

Jeong, Himchan; Dey, Dipak

Article

Application of a vine copula for multi-line insurance reserving

Risks

Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Jeong, Himchan; Dey, Dipak (2020) : Application of a vine copula for multi-line insurance reserving, Risks, ISSN 2227-9091, MDPI, Basel, Vol. 8, Iss. 4, pp. 1-23, <https://doi.org/10.3390/risks8040111>

This Version is available at:

<https://hdl.handle.net/10419/258064>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Article

Application of a Vine Copula for Multi-Line Insurance Reserving

Himchan Jeong ^{1,*}  and Dipak Dey ² 

¹ Department of Statistics and Actuarial Science, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada

² Department of Statistics, University of Connecticut, 215 Glenbrook Rd. U-4120, Storrs, CT 06269-4120, USA; dipak.dey@uconn.edu

* Correspondence: himchan_jeong@sfu.ca

Received: 11 September 2020; Accepted: 14 October 2020; Published: 21 October 2020



Abstract: This article introduces a novel use of the vine copula which captures dependence among multi-line claim triangles, especially when an insurance portfolio consists of more than two lines of business. First, we suggest a way to choose an optimal joint loss development model for multiple lines of business that considers marginal distribution, vine copula structure, and choice of family for each pair of copulas. The performance of the model is also demonstrated with Bayesian model diagnostics and out-of-sample validation measures. Finally, we provide an implication of the dependence modeling, which allows a company to analyze and establish the risk capital for whole portfolio.

Keywords: Bayesian inference; model selection; multi-line reserving; property and casualty insurance; vine copula

1. Introduction

There have been various approaches to determine loss reserve for a single line of business with statistical models. According to Mack (1993), the chain-ladder method, which has been used as a rule of thumb for the determination of reserve, can be interpreted as a nonparameteric stochastic model. After that, many univariate stochastic reserving models have been developed and used for both determination of point estimates of reserve and risk management. For detailed discussions on univariate stochastic reserving models, see England and Verrall (2002).

However, it is also certain that most of the insurance companies do not run only a single line of business so one needs to consider possible dependence among the lines of business in reserve modeling. Consequently, stochastic reserving models need to be extended to multivariate frameworks. In this regard, some actuarial literature focused on the extension of the chain-ladder method to multivariate cases, such as Braun (2004), Schmidt (2006), and Merz and Wüthrich (2008) which are followed by Shi et al. (2012), which used a bivariate normal distribution to model multi-line reserves. Besides the dependence among multiple lines of business, there have been some works on the dependence between paid claim triangles and incurred claim triangles such as Zhang (2010) and Merz and Wüthrich (2010).

Note that the methods mentioned above are less flexible since it is not allowed to disentangle marginal distributions and multivariate association structure. Therefore, use of a copula was introduced to the multivariate reserving problem since this allows us to consider marginal distribution and association structure in a separate way. Based on this idea, Shi and Frees (2011) proposed the use of a bivariate Gaussian copula to model dependence among multi-line reserves where the marginal distributions are chosen as Gaussian and gamma, respectively. Peters et al. (2014) considered dependence between paid-incurred claim triangles via bivariate copula. Further, Abdallah et al. (2015)

used the idea of the hierarchical archimedean copula to capture calendar year effect and multi-line dependence simultaneously.

Although there have been some works on the multivariate reserving in actuarial literature, applications of all the aforementioned works are restricted only to possible dependence between two lines of business. Indeed, it is natural that a property and casualty insurance company runs more than two lines of business, so we need to incorporate such a high-dimensional dependence structure in our reserve modeling.

In that regard, we apply the idea of a vine copula in this article, which uses bivariate copulas as its building blocks and connects them with vine structure to describe the high-dimensional association in a flexible way. The vine copula has been widely used in the actuarial and financial literature. [Loaiza Maya et al. \(2015\)](#) investigated dependency among the exchange rates of Latin American countries using the vine copula. [Reboredo and Ugolini \(2016\)](#), [Arreola Hernandez et al. \(2017\)](#), and [Trucíos et al. \(2020\)](#) used the vine copula to assess systemic risks due to possible dependence among economic subjects. Surprisingly, usage of the vine copula in property and casualty insurance literature does exist but is scarce. For example, [Shi and Yang \(2018\)](#) used a vine copula to capture the serial dependence of claim amounts and derive the experience ratemaking factor.

This paper has been organized as follows. In Section 2, basic concept of the vine copula is introduced and the model selection procedure to be implemented is specified. In Section 3, we describe the data used for our empirical analysis. In Section 4, we go through the model selection procedure to determine marginal distributions and the vine copula structure to be used in our analysis. In Section 5, we discuss the implications of estimation results from the perspective of enterprise risk management using the predictive distribution of unpaid claims. Section 6 addresses practical issues for implementing the proposed methodology. Finally, we conclude this article in Section 7 by providing some future directions of research.

2. Proposed Methodology

Suppose an insurance company owns a portfolio which consists of N multiple lines of business. By assuming balanced observations in multiple triangles, one can write the multivariate cumulative paid claims as $\mathbf{Y}_{ij} = (Y_{ij}^{(1)}, Y_{ij}^{(2)}, \dots, Y_{ij}^{(N)})$ where $n = 1, \dots, N$ indicates a claim triangle from the n^{th} line of business, $i = 1, \dots, I$ means the i^{th} accident years, and $j = 1, \dots, J$ denotes j^{th} development lag. In general, it is of interest to predict the cumulative paid claim for the next year given information up to the current year, which can be written as $\mathbb{E}[Y_{i,j+1}^{(n)} | Y_{i,j}^{(n)}]$. Therefore, instead of working with \mathbf{Y}_{ij} , one can directly model the incremental development of claims, or age-to-age factors as follows:

$$\mathbf{D}_{ij} = (D_{ij}^{(1)}, D_{ij}^{(2)}, \dots, D_{ij}^{(N)}),$$

where $D_{i,j+1}^{(n)} := Y_{i,j+1}^{(n)} / Y_{i,j}^{(n)}$ so that $\mathbb{E}[Y_{i,j+1}^{(n)} | Y_{i,j}^{(n)}] = Y_{i,j}^{(n)} \mathbb{E}[D_{i,j+1}^{(n)}]$. Since each of $D_{ij}^{(n)} = (D_{ij}^{(n)})_{i=1, \dots, I, j=1, \dots, J}$ are observed from the same business line, it is natural that we model the marginal distribution of $D^{(n)}$ and the dependence structure among $D^{(n)}$ via copulas to jointly model $D^{(n)}$, $n = 1, 2, \dots, N$.

According to [Sklar \(1959\)](#), if all marginal distribution functions are continuous, then there is a unique function $C : [0, 1] \rightarrow [0, 1]$ such that

$$H(x_1, x_2, \dots, x_m) = C(F_1(x_1), F_2(x_2), \dots, F_m(x_m)),$$

where F_i denotes marginal distribution function of X_i and H denotes joint distribution function of (X_1, X_2, \dots, X_m) . Therefore, use of copulas allows us to capture the association among the joint response random variables, which may follow different marginal distributions. In that sense, by letting

$\theta = (\theta^{(1)}, \dots, \theta^{(N)})$, where $\theta^{(n)}$ is the parameter vector for the marginal distribution of $D^{(n)}$ and ϕ is the copula parameter, the likelihood for joint distribution is given as follows:

$$\begin{aligned} \ell(\theta, \phi | \mathbf{d}) &= \sum_{i=1}^I \sum_{j=2}^{J+1-I} \log h \left(d_{ij}^{(1)}, \dots, d_{ij}^{(N)} \mid \theta, \phi \right) \\ &= \sum_{i=1}^I \sum_{j=2}^{J+1-I} \left[\sum_{n=1}^N \log f^{(n)} \left(d_{ij}^{(n)} \mid \theta^{(n)} \right) + \log c_{\phi} \left(F^{(1)}(d_{ij}^{(1)} | \theta^{(1)}), \dots, F^{(N)}(d_{ij}^{(N)} | \theta^{(N)}) \right) \right]. \end{aligned}$$

Here h means the joint density of $(D^{(1)}, \dots, D^{(N)})$ and we assume that $d_{ij}^{(n)}$ ($i = 1, \dots, I, j = 2, \dots, J$) are independent for fixed n .

