the estimation methods, we now focus on the prediction of the values on the last diagonal of all run-off triangles. To measure the accuracy of the predictions, we consider their MSEP. More specifically, we leave out the last diagonal of all three run-off triangles, apply the different methods on the remaining data and calculate the mean squared relative prediction error for each method. The results are given in Table 3 for each subportfolio separately as well as all portfolios jointly. While the three methods perform quite similar on the first two run-off triangles, this

Table 3. MSEP for the last diagonal of all run-off triangles (and totals) of a real insurance portfolio obtained from SCL-LS, GMCL-FGLS and GMCL-MM.

Method	**Run-off Triangle**			**Total**
	MTPL paid	MTPL incurred	GTPL paid	
SCL-LS	0.024	0.021	0.142	0.187
GMCL-FGLS	0.032	0.057	0.337	0.426
GMCL-MM	0.024	0.040	0.076	0.140

is not the case for the GTPL paid data as can be seen from Table 3. The MSEP of GMCL-FGLS is large for this run-off triangle. SCL-LS performs better, but not as good as GMCL-MM which is the only method that yields reasonable performance for these data. As a result, GMCL-MM also shows the best overall performance which illustrates that the outliers in these run-off triangles affect the predictions of the non-robust methods.

7. Conclusion

In this paper we have presented a robust estimation method for the general multivariate chain ladder model proposed by Zhang (2010). Hence, our proposed methodology takes into account contemporaneous correlations and structural connections between different run-off triangles and still yields reliable results when the data are contaminated. Moreover, it allows us to automatically identify the most influential and atypical claims in the run-off triangles.

It is important to further inspect the detected outliers and to understand the reasons for their atypical behavior. If the outliers are errors or due to causes that are not likely to happen again in future, then the robust results can be used as reserve estimates. However, if such atypical observations are expected to re-occur in the future, it is necessary to model also their process (which is outside the scope of this paper) and to predict how much extra reserve besides the robust total reserve estimate is needed to cope with such atypical observations in future years. In such a case the final estimate may for instance be equal to the robust total reserve estimate plus a safe margin when outliers lead to an overestimation of the total reserve estimate. Note that it can also happen that outliers lead to an underestimation of the total reserve estimate even if the atypical claims are larger than the expected claims.

The robust GMCL method was applied on simulated run-off triangles illustrating its excellent performance. From a portfolio analysis of real run-off triangles from a small non-life insurance company in Belgium it was clear that the proposed robust methodology is helpful to gain insight in the data and to build up a more realistic reserve, certainly when it is used in addition to the classical multivariate chain ladder method.

Author Contributions: The authors contributed equally to this work.

Funding: This research was funded by International Funds KU Leuven grant numbers C16/15/068 and C24/15/001, the CRoNoS COST Action grant number IC1408 and the Flemish Science Foundation (FWO) grant number 1523915N.

Conflicts of Interest: "The authors declare no conflict of interest."

Appendix

Table A1. Development parameter estimates and empirical correlation estimates obtained from SCL-LS for a real insurance portfolio.

k	$\hat{\beta}_{11}$	$\hat{\beta}_{22}$	$\hat{\beta}_{33}$	$\tilde{\rho}_{12}$	$\tilde{\rho}_{13}$	$\tilde{\rho}_{23}$
1	1.29	1.04	1.88	0.13	0.51	0.04
2	1.14	1.01	1.18	-0.22	-0.08	0.13
3	1.08	0.99	1.35	0.20	-0.08	-0.08
4	1.05	1.01	1.06	0.26	-0.02	-0.09
5	1.04	1.00	1.12	0.11	-0.02	0.18
6	1.03	1.00	1.05	-0.22	-0.01	0.08
7	1.03	1.00	1.01	-0.14	-0.11	0.53
8	1.02	0.99	1.03	0.38	0.14	0.26
9	1.02	0.99	1.02	0.39	0.14	0.01
10	1.01	1.01	1.01	0.36	-0.11	0.17
11	1.02	1.00	1.01	-0.35	-0.01	-0.03
12	1.01	0.99	1.03	0.26	0.16	0.08
13	1.01	1.01	1.02	-0.29	-0.13	-0.28
14	1.01	0.99	1.03	0.17	0.05	-0.28
15	1.02	0.99	1.02	0.11	-0.23	-0.01
16	1.01	0.99	1.01	0.09	0.43	0.49
17	1.01	1.00	1.03	-0.23	-0.17	0.24
18	1.01	0.99	1.03	-0.54	-0.18	-0.08
19	1.01	0.99	1.03	0.08	-0.28	0.32
20	1.04	0.99	1.01	-0.37	-0.07	-0.04

Ajne, B.. 1994. Additivity of chain-ladder projections. *Astin Bulletin 24*(2), 311–318.