Depending on the sign of dependence, we may use different families of copulas. For example, in order to capture positive dependence, one can suggest the following bivariate copulas:

$$\text{Gumbel: } C^G(u, v) = \exp \left[- \left((-\log(u))^{\phi} + (-\log(v))^{\phi} \right)^{1/\phi} \right], \quad \phi \in [1, \infty)$$

$$\text{Clayton: } C^C(u, v) = \{ u^{-\phi} + v^{-\phi} - 1 \}^{-1/\phi}, \quad \phi \in (0, \infty).$$

Although those two copulas can only capture positive dependence, one can easily rotate those in order to capture reversed tail behavior and negative dependence as follows:

$$90^\circ \text{ rotated Gumbel: } C^{G90}(u, v) = v - C^G(1 - u, v),$$

$$90^\circ \text{ rotated Clayton: } C^{C90}(u, v) = v - C^C(1 - u, v),$$

$$\text{Survival Gumbel: } C^{SG}(u, v) = u + v - 1 + C^G(1 - u, 1 - v),$$

$$\text{Survival Clayton: } C^{SC}(u, v) = u + v - 1 + C^C(1 - u, 1 - v),$$

$$270^\circ \text{ rotated Gumbel: } C^{G270}(u, v) = u - C^G(u, 1 - v),$$

$$270^\circ \text{ rotated Clayton: } C^{C270}(u, v) = u - C^C(u, 1 - v).$$

Further, the Gaussian copula and Frank copula are also prevalent choices which can capture both positive and negative dependence:

$$\text{Gaussian: } C^N(u, v) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \phi), \quad \phi \in [-1, 1],$$

$$\text{Frank: } C^F(u, v) = -\frac{1}{\phi} \log \left[1 + \frac{(\exp(-\phi u) - 1)(\exp(-\phi v) - 1)}{\exp(-\phi) - 1} \right], \quad \phi \in \mathbb{R} \setminus \{0\},$$

where $\Phi_2(\cdot, \cdot)$ stands for the cumulative distribution function of a bivariate, standard, normal, random variable with correlation ϕ and $\Phi(\cdot)$ stands for the cumulative distribution function of a univariate, standard, normal random variable.

Given families of copulas, allow us to consider various facets of possible dependence among the lines of business, including the upper and lower tail dependence properties. For example, the Clayton copula can capture a positive association and has lower tail dependence but no upper tail dependence. The Gumbel copula can capture positive associations and has upper tail dependence but no lower tail dependence. Further, both Frank and Gaussian copulas are symmetric and able to capture positive and negative dependence, but they have no tail dependencies. Note that upper tail dependence of original copulas corresponds to lower tail dependence of survival copulas, and vice versa. Further, if an original copula can capture a positive association, then the corresponding 90° or 270° rotated copula can capture a negative association; for example, 90° rotated Clayton and 270° rotated Gumbel copulas can be used to capture negative associations. We refer the readers to [Nelsen \(1999\)](#), [Embrechts et al. \(2001\)](#), and [Hua and Joe \(2011\)](#) for a more detailed explanation. The variability of tail behaviors is potentially important considering the data used in our empirical analysis. Here we capture the dependencies among the incremental development factors of different lines of business so that it is natural that the

incremental development factors will be quite large in the initial years and become smaller in the mature years. In this regard, upper tail dependence corresponds to the associations among the loss development factors in the initial years, whereas lower tail dependence corresponds to the associations among the loss development factors in the mature years.

If $N = 2$, then the dependence can be captured by a bivariate copula, which has been explored in previous works such as [Shi and Frees \(2011\)](#). However, we may not preclude the possibility that $N > 2$, and one can suggest the use of vine copula when we have more than two lines of business to capture the multiple dependence. The concept of vine copula was introduced by [Aas et al. \(2009\)](#) who extended the research conducted by [Bedford and Cooke \(2001\)](#) and [Bedford and Cooke \(2002\)](#) to show how multivariate data, which exhibit complex patterns of dependence in the tails, can be modeled using a cascade of pair-copulas, acting on two variables at a time. Later [Czado \(2010\)](#) and [Joe and Kurowicka \(2011\)](#) developed this idea further and provided us with a definite road map for constructing vine copulas to represent complex multivariate data. The core concept behind construction of vine copulas is that a d -dimensional density $f(x_1, x_2, \dots, x_d)$ can be represented as a product of pair copula densities and marginal densities. For example, let us consider $d = 3$ dimension. A possible way to represent the joint density $f(x_1, x_2, x_3)$ is as follows:

$$f(x_1, x_2, x_3) = f_{3|12}(x_3|x_1, x_2)f_{2|1}(x_2|x_1)f_1(x_1).$$

Additionally, since

$$\begin{aligned} f_{2|1}(x_2|x_1) &= c_{12}(F_1(x_1), F_2(x_2))f_2(x_2), \\ f_{3|12}(x_3|x_1, x_2) &= c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)|x_2)f_{3|2}(x_3|x_2), \\ f_{3|2}(x_3|x_2) &= c_{23}(F_2(x_2), F_3(x_3))f_3(x_3), \end{aligned}$$

where $c_{ij}(F_i(x_i), F_j(x_j))$ is the joint copula density of random variables X_i and X_j . One can write out the joint density of X_1, X_2, X_3 as follows:

$$\begin{aligned} f(x_1, x_2, x_3) &= f_1(x_1)f_2(x_2)f_3(x_3) \quad (\text{marginals}) \\ &\times c_{12}(F_1(x_1), F_2(x_2))c_{23}(F_2(x_2), F_3(x_3)) \quad (\text{unconditional pair}) \\ &\times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)|x_2) \quad (\text{conditional pair}). \end{aligned}$$

Note that in general, $c_{13|2}$ is affected by x_2 , not only through $F_{1|2}(x_1|x_2)$ and $F_{3|2}(x_3|x_2)$, but also directly through x_2 . However, such a general form of conditional copula density is cumbersome to be used in statistical inference. For this reason, it is tempting to use a simplified form of conditional copula density so that $c_{13|2}$ is affected by x_2 only through $F_{1|2}(x_1|x_2)$ and $F_{3|2}(x_3|x_2)$. [Haff et al. \(2010\)](#) showed that we may write each conditional pair copula density in a simplified form for elliptical distributions with a positive definite scale matrix. Further, it was also shown that use of the simplified form could be a good approximation even in a case wherein such an assumption could not be fully validated. Hence, we will use a simplified form of conditional copula density in this article hereafter.

The sequential dependence structure via pair copulas is described with a regular vine, or R-vine. Let $\mathcal{V} = (T_1, \dots, T_{d-1})$ be a set of trees where each tree $T_j = (N_j, E_j)$ is connected and $N_1 = \{1, \dots, d\}$. Here N_j and E_j mean the nodes and edges of T_j , respectively. If $N_j = E_{j-1}$ for $j = 2, \dots, d$, \mathcal{V} is called a vine. Further, if

$$|e \cap e'| = 1 \text{ for } \{e, e'\} \in E_j \text{ and } j = 2, \dots, d - 1, \tag{1}$$

then \mathcal{V} is called a R-vine. There are two important subclasses of R-vine, C-vine and D-vine. If there is a node $n \in N_j$ so that the degree of that node is $d - j$ in an R-vine, it is called a C-vine. D-vine is an R-vine wherein the degree of every node is at most 2.

The above decomposition of the joint density in terms of vine structure is not unique. For example, when $d = 3$, we have the following three available R-vine structures as in Figure 1. Note that in the case of 3-dimensional R-vine structures, C-vine and D-vine are identical, which is no more true for $d > 3$.

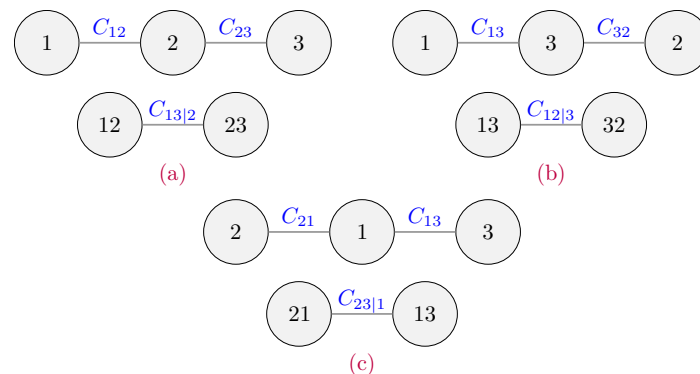


Figure 1. R-vine structures for 3 dimensions.

Note that the number of regular vine structures on d -dimensional variables grows super-exponentially so that there are $d! \cdot 2^{\binom{d-2}{2}}$ regular vine structures on d variables, as shown in Joe and Kurowicka (2011). Therefore, it is already meaningless to display every possible vine structure if $d \geq 5$.

Therefore, for the calibration of vine copula model, we need to carefully determine the following components based on observed data:

- Marginal distribution for each of $D^{(n)}$;
- Optimal d -dimensional vine structure;
- Copula family for each of pair copula in the selected vine structure.

For statistical inference, one may consider either the frequentist or the Bayesian approach and subsequently choose the optimal structure via a model selection procedure. However, it can be quite computationally demanding to choose an optimal model in both approaches. For example, as mentioned in Table 1 of Gruber and Czado (2015), the size of search space of vine copula model for $d = 5$ is already (1.3559×10^{11}) when we consider seven families of copula for each of the pair copulas, which is definitely computationally infeasible. In this regard, Dissmann et al. (2013) proposed a top-down approach to choose the vine structure from a frequentist view. Gruber and Czado (2018) proposed sequential vine copula model selection from a Bayesian perspective. Schamberger et al. (2017) also considered a Bayesian approach to model dependence structure with a factor copula in order to handle the dimensionality issue. Recently, Kreuzer and Czado (2019) showed that the pair copula family can be chosen together with their parameters using the Hamiltonian MCMC.

The Bayesian approach can be beneficial in reserve modeling since it can incorporate the uncertainty of parameter estimation and reflect it to compute the predictive distribution of future unpaid claims, as proposed in Shi et al. (2012). Therefore, we propose the following two-step approach, which can be considered as a compromise of Dissmann et al. (2013) and Gruber and Czado (2018):

1. Estimate the parameters with Bayesian inference and choose marginal distribution based on Bayesian model selection criteria, such as deviance information criterion (DIC) and the logarithm of the pseudomarginal likelihood (LPML). The DIC for each marginal distribution of n^{th} line, proposed by Spiegelhalter et al. (2002), is defined as

$$DIC^{(n)} = -4 \int \ell(\theta^{(n)} | \mathbf{d}^{(n)}) \pi(\theta^{(n)} | \mathbf{d}^{(n)}) d\theta + 2\ell(\tilde{\theta}^{(n)} | \mathbf{d}^{(n)}),$$

where $\tilde{\theta}^{(n)} = \mathbb{E} \left[\theta^{(n)} | \mathbf{d}^{(n)} \right]$ and $\mathbf{d}^{(n)} = \{d_{i,2}^{(n)}, \dots, d_{i,J-i+1}^{(n)} | i = 1, \dots, I\}$ are the observed values of incremental loss developments from n^{th} line of business. Since this integral is hardly available in a closed form, one can estimate DIC as follows:

$$\widehat{\text{DIC}}^{(n)} = -4 \sum_{s=1}^S \ell(\theta_{[s]}^{(n)} | \mathbf{d}^{(n)}) + 2\ell \left(\frac{1}{S} \sum_{s=1}^S \theta_{[s]}^{(n)} \middle| \mathbf{d}^{(n)} \right),$$

where $\{\theta_{[s]}^{(n)}\}_{s=1}^S$ are MCMC samples generated from the posterior density. Note that we prefer models with smaller DIC values.

LPML is calculated based on the conditional predictive ordinate (CPO), which was proposed by Gelfand et al. (1992) and Geisser (2017). The CPO for $d_{i,j}^{(n)}$ is defined as follows:

$$\text{CPO}_{i,j}^{(n)} = \int f(d_{i,j}^{(n)} | \theta) \pi(\theta^{(n)} | \mathbf{d}_{(-i,j)}^{(n)}) d\theta^{(n)},$$

where $\mathbf{d}_{(-i,j)}^{(n)} = \mathbf{d}^{(n)} - d_{i,j}^{(n)}$. Since CPO is usually not readily available in closed form, one can estimate $\widehat{\text{CPO}}_{i,j}^{(n)}$ as follows, according to Gelfand and Dey (1994):

$$\widehat{\text{CPO}}_{i,j}^{(n)} = S \left(\sum_{s=1}^S \frac{1}{f(d_{i,j}^{(n)} | \theta_{[s]})} \right)^{-1}.$$

Finally, according to Ibrahim et al. (2014), $\text{CPO}_{i,j}^{(n)}$ can be summarized as LPML as follows:

$$\text{LPML}^{(n)} = \sum_{i=1}^I \sum_{j=2}^{J+1-I} \log(\widehat{\text{CPO}}_{i,j}^{(n)}).$$

We prefer models with larger LPML values.

- Based on the fitted marginal model and corresponding posterior means of marginal parameters $\hat{\theta}^{(n)}$ for $n = 1, \dots, N$, generate probability integral transform (PIT) $\hat{U}_{ij}^{(n)} = F^{(n)} \left(d_{ij}^{(n)} \middle| \hat{\theta}^{(n)} \right)$ and optimize the following:

$$\ell \left(\phi \middle| \hat{U} \right) = \sum_{i=1}^I \sum_{j=2}^{J+1-I} \left[\log c_{\phi} \left(\hat{U}_{ij}^{(1)}, \dots, \hat{U}_{ij}^{(N)} \right) \right], \tag{2}$$

which ends up with the optimal vine copula structure and copula family for each pair copula using the top-down algorithm of Dissmann et al. (2013). The algorithm starts with the choice of first tree T_1 and subsequently chooses T_j for $j = 2, \dots, d - 1$ while we keep the constraint of having R-vine structure in (1). The optimal tree T_j is chosen by solving the following optimization problem:

$$\underset{E_j, N_j}{\operatorname{argmin}} \sum_{e \in E_j} w_e,$$

where w_e is pairwise BIC of an edge e in our implementation.

In our search for the best family for each of pair copula, we use the aforementioned bivariate Gaussian, Frank, Clayton, Gumbel, and their rotated copulas. Note that this approach also utilizes the inference by margin (IFM) method proposed by Joe and Xu (1996), since it decomposes estimation of marginal distribution and copula structure separately for model selection a with

lesser computational burden. It should be noted that such computational convenience is obtained at the expense of potential estimation bias, as shown in [Louzada and Ferreira \(2016\)](#).

3. Data Description

For the empirical analysis, we used a publicly available dataset from ACE Limited 2013 Global Loss Triangles, which consists of aggregated claim developments for insurance operations in North America (4 lines of business), insurance operations overseas (3 lines of business), and global reinsurance operations (2 lines of business). A description and the corresponding index for each line of business are given in Table 1. We incorporate a 9-dimensional vine copula structure to capture possible dependence among all lines of business. Indeed, the current COVID-19 situation could be a clear example of showing the inappropriateness of ignoring potential dependencies among different countries. Note that the dataset and code for data analysis are attached as Supplementary Materials.

Table 1. Description of the lines of business.

Index	Description
1	North American Workers' Compensation
2	North American General Liability
3	North American Other Casualty
4	North American Non-Casualty
5	Overseas General Casualty
6	Overseas General Non-Casualty
7	Overseas General Personal Accident
8	Global Reinsurance Property
9	Global Reinsurance Non-Property

Although we consider all possible lines of business simultaneously here, an actuary who applies vine copulas to model dependence among the lines of business should be careful, since the complexity of the vine copula structure increases super-exponentially as the dimensionality grows. Therefore, it is a role of an actuary to apply his/her industrial experience or knowledge for the dependence modeling. Further, it is also possible to observe more than 10 lines of business according to the classification in Schedule P of the National Association of Insurance Commissioners (NAIC) annual report, including but not limited to:

- Homeowners/farmowners;
- Private passenger auto liability/medical;
- Commercial auto/truck liability/medical;
- Worker's compensation;
- Special liability;
- Other liability;
- Fidelity/surety.

Therefore, it is also possible to merge some lines of business into one category in order to not only enhance homogeneity and credibility in the analysis but also avoid too much complexity of the vine structure due to the large number of lines of business.