Bilodeau, M. and P. Duchesne. 2000. Robust estimation of the SUR model. *The Canadian Journal of Statistics 28*(2), 277–288. doi:10.2307/3315978.

Braun, C.. 2004. The prediction error of the chain ladder method applied to correlated run-off triangles. *Astin Bulletin 34*(2), 399–423.

Brazauskas, V.. 2009. Robust and efficient fitting of loss models: diagnostic tools and insights. *North American Actuarial Journal 13*(3), 356–369.

Brazauskas, V., B. L. Jones, and R. Zitikis. 2009. Robust fitting of claim severity distributions and the method of trimmed moments. *Journal of Statistical Planning and Inference 139*(6), 2028–2043.

Dreksler, S., C. Allen, A. Akoh-Arrey, J. A. Courchene, B. Junaid, J. Kirk, W. Lowe, S. O'Dea, J. Piper, M. Shah, G. Shaw, D. Storman, S. Thaper, L. Thomas, M. Wheatley, and M. Wilson. 2015. Solvency II Technical Provisions For General Insurers. *British Actuarial Journal 20*(1), 7–129.

England, P. D. and R. J. Verrall. 2002. Stochastic Claims Reserving in General Insurance. *British Actuarial Journal 8*(3), 443–518.

Hubert, M., T. Verdonck, and O. Yorulmaz. 2017. Fast robust SUR with economical and actuarial applications. *Statistical Analysis and Data Mining 10*(2), 77–88. doi:10.1002/sam.11313.

Koenker, R. and S. Portnoy. 1990. M-estimation of multivariate regressions. *Journal of the American Statistical Association 85*(412), 1060–1068. doi:10.1080/01621459.1990.10474976.

Maronna, R. A., D. R. Martin, and V. J. Yohai. 2006. *Robust Statistics: Theory and Methods*. New York: John Wiley and Sons.

Merz, M. and M. V. Wüthrich. 2007. Prediction error of the chain ladder reserving method applied to correlated run-off triangles. *Annals of Actuarial Science 2*(1), 25–50.

Merz, M. and M. V. Wüthrich. 2008. Prediction error of the multivariate chain ladder reserving method. *North American Actuarial Journal 12*(2), 175–197.

Nelder, J. A. and R. W. M. Wedderburn. 1972. Generalized Linear Models. *Journal of the Royal Statistical Society. Series A (General) 135*(3), 370–384.

Peremans, K, P. Segaert, S. Van Aelst, and T. Verdonck. 2017. Robust bootstrap procedures for the chain-ladder method. *Scandinavian Actuarial Journal 2017*(10), 870–897. doi:10.1080/03461238.2016.1263236.

Peremans, K. and S. Van Aelst. 2018. Robust Inference for Seemingly Unrelated Regression Models. *Journal of Multivariate Analysis to appear, arXiv:1801.04716*.

Pitselis, G., V. Grigoriadou, and I. Badounas. 2015. Robust loss reserving in a log-linear model. *Insurance: Mathematics and Economics 64*, 14–27.

Pröhl, C. and K. D. Schmidt. 2005. Multivariate chain-ladder. *Dresdner Schriften zur Versicherungsmathematik*.

Quarg, G. and T. Mack. 2004. Munich chain ladder. *Blätter der DGVFM 26*(4), 597–630.

Salibian-Barrera, M. and V. J. Yohai. 2006. A fast algorithm for S-regression estimates. *Journal of Computational and Graphical Statistics 15*(2), 414–427. doi:10.1198/106186006X113629.

Schmidt, K. D.. 2006. Optimal and additive loss reserving for dependent lines of business. *Casualty Actuarial Society Forum*, 319–351.

Verdonck, T. and M. Debruyne. 2011. The influence of individual claims on the chain-ladder estimates: Analysis and diagnostic tool. *Insurance: Mathematics and Economics 48*(1), 85–98.

Verdonck, T. and M. Van Wouwe. 2011. Detection and correction of outliers in the bivariate chain–ladder method. *Insurance: Mathematics and Economics 49*(2), 188–193.

Verdonck, T., M. Van Wouwe, and J. Dhaene. 2009. A robustification of the chain-ladder method. *North American Actuarial Journal 13*(2), 280–298.