The loss development triangles, which are training sets in terms of predictive analytics, can be expressed as follows:

$$\mathcal{D}_{1:I} = \{Y_{ij}^{(n)} : 1 \leq i \leq I \text{ and } 1 \leq j \leq \min(I, I + 1 - i), n = 1, \dots, 9\}. \quad (3)$$

Our task is to predict the cumulative paid claims for the next years, which are described as follows:

$$\mathcal{D}_{I+k} = \{Y_{ij}^{(n)} : 1 + k \leq i \leq I \text{ and } j = I + 1 + k - i, n = 1, \dots, 9\}. \quad (4)$$

For enterprise risk management (ERM) purposes, insurance companies are usually interested in the distribution of aggregate unpaid claims after one year of loss development, which is determined by incremental paid losses defined as follows:

$$L^{(n)} = \sum_{i=2}^I (Y_{i,I+2-i}^{(n)} - Y_{i,I+1-i}^{(n)}) = \sum_{i=2}^I D_{i,I+2-i}^{(n)} \cdot Y_{i,I+1-i}^{(n)} \quad L = \sum_{n=1}^N L^{(n)}. \tag{5}$$

This formulation validates us to model $D_{i,j}^{(n)}$ rather than $Y_{i,j}^{(n)}$ because one can see that

$$\mathbb{E}[L|\mathcal{D}_{1:I}] = \sum_{n=1}^N \sum_{i=2}^I \mathbb{E}[D_{i,I+2-i}^{(n)}|\mathcal{D}_{1:I}] \cdot Y_{i,I+1-i}^{(n)}.$$

A simple approach for the analysis of L is the silo approach, which means we just aggregate all lines of business so that the paid losses of the same development lag and accident year are merged and modeled altogether. In this case, one can write L as follows:

$$L = \sum_{i=2}^I (Y_{i,I+2-i}^\bullet - Y_{i,I+1-i}^\bullet) = \sum_{i=2}^I D_{i,I+2-i}^\bullet \cdot Y_{i,I+1-i}^\bullet$$

where $Y_{i,j}^\bullet = \sum_{n=1}^N Y_{i,j}^{(n)}$ and $D_{i,j}^\bullet = Y_{i,j+1}^\bullet / Y_{i,j}^\bullet$. However, such an approach ignores the heterogeneity of the lines of business and could be problematic. Suppose there are two lines of business, personal and commercial auto insurances, and reported losses of these lines are given in Table 2. Note that the incremental development is calculated as $\log Y_{j+1} / Y_j$. One can see that due to the difference in volume, the volatility of loss development for the commercial auto line is wiped out if we analyze the aggregated data. Therefore, the vine copula model can be considered as a flexible generalization of two extreme models, silo and independent approaches, so that one can consider possible dependence and heterogeneity among the lines of business simultaneously.

Table 2. An illustrative example of heterogeneity among business lines.

Year	Paid Losses			Incremental Development		
	Personal	Commercial	Aggregate	Personal	Commercial	Aggregate
2	5,000,000	800,000	5,800,000	-	-	-
3	5,200,000	1,200,000	6,400,000	3.92%	40.55%	9.84%
4	5,300,000	1,500,000	6,800,000	1.90%	22.31%	6.06%

4. Model Selection and Parameter Estimation

For marginal distribution, we apply the idea of the cross-classified model for each line of business as in Shi and Frees (2011) and Taylor and McGuire (2016), which regresses $Y_{ij}^{(n)}$ to linear predictor $\gamma^{(n)} + \alpha_i^{(n)} + \delta_j^{(n)}$ as follows:

$$\begin{aligned} \text{Lognormal:} \quad & \mathbb{E}[\log Y_{ij}^{(n)}] = \gamma^{(n)} + \alpha_i^{(n)} + \delta_j^{(n)}, \\ \text{Gamma:} \quad & \mathbb{E}[Y_{ij}^{(n)}] = \exp(\gamma^{(n)} + \alpha_i^{(n)} + \delta_j^{(n)}), \end{aligned} \tag{6}$$

where $\alpha_i^{(n)}$ corresponds to the effect for the i^{th} accident year and $\delta_j^{(n)}$ corresponds to the cumulative effect up to j^{th} development lag, for n^{th} line of business, respectively.

Recall that it is in our interests to predict $\mathbb{E} \left[Y_{i,j+1}^{(n)} | Y_{i,j}^{(n)} \right] = Y_{i,j}^{(n)} \mathbb{E} \left[D_{i,j+1}^{(n)} \right]$. Therefore, one can rewrite (6) in the following way:

$$\begin{aligned} \text{Lognormal: } & \mathbb{E} \left[\log D_{i,j+1}^{(n)} \right] = \eta_{j+1}^{(n)}, \\ \text{Gamma: } & \mathbb{E} \left[D_{i,j+1}^{(n)} \right] = \exp(\eta_{j+1}^{(n)}), \end{aligned} \tag{7}$$

where $\eta_{j+1}^{(n)} = \delta_{j+1}^{(n)} - \delta_j^{(n)}$ and $D_{i,j+1}^{(n)} = Y_{i,j+1}^{(n)} / Y_{i,j}^{(n)}$. Therefore, $\eta_{j+1}^{(n)}$ can be interpreted as the incremental loss development factor from j^{th} lag to $j + 1^{\text{th}}$ lag. In this regard, we utilize two candidate distributions in (7) and compare the model selection diagnostics to choose a suitable distribution for each of marginal models.

Further, it is natural to expect that for a fixed accident year, paid claim amounts gradually increase until they are fully mature and developed, while the magnitude of development gets smaller as development lag increases, which means the incremental loss development factor is expected to be greater than 1 but monotonically decreasing until it converges to 1. In terms of cross-classified model, such a statement is equivalent to the following mathematical condition:

$$\eta_2^{(n)} \geq \eta_3^{(n)} \geq \dots \eta_{L(n)}^{(n)} = 0, \tag{8}$$

for a large enough integer $L(n)$, or there are $\zeta_t^{(n)} \in [0, \infty)$ for $t = 2, \dots, L(n)$ and $n = 1, \dots, 9$ such that $\eta_j^{(n)} = \sum_{t=j}^{L(n)} \zeta_t^{(n)}$. Therefore, we may suggest the following four models as the candidates for marginal distribution of each line of business:

- LNU: The cross-classified lognormal model in (7);
- LNC: The cross-classified lognormal model in (7) with (8);
- GamU: The cross-classified gamma model in (7);
- GamC: The cross-classified gamma model in (7) with (8).

All of the four models are fitted via RStan because of its flexibility. For example, one can incorporate the constraint in (8) by forcing a lower limit of $\zeta_t^{(n)}$ as 0 in RStan and using a diffuse uniform prior on a positive real number for $\zeta_t^{(n)}$, $n = 1, \dots, 9$, and $t = 1, \dots, J$. For the unconstrained models, a diffuse uniform prior on the positive real number is directly used for $\eta_{j+1}^{(n)}$, $n = 1, \dots, 9$, and $j = 2, \dots, J$. For each marginal model, four chains with 1000 iterates are used; the first 500 iterates are discarded for burn-in, which usually requires computation time for MCMC sampling of less than a second.

Table 3 provides a summary of the estimates of parameters in all models for the marginal components, which are summarized with posterior mean and upper/lower bounds of 90% of the Bayesian credible interval for each parameter. One can see that $\hat{\eta}_9^{(n)}$ tends to be greater than or equal to $\hat{\eta}_8^{(n)}$ in the unconstrained models, whereas $\hat{\eta}_8^{(n)} > \hat{\eta}_9^{(n)}$ for all $n = 1, \dots, 9$ in the constrained models, which is a more natural pattern in loss development analysis.

When posterior MCMC samples of parameters are obtained, it is quite important to make sure that these samples converge to (proper) posterior distributions. In this regard, we use \hat{R} statistics proposed by Gelman and Rubin (1992) and traceplots which enable us to judge the convergence of MCMC sampling with the generated samples. Basically, \hat{R} statistics tell us that the samples from distinct chains are well-mixed so that the chains are considered converged if $\hat{R} \simeq 1$. Table A1 in Appendix A shows that $\hat{R} \simeq 1$ for all parameters, which tells us that the chains are well-mixed. One can also see that MCMC chains are well-mixed by looking at selected traceplots of maginal models, as shown in Figure A1. It indicates different MCMC chains end up with similar empirical distributions, which is a necessary condition of the convergence of the MCMC algorithm to the correct posterior distribution.

After the convergence of Bayesian models is assessed, we compare the goodness-of-fit of Bayesian models using DIC and LPML. Based on the DIC and LPML values for fitted marginal models in Table 4, it turns out that lognormal distribution is favored compared to gamma distribution. Further, DIC values of constrained lognormal distribution are always less than ones of unconstrained lognormal distribution, while LPML values have no definitive patterns between constrained and unconstrained lognormal distributions. Therefore, we suggest to use constrained lognormal distribution as the marginal distribution of every line of business, based on the model selection diagnostics and interpretability of the regression coefficients while unconstrained lognormal distributions are considered as a benchmark.

Once a marginal distribution for each triangle is specified, the dependence structure can be modeled with a 9-dimensional vine copula. Since marginal distributions are fixed and calibrated, it is possible to optimize (2). For the optimization routine, we use the algorithm of [Dissmann et al. \(2013\)](#) which is readily available in an R package *VineCopula*, as a function `RVineStructureSelect`, which explores both the optimal vine structure and the choice of family for each of the pair copulas sequentially. For a detailed explanation of such an implementation, see Chapter 8 of [Czado \(2019\)](#).

In our search for the vine structure, we also consider the sparsity of the vine copula. Since we need $\binom{d}{2}$ pair copulas to construct a d -dimensional vine structure, the number of required pair copulas quadratically increases as the dimension increases, which adds both complexity and computational burden to the optimization scheme. Therefore, in this article, we also incorporate the idea of a sparse copula so that a pair copula is considered to be an independent copula unless the estimated association Kendall's Tau is significantly different from 0 for the pair copula. More specifically, we use the following test statistic T based on the estimated value of Kendall's Tau, $\hat{\tau}$ as proposed in [Genest and Favre \(2007\)](#):

$$T = \sqrt{\frac{9m(m-1)}{2(2m+5)}} \times |\hat{\tau}|,$$

where T asymptotically follows a standard normal distribution if m is large enough. For more approaches of incorporating sparsity in a high-dimensional vine copula, see [Gruber and Czado \(2018\)](#) and [Nagler et al. \(2019\)](#).

According to the optimization with the data, the optimal vine structure is given Table 5 which only displays non-independent pair copulas. In the table, ϕ and τ mean estimated copula parameter and Kendall's Tau for each pair copula, respectively. Some empirical relationships are observed by the estimated copula parameters, which seem intuitive considering the nature of the lines of business. For example, it might be quite natural to have strongly positive dependence among the development patterns of line 4 (North American non-casualty) and line 6 (overseas general non-casualty) due to inherent characteristics of the same claim type. Dependence between line 8 (global reinsurance property) and line 9 (global reinsurance non-property) is captured by the Gumbel copula, which indicates there is positive association between claim settlements in the initial years (that corresponds to upper tail part) for a global reinsurance business unit. Such a positive tail dependence might have originated from an internal decision on the claim processing in the same business unit. Note that a company has limited resources to deal with litigations related to the claim adjustments so that claim adjustments can be delayed in a line by focusing on expeditious claim adjustments in another line, which might explain negative associations between lines 6 and 9, lines 2 and 7, and lines 5 and 9. One can see that the estimated vine structure exhibits enough of a degree of sparsity because only 10 pair copulas are non-independent among $\binom{9}{2} = 36$ pair copulas in the vine structure so that most of the dependencies arise from tree 1, as shown in Figure 2. There are nine lines of business that correspond to the nodes of tree 1 so that we have nine nodes and eight edges as an R-vine structure which captures the first layer of dependence among the lines of business. Since the dependence of a pair of copulas is quite weak except for those in tree 1, we only display the R-vine structure for tree 1.

Table 3. Estimated regression coefficients for marginal models.

Unconstrained Lognormal Model																											
	T1			T2			T3			T4			T5			T6			T7			T8			T9		
	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
η_2	0.79	0.73	0.85	1.26	1.19	1.33	0.66	0.62	0.70	0.36	0.32	0.40	0.72	0.70	0.74	0.81	0.79	0.83	0.53	0.53	0.54	1.45	1.28	1.61	1.80	1.74	1.86
η_3	0.32	0.26	0.38	0.61	0.54	0.69	0.22	0.18	0.26	0.05	0.02	0.09	0.27	0.24	0.29	0.19	0.18	0.21	0.09	0.08	0.09	0.28	0.10	0.47	0.68	0.62	0.74
η_4	0.19	0.13	0.26	0.43	0.35	0.51	0.13	0.09	0.17	0.03	0.00	0.07	0.16	0.14	0.19	0.07	0.05	0.10	0.03	0.03	0.04	0.14	0.02	0.31	0.38	0.31	0.44
η_5	0.14	0.07	0.21	0.24	0.16	0.33	0.08	0.03	0.13	0.03	0.00	0.07	0.11	0.08	0.14	0.03	0.01	0.05	0.01	0.01	0.02	0.12	0.01	0.28	0.26	0.19	0.33
η_6	0.10	0.02	0.18	0.21	0.12	0.31	0.06	0.01	0.11	0.03	0.00	0.07	0.07	0.04	0.10	0.02	0.00	0.04	0.01	0.00	0.02	0.12	0.01	0.30	0.17	0.10	0.25
η_7	0.08	0.01	0.16	0.10	0.02	0.21	0.04	0.00	0.09	0.03	0.00	0.08	0.05	0.01	0.08	0.02	0.00	0.04	0.01	0.00	0.02	0.14	0.01	0.34	0.12	0.03	0.20
η_8	0.08	0.01	0.18	0.10	0.01	0.21	0.04	0.00	0.10	0.04	0.00	0.10	0.04	0.01	0.08	0.02	0.00	0.05	0.01	0.00	0.02	0.17	0.01	0.43	0.08	0.01	0.17
η_9	0.10	0.01	0.24	0.11	0.01	0.26	0.06	0.01	0.13	0.06	0.00	0.14	0.04	0.00	0.09	0.03	0.00	0.06	0.01	0.00	0.03	0.23	0.02	0.57	0.10	0.01	0.22
Unconstrained Gamma Model																											
	T1			T2			T3			T4			T5			T6			T7			T8			T9		
	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
η_2	0.81	0.76	0.86	1.27	1.21	1.34	0.67	0.63	0.70	0.37	0.33	0.41	0.72	0.70	0.74	0.81	0.79	0.83	0.53	0.53	0.54	1.60	1.43	1.76	1.81	1.76	1.87
η_3	0.32	0.27	0.38	0.62	0.55	0.69	0.22	0.18	0.26	0.05	0.01	0.10	0.27	0.24	0.29	0.19	0.17	0.21	0.09	0.08	0.09	0.29	0.11	0.47	0.68	0.63	0.74
η_4	0.19	0.13	0.25	0.44	0.36	0.51	0.13	0.09	0.17	0.03	0.00	0.07	0.16	0.14	0.19	0.08	0.05	0.10	0.03	0.03	0.04	0.14	0.02	0.31	0.38	0.32	0.44
η_5	0.14	0.07	0.20	0.25	0.17	0.33	0.08	0.03	0.13	0.03	0.00	0.07	0.11	0.09	0.14	0.03	0.01	0.05	0.01	0.01	0.02	0.12	0.01	0.29	0.26	0.20	0.33
η_6	0.09	0.02	0.17	0.21	0.13	0.30	0.06	0.02	0.11	0.03	0.00	0.07	0.07	0.04	0.10	0.02	0.00	0.04	0.01	0.00	0.02	0.13	0.01	0.31	0.17	0.10	0.25
η_7	0.07	0.01	0.15	0.10	0.02	0.20	0.03	0.00	0.08	0.03	0.00	0.09	0.04	0.01	0.08	0.02	0.00	0.04	0.01	0.00	0.02	0.14	0.01	0.35	0.12	0.04	0.20
η_8	0.08	0.01	0.17	0.10	0.01	0.20	0.04	0.00	0.10	0.04	0.00	0.11	0.04	0.01	0.08	0.02	0.00	0.05	0.01	0.00	0.02	0.18	0.01	0.45	0.08	0.01	0.17
η_9	0.10	0.01	0.22	0.10	0.00	0.26	0.05	0.00	0.13	0.06	0.01	0.15	0.03	0.00	0.08	0.02	0.00	0.06	0.01	0.00	0.02	0.25	0.02	0.64	0.09	0.01	0.22
Constrained Lognormal Model																											
	T1			T2			T3			T4			T5			T6			T7			T8			T9		
	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
η_2	0.79	0.74	0.85	1.26	1.19	1.32	0.66	0.63	0.70	0.36	0.32	0.40	0.72	0.70	0.74	0.81	0.79	0.83	0.53	0.53	0.54	1.45	1.29	1.61	1.80	1.75	1.85
η_3	0.33	0.27	0.38	0.61	0.54	0.69	0.21	0.18	0.25	0.07	0.04	0.11	0.27	0.24	0.29	0.19	0.17	0.21	0.09	0.08	0.09	0.33	0.20	0.48	0.68	0.62	0.73
η_4	0.20	0.15	0.26	0.43	0.36	0.50	0.14	0.10	0.17	0.05	0.03	0.08	0.16	0.14	0.19	0.08	0.05	0.10	0.03	0.03	0.04	0.21	0.12	0.33	0.38	0.32	0.44
η_5	0.15	0.10	0.20	0.27	0.20	0.34	0.09	0.06	0.12	0.04	0.02	0.06	0.11	0.09	0.14	0.04	0.02	0.06	0.02	0.01	0.02	0.16	0.08	0.25	0.26	0.20	0.32
η_6	0.11	0.07	0.16	0.20	0.13	0.27	0.06	0.04	0.09	0.03	0.01	0.05	0.08	0.05	0.10	0.02	0.01	0.04	0.01	0.01	0.02	0.12	0.05	0.20	0.18	0.13	0.24
η_7	0.08	0.03	0.12	0.13	0.06	0.20	0.04	0.02	0.07	0.02	0.01	0.04	0.05	0.03	0.07	0.02	0.01	0.03	0.01	0.00	0.01	0.09	0.03	0.16	0.12	0.07	0.18
η_8	0.05	0.01	0.10	0.08	0.02	0.15	0.03	0.01	0.05	0.01	0.00	0.03	0.03	0.01	0.06	0.01	0.00	0.02	0.00	0.00	0.01	0.06	0.01	0.12	0.08	0.03	0.13
η_9	0.03	0.00	0.07	0.04	0.00	0.10	0.01	0.00	0.03	0.01	0.00	0.02	0.02	0.00	0.04	0.01	0.00	0.01	0.00	0.00	0.01	0.03	0.00	0.08	0.04	0.00	0.09