Wüthrich, M. V. and M. Merz. 2008. *Stochastic claims reserving methods in insurance*. John Wiley & Sons.

Zellner, A.. 1962. An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association 57*(298), 348–368. doi:10.2307/2281644.

Zhang, Y.. 2010. A general multivariate chain ladder model. *Insurance: Mathematics and Economics 46*(3), 588–599. doi:10.1016/j.insmatheco.2010.03.002.

doi:10.20944/preprints201807.0556.v1

Peer-reviewed version available at *Risks* **2018**, *6*, 108; doi:10.3390/risks6040108

k	$\hat{\beta}_{01}$	$\hat{\beta}_{11}$	$\hat{\beta}_{21}$	$\hat{\beta}_{31}$	$\hat{\beta}_{02}$	$\hat{\beta}_{12}$	$\hat{\beta}_{22}$	$\hat{\beta}_{32}$	$\hat{\beta}_{03}$	$\hat{\beta}_{13}$	$\hat{\beta}_{23}$	$\hat{\beta}_{33}$	$\hat{\rho}_{12}$	$\hat{\rho}_{13}$	$\hat{\rho}_{23}$
1	23397.72	1.14	0.02	0.93	-11.47	0.08	1.00	0.83	21694.08	-0.02	0.00	1.22	0.20	0.50	0.03
2	15223.35	1.09	0.01	-0.15	20020.27	0.12	0.95	-0.23	1727.03	0.01	0.00	1.07	-0.22	-0.10	0.04
3	16228.14	0.99	0.04	-0.14	15116.47	-0.01	0.99	-0.03	-12277.95	0.05	-0.02	1.57	0.23	0.02	-0.07
4	10350.14	1.00	0.03	-0.06	50876.00	-0.07	1.03	-0.11	4182.92	0.00	0.00	1.00	0.23	-0.19	-0.23
5	1028.93	0.94	0.05	-0.01	-6957.99	-0.05	1.04	0.01	-1377.61	0.02	-0.01	1.01	-0.01	0.04	0.19
6	12243.16	0.97	0.03	-0.03	8286.35	0.06	0.98	-0.36	10968.80	0.00	0.00	0.97	-0.29	-0.01	-0.01
7	-3719.21	1.02	0.00	0.04	-6260.32	-0.13	1.07	0.00	-379.22	0.00	0.00	1.00	-0.22	-0.05	0.62
8	-755.07	1.03	0.00	-0.01	5287.58	-0.04	1.00	0.17	-1120.14	0.00	0.00	1.01	0.41	0.19	0.45
9	-11302.41	1.05	-0.01	-0.07	-4825.36	0.00	1.00	-0.08	904.91	0.00	0.00	1.00	0.36	0.08	-0.05
10	6920.22	0.97	0.03	0.01	37848.84	-0.15	1.09	-0.06	502.78	0.00	0.00	1.00	0.17	-0.02	0.24
11	9660.89	0.95	0.04	0.00	-27830.17	0.09	0.96	-0.05	-438.20	0.00	0.00	1.02	-0.26	0.20	-0.04
12	-16214.89	1.01	0.01	0.00	8784.70	-0.07	1.03	0.10	-1370.21	0.01	0.00	1.00	0.20	0.14	0.16
13	-18821.47	1.00	0.02	0.01	-30184.25	-0.08	1.07	0.08	-1385.69	0.01	0.00	1.00	-0.44	0.02	-0.16
14	-17224.86	1.00	0.02	0.00	40874.99	-0.06	1.02	-0.19	-11617.13	0.01	0.00	1.02	0.08	0.08	-0.32
15	-20373.50	1.02	0.00	0.12	-24051.79	0.06	0.97	-0.11	-7141.82	0.00	0.00	1.01	0.20	-0.21	-0.12
16	-2082.74	1.02	0.00	-0.02	17582.20	0.02	0.98	-0.05	1397.56	0.00	0.00	1.00	0.05	0.36	0.43
17	-44523.11	1.04	0.00	-0.02	61268.64	-0.04	1.00	0.03	2554.84	0.00	0.00	1.05	-0.07	-0.04	-0.11
18	-13650.45	1.02	0.00	-0.03	-51338.15	0.02	0.99	0.05	-6862.07	0.01	0.00	1.04	-0.56	-0.13	-0.10
19	-37910.90	1.01	0.01	0.06	4693.44	0.00	0.99	0.02	-55064.83	0.04	0.00	0.97	0.06	-0.45	0.56
20	874470.74	0.53	0.07	-0.48	-76063.75	0.04	0.99	0.01	-1304.77	0.00	0.00	1.00	-0.31	-0.18	-0.03