Table 3. Cont.

	Constrained Gamma Model																										
	T1			T2			T3			T4			T5			T6			T7			T8			T9		
	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
η_2	0.81	0.76	0.86	1.27	1.21	1.33	0.67	0.63	0.70	0.37	0.33	0.40	0.72	0.70	0.74	0.81	0.79	0.83	0.53	0.53	0.54	1.60	1.44	1.76	1.81	1.76	1.86
η_3	0.33	0.27	0.38	0.62	0.55	0.69	0.22	0.18	0.25	0.07	0.04	0.10	0.27	0.24	0.29	0.19	0.18	0.21	0.09	0.08	0.09	0.33	0.21	0.48	0.68	0.63	0.74
η_4	0.20	0.15	0.25	0.44	0.37	0.51	0.13	0.10	0.17	0.05	0.03	0.07	0.16	0.14	0.19	0.08	0.06	0.10	0.03	0.03	0.04	0.22	0.12	0.32	0.38	0.32	0.44
η_5	0.15	0.10	0.19	0.27	0.20	0.34	0.09	0.06	0.12	0.04	0.02	0.06	0.11	0.09	0.14	0.04	0.02	0.06	0.02	0.01	0.02	0.16	0.08	0.25	0.26	0.21	0.32
η_6	0.11	0.06	0.15	0.20	0.14	0.27	0.06	0.03	0.09	0.03	0.01	0.05	0.08	0.05	0.10	0.02	0.01	0.04	0.01	0.01	0.02	0.12	0.05	0.20	0.18	0.13	0.24
η_7	0.08	0.04	0.12	0.13	0.06	0.19	0.04	0.02	0.07	0.02	0.01	0.04	0.05	0.03	0.07	0.02	0.01	0.03	0.01	0.00	0.01	0.09	0.03	0.16	0.12	0.07	0.18
η_8	0.05	0.02	0.09	0.08	0.02	0.15	0.03	0.01	0.05	0.01	0.00	0.03	0.03	0.01	0.06	0.01	0.00	0.02	0.00	0.00	0.01	0.06	0.01	0.12	0.08	0.03	0.13
η_9	0.03	0.00	0.06	0.04	0.00	0.10	0.01	0.00	0.04	0.01	0.00	0.02	0.02	0.00	0.04	0.01	0.00	0.01	0.00	0.00	0.01	0.03	0.00	0.08	0.04	0.00	0.09

Table 4. DIC and LPML for marginal components.

		T1	T2	T3	T4	T5	T6	T7	T8	T9
DIC	LNU	-57.1	-46.5	-90.3	-84.2	-125.1	-140.5	-224	16.8	-61.8
	GamU	-38.6	-7.7	-74.4	-75.3	-105.9	-123.6	-214.1	47.6	-14.9
	LNC	-62.9	-50.5	-95.9	-88.4	-129.2	-144.1	-228.5	11.9	-66.7
	GamC	-44.2	-11.8	-79.2	-79.6	-110.5	-127.3	-218	42.8	-19.2
LPML	LNU	-326.2	-129.3	-117.2	-291.7	-83.2	-127.3	-43.2	-241.8	-155.2
	GamU	-379.2	-168.6	-144.3	-364.8	-103.1	-162	-62.1	-342.7	-215.2
	LNC	-368.1	-132.7	-123.4	-300.6	-82.6	-118.1	-43.4	-242	-144.8
	GamC	-419.6	-176.4	-145.5	-392.5	-100.1	-161.3	-57.5	-350.5	-219.5

Table 5. Specification of vine structure.

Tree	Edge	Family	ϕ	τ
1	9,8	Gumbel	1.87	0.46
1	9,6	90° rotated Clayton	-1.20	-0.37
1	8,1	Frank	9.40	0.65
1	1,7	270° rotated Clayton	-0.20	-0.09
1	7,2	90° rotated Clayton	-0.95	-0.32
1	6,4	Frank	5.87	0.50
1	9,5	270° rotated Gumbel	-1.24	-0.20
1	4,3	Survival Gumbel	1.61	0.38
2	9,4;6	90° rotated Gumbel	-1.35	-0.26
3	9,7;8,1	Survival Clayton	0.52	0.21

We also visualize the dependence structure via normalized contour plots in Figure 3, which displays the association between two probability integral transforms in a normalized scale. For example, if a pair copula $c(u_1, u_2)$ is given, then the corresponding normalized contour plot gives us a contour plot with the following density g :

$$g(\Phi^{-1}(u_1), \Phi^{-1}(u_2)) = c(u_1, u_2)\phi(\Phi^{-1}(u_1))\phi(\Phi^{-1}(u_2)).$$

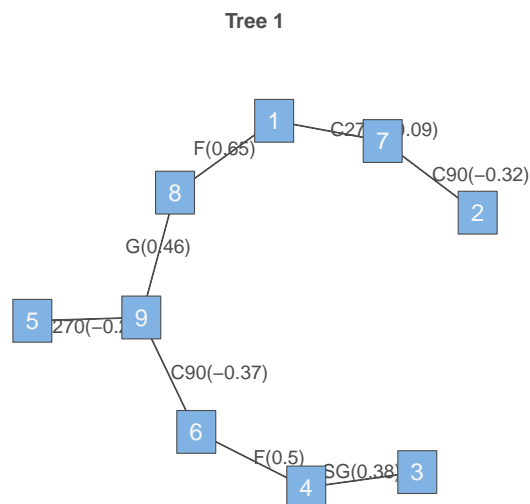


Figure 2. R-vine structure for tree 1.

Normalized contour plots for trees 1, 2, and 3 in the chosen vine structure are provided in Figure 3. One can see that some pair copulas (for example, $c_{4,3}$ and $c_{8,1}$) demonstrate strong tail dependence or skewed shape of the contours, which cannot be captured with naive use of eillptical copulas such as Gaussians or t -copulas. Therefore, Figure 3 substantiates that the choice of pair copula families are enough to describe the dependence structure observed in the data. In the figure, the first row corresponds to normalized contour plots for tree 3 where one can see concentric circles as contour plots, which indicates independence of corresponding pair copulas. On the other hand, one can see many skewed contour plots for tree 3 from the third row of the figure, which indicates varying dependence of corresponding pair copulas.

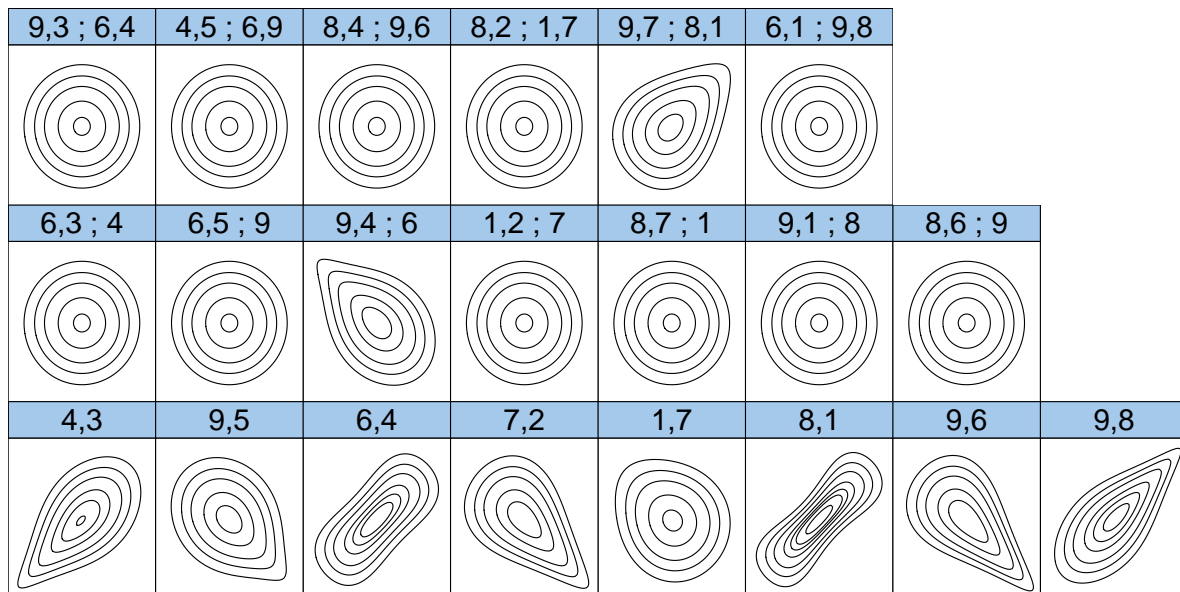


Figure 3. Normalized contour plots for trees 1, 2, and 3.

5. Validation and Actuarial Implication

Once a predictive model is calibrated based on the upper loss triangles $\mathcal{D}_{1:9}$, which is defined in (3), it needs to be validated based on available \mathcal{D}_{9+k} , which is defined in (4). For this purpose, $\mathcal{D}_{10} = \{Y_{ij}^{(n)} : 2 \leq i \leq 9 \text{ and } j = 11 - i, n = 1, \dots, 9\}$, the latest diagonals are used as a validation set.

Under the proposed lognormal model for marginal components, one can show that

$$\mathbb{E} \left[Y_{i,j+1}^{(n)} | Y_{i,j}^{(n)} \right] = \mathbb{E} \left[D_{i,j+1}^{(n)} \right] Y_{i,j}^{(n)} = \exp \left(\eta_{i,j+1}^{(n)} + \frac{1}{2} \sigma^{2(n)} \right) Y_{i,j}^{(n)},$$

so that a point estimate of $Y_{i,11-i}^{(n)}$ is given as $\hat{Y}_{i,11-i}^{(n)} = Y_{i,10-i}^{(n)} \times \exp \left(\hat{\eta}_{11-i}^{(n)} + \frac{1}{2} \hat{\sigma}^{2(n)} \right)$, where $\hat{\eta}^{(n)}$ and $\hat{\sigma}^{(n)}$ are obtained as posterior means from the MC samples of parameters. Since it is of interest to predict the future unpaid claim L as defined in (5), here we apply three models to project the unpaid claims; the independent model which assumes independence among all lines of business, the silo model which assumes a perfect positive association among all lines of business, and the copula model which utilizes the vine copula structure specified and optimized in Section 4.

Due to linearity of expectation, it is natural to expect that point estimates of reserves under the copula model would be more or less the same as the estimates under independent model. However, insurance companies are often interested in not only the predictive mean, but also the confidence interval of the estimated mean and risk measures for enterprise risk management. In this regard, the proposed vine copula model allows us to consider the impacts of associations among the lines of business that might lead to more accurate estimation of risk measures. For example, if there are positive associations among the lines of business, then the independent model will underestimate

Value at Risk (VaR) or Conditional tail expectation (CTE) since a worse scenario can affect all lines of business adversely.

Table 6 summarizes the point estimates of unpaid claims for the subsequent calendar year based on the copula model and the silo model. Note that point estimates of unpaid claims from the independent model should be the same as ones from the copula model, as long as we model each line of business with the same marginal distribution. Further, although we primarily use constrained models for the marginal components, here we also provide the point estimates of unpaid claims from the models where we do not impose any constraints on the parameters.

From the table, it is shown that copula model with constraints turns out to be the best model in terms of prediction of aggregate unpaid claims for the subsequent year. We observe that under unconstrained models, the predicted values of unpaid claims can be severely exaggerated, especially for mature years (in our case, AY = 2005). That is because the value of $\eta_9^{(n)}$ can be overestimated in the unconstrained models, as observed in Table 3. Therefore, it implies that naive implementation of unconstrained model for multi-line reserving problem may end up with poor prediction.

Table 6. Summary of unpaid claims prediction with the fitted models.

	Unconstrained		Constrained		Actual
	Copula	Silo	Copula	Silo	
AY = 2005	529,426	219,222	148,767	81,657	90,662
AY = 2006	264,689	167,787	185,653	131,065	109,420
AY = 2007	299,024	196,099	319,881	227,435	209,649
AY = 2008	454,503	404,983	461,885	414,982	338,359
AY = 2009	405,808	386,094	447,697	409,172	304,544
AY = 2010	651,228	610,943	703,166	615,770	450,724
AY = 2011	942,324	1,071,867	986,753	1,070,549	762,328
AY = 2012	2,182,202	2,649,858	2,181,708	2,648,604	1,837,070
Total	5,729,205	5,706,854	5,435,511	5,599,236	4,102,756

In order to evaluate the prediction performance, here we use validation measures such as root mean squared error (RMSE) and mean absolute error (MAE) defined as follows:

$$RMSE =: \sqrt{\frac{1}{8} \sum_{i=2}^9 (\hat{Y}_{i,11-i}^\bullet - Y_{i,11-i}^\bullet)^2}, \quad MAE =: \frac{1}{8} \sum_{i=2}^9 |\hat{Y}_{i,11-i}^\bullet - Y_{i,11-i}^\bullet|,$$

where $\hat{Y}_{i,11-i}^\bullet$ is the estimate of $Y_{i,11-i}^\bullet$ under each method. We prefer the models with smaller values of RMSE and/or MAE. According to Table 7, the copula model with constraints still turns out to be the best model in terms of RMSE and MAE, which evaluate the performance of predicting the point estimates.

Table 7. Summary of validation performance measures of the fitted models.

	Unconstrained		Constrained	
	Copula	Silo	Copula	Silo
RMSE	234,540	318,849	190,381	315,934
MAE	203,306	203,900	166,594	189,311

However, it is of less concern to get point estimates of L , the aggregate unpaid claims. Actuaries (and companies) want to know a range of estimates, or the predictive distribution of the aggregate unpaid claims to assess the quality of decision making on reserve and retain appropriate amounts of risk capital for the whole company. Thus, here we provide a way to simulate predictive samples of

$L^{(n)}$ and describe the predictive distribution of $L^{(n)}$ for $n = 1, \dots, 9$ and subsequently $L = \sum_{n=1}^9 L^{(n)}$ with the following steps, which are similar to those of Gao (2018):

1. Generate a 9-dimensional uniform random vector $(u_{ij}^{r:(1)}, \dots, u_{ij}^{r:(9)})$ based on the specified copula structure in a sequential way. For example, one can first generate $(u_{ij}^{r:(1)}, u_{ij}^{r:(2)})$ from C_{12} . After that, $u_{ij}^{r:(3)}$ is generated subsequently due to the following identity:

$$\mathcal{U}[0, 1] \stackrel{d}{=} w_3 = C_{3|12}(u_3|u_1, u_2) = h_{3|1,2} \left(C_{3|2}(u_3|u_2), C_{1|2}(u_1|u_2) \right),$$

where $h_{3|1,2}(w, v) = \frac{\partial}{\partial w} C_{13;2}(w, v)$. One can continue such implementation to obtain a random sample of d -dimensional uniform vector, which is readily available with RVineSim function in R package VineCopula. (For details, see Chapter 6 of Czado (2019).)