Table A2. Development parameter estimates and correlation estimates obtained from GMCL-FGLS for a real insurance portfolio.

k	$\hat{\beta}_{01}$	$\hat{\beta}_{11}$	$\hat{\beta}_{21}$	$\hat{\beta}_{31}$	$\hat{\beta}_{02}$	$\hat{\beta}_{12}$	$\hat{\beta}_{22}$	$\hat{\beta}_{32}$	$\hat{\beta}_{03}$	$\hat{\beta}_{13}$	$\hat{\beta}_{23}$	$\hat{\beta}_{33}$	$\hat{\rho}_{12}$	$\hat{\rho}_{13}$	$\hat{\rho}_{23}$
1	7820.38	1.15	0.02	1.16	-3680.41	0.08	1.00	1.03	1717.07	0.01	0.00	1.11	0.24	-0.31	0.06
2	12144.56	1.09	0.01	-0.11	16619.03	0.13	0.95	-0.19	873.94	0.01	0.00	1.06	0.20	0.11	-0.10
3	23528.36	1.00	0.03	-0.20	22422.99	0.00	0.98	-0.10	1918.65	0.01	0.00	0.99	0.08	0.30	0.10
4	8438.94	1.01	0.02	-0.04	891.69	0.06	0.97	-0.11	4896.14	0.00	0.00	1.00	0.03	0.00	0.21
5	-2355.67	0.98	0.03	-0.02	-30886.96	-0.06	1.05	0.04	1715.40	-0.01	0.00	1.03	-0.03	-0.21	-0.04
6	8351.98	0.97	0.04	-0.04	9538.34	0.07	0.97	-0.33	-209.96	0.00	0.00	1.00	-0.29	-0.20	-0.08
7	-2873.28	1.02	0.00	0.03	-4771.62	-0.12	1.07	0.02	-243.36	0.00	0.00	1.00	-0.23	-0.17	0.64
8	-806.41	1.00	0.01	0.01	821.12	-0.06	1.02	0.13	-1135.19	0.00	0.00	1.01	0.06	0.09	0.32
9	-6931.74	1.03	0.00	-0.03	1925.54	-0.03	1.01	-0.02	1272.45	0.00	0.00	1.00	-0.19	0.21	0.02
10	8446.18	0.97	0.02	0.00	13573.18	0.00	0.99	-0.06	44.18	0.00	0.00	1.00	-0.46	-0.05	-0.17
11	-1481.68	0.98	0.03	0.00	-3558.47	0.04	0.97	0.04	-588.16	0.00	0.00	1.02	-0.02	0.15	-0.03
12	-19036.01	1.01	0.01	0.00	10657.18	-0.07	1.03	0.08	-1020.05	0.00	0.00	1.00	0.13	0.77	0.10
13	-17979.52	1.03	0.00	-0.02	21175.00	-0.07	1.03	0.08	-1469.87	0.01	0.00	1.00	0.23	-0.25	-0.18
14	-6110.32	1.01	0.00	0.00	-21779.08	0.02	1.00	-0.08	-4066.28	0.00	0.00	1.00	-0.34	0.66	0.16
15	-2628.61	0.99	0.00	0.15	-20629.80	0.05	0.97	-0.10	219.13	0.00	0.00	1.01	-0.07	-0.50	-0.11
16	621.54	1.02	0.00	-0.03	-42626.20	0.03	1.00	-0.05	-2510.24	0.00	0.00	0.99	-0.59	0.79	-0.22
17	-39374.59	1.04	0.00	-0.10	70972.58	-0.07	1.00	0.25	2017.96	0.00	0.00	1.00	0.15	-0.12	0.60
18	25424.10	0.98	0.01	-0.02	101648.10	-0.11	1.03	0.09	-25270.86	0.02	-0.01	1.02	0.12	0.09	-0.97
19	-42462.66	1.02	0.02	-0.11	74563.74	-0.04	1.01	-0.13	4055.82	0.00	0.00	1.02	0.83	-0.89	-0.99
20	-23405.29	1.01	0.00	0.03	-61530.32	0.04	0.99	0.00	2593.46	0.00	0.00	1.00	0.21	0.52	-0.08

Table A3. Development parameter estimates and correlation estimates obtained from GMCL-MM for a real insurance portfolio.