2. Based on the lognormal assumption of marginal components, $(D_{ij}^{r:(1)}, \dots, D_{ij}^{r:(9)})$ are generated as follows; $D_{ij}^{r:(n)} = \exp \left(\hat{\eta}_j^{r:(n)} + \Phi^{-1}(u_{ij}^{r:(n)}) \hat{\sigma}^{r:(n)} \right)$ where $\hat{\eta}_j^{r:(n)}$ and $\hat{\sigma}^{r:(n)}$ are r^{th} MC samples of $\eta_j^{(n)}$ and $\sigma^{(n)}$ from the marginal model for the n^{th} line of business, respectively.
3. Repeat steps 1 to 2 to get $L^{r:(n)}$, the MC samples of $L^{(n)}$ for $r = 1, \dots, R$ where

$$L^{r:(n)} = \sum_{i=2}^9 \hat{Y}_{i,11-i}^{r:(n)} - Y_{i,10-i}^{(n)} = \sum_{i=2}^9 \left(D_{i,11-i}^{r:(n)} - 1 \right) Y_{i,10-i}^{(n)} \text{ and } L^{r \bullet} = \sum_{n=1}^9 L^{r:(n)}.$$

Since $\{Y_{i,10-i}^{(n)} : 2 \leq i \leq 9, n = 1, \dots, 9\} \subset \mathcal{D}_{1:9}$, only $\{D_{i,11-i}^{(n)} : 2 \leq i \leq 9, n = 1, \dots, 9\}$ needs to be simulated to describe the distribution of L . In the simulation scheme, possible uncertainties of estimated parameter values are already considered since the marginal parameters are estimated with the Bayesian approach.

Figure 4 shows the predictive kernel densities of L based on the Monte Carlo samples generated with the aforementioned approach from the constrained model. In the silo approach, all nine triangles are aggregated into one triangle based on the development lags and accident years, so that loss development patterns of all business lines are implicitly assumed to be identical, which may not be true in some cases. One can see that the predictive distribution of L under the copula model is more jagged than one under the independent model, but the copula model is still closer to the independent model than the silo model. It agrees with the proposed vine copula structure described in Figure 2, which shows a sufficient degree of sparsity.

The discrepancy among predictive distributions based on copula, independent, and silo models also can be quantified via Hellinger distance. Hellinger distance, originally proposed by Hellinger (1909), is used to quantify the "distance" between two distributions. For example, if there are two functions f and g which are probability density functions of two distributions F and G , the square of Hellinger distance is calculated as follows:

$$H^2(F, G) = 1 - \int \sqrt{f(x)g(x)} dx.$$

One can see that $0 \leq H(F, G) \leq 1$ is always satisfied with Hellinger distance so that it is well calibrated and interpreted, where $H(F, G) = 0$ means F and G are identical almost surely and $H(F, G) = 1$ means F and G are totally different. Due to this property, Hellinger distance has been widely used in the statistics literature, such as Beran (1977), and estimation of $H(F, G)$ using generated samples from F and G can be easily implemented via R packages such as statip. Here, $\hat{H}^2(L|Copula, L|Independent)$, the square of the Hellinger distance between the kernel density of L under copula model with constraints (illustrated with blue solid line in Figure 4) and the density under the independent model with constraints (illustrated with black dotted line in Figure 4) is about 4.33%

while $\hat{H}^2(L|Copula, L|Silo) = 13.54\%$. Therefore, one can see that the proposed vine copula structure can capture weak dependencies among the lines of business.

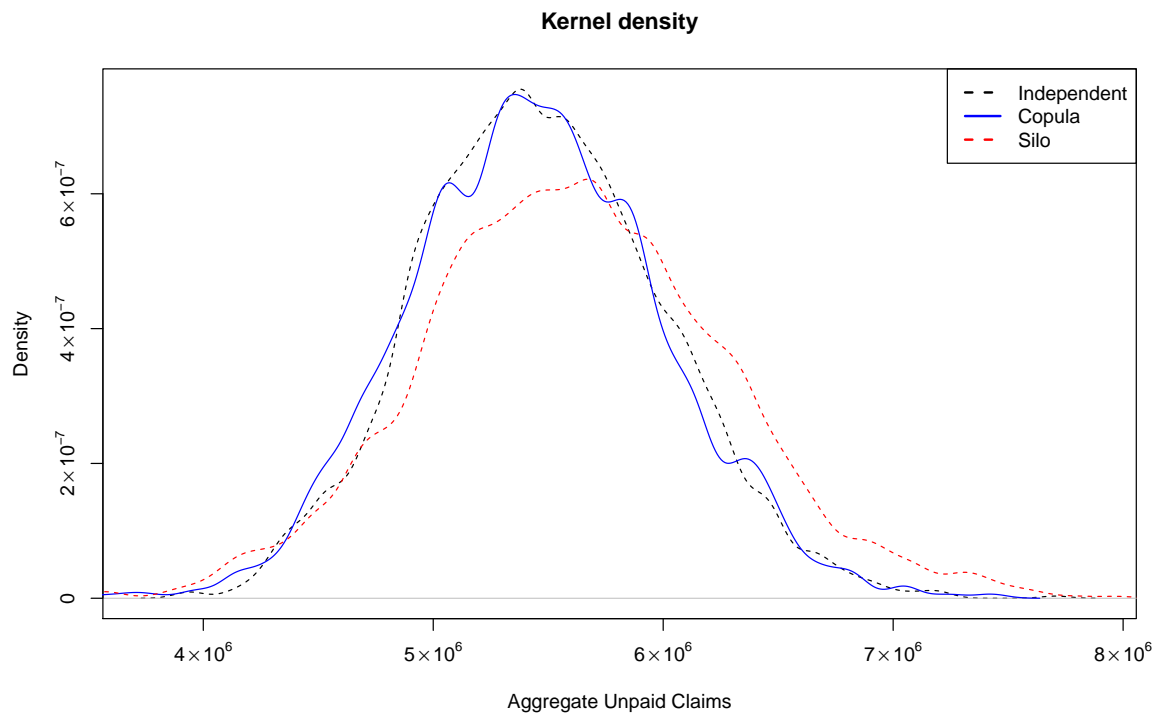


Figure 4. Predictive density of aggregate unpaid claims for constrained models.

Note that Hellinger distance is a special case of the functional Bregman divergence proposed in Goh and Dey (2014) and Jeong (2020), which is given as follows:

$$D_\psi(F, G) = \int \psi \left(\frac{f(x)}{g(x)} \right) g(x) dx,$$

where $\psi(\cdot)$ is a strictly convex and differentiable function with positive support and $\psi(1) = 0$. One can see that $D_\psi(F, G) = H^2(F, G)$ if $\psi(z) = 1 - \sqrt{z}$, and $D_\psi(F, G)$ is Kullback–Leibler divergence if $\psi(z) = -\log z$.

Based on the predictive distributions of L and $L^{(n)}$ for $n = 1, \dots, 9$, one can calculate the risk margins for the unpaid claims as in Table 8. Both VaR and CTE are estimated in an empirical way—in other words, based on the generated samples of the predictive distribution of L or $L^{(n)}$. Usually, it is expected that a company can get a benefit from the diversification of risks for multiple lines of business so that the risk capital for aggregate reserve is less than the sum of the risk capitals for the reserve of each line, which is known as subadditivity of risk measure. Although subadditivity is not guaranteed for VaR, it is guaranteed for CTE and we can quantify the diversification effects by taking the differences of risk capitals by subtracting the risk capital for aggregate reserve from the sum of the risk capitals for the reserve of each line. Table 9 summarizes measured diversification effects for given models. We can observe that the measured diversification effect of the copula model is much greater than that of the silo model because of weak dependence captured in the vine structure, although it is still less than that of the independent model, which assumes perfect independence among the lines of business.

Table 8. Estimated risk margins for the unpaid claims.

	VaR			CTE		
	90%	95%	99%	90%	95%	99%
Marginal T1	384,383	420,016	494,059	432,782	464,901	521,869
Marginal T2	1,403,464	1,520,387	1,720,617	1,560,576	1,657,167	1,893,831
Marginal T3	655,532	694,812	766,766	707,338	740,161	805,153
Marginal T4	1,657,500	1,802,081	2,126,876	1,858,735	1,993,388	2,313,396
Marginal T5	777,885	807,011	869,457	818,302	844,170	895,520
Marginal T6	691,142	714,401	757,403	721,538	741,557	791,119
Marginal T7	405,697	413,424	428,243	416,023	423,117	439,641
Marginal T8	586,402	696,629	1,009,717	757,769	878,869	1,199,343
Marginal T9	556,892	591,841	658,353	603,369	633,144	703,385
Marginal Total	7,118,897	7,660,602	8,831,491	7,876,431	8,376,473	9,563,257
Independent Aggregate	6,163,307	6,363,592	6,755,125	6,428,484	6,608,753	6,978,484
Copula Aggregate	6,145,848	6,376,847	6,774,463	6,435,712	6,610,684	6,995,674
Silo Aggregate	6,434,983	6,692,370	7,309,860	6,815,535	7,076,775	7,586,896

Table 9. Risk reduction from diversification.

	VaR			CTE		
	90%	95%	99%	90%	95%	99%
Independent	955,590	1,297,010	2,076,366	1,447,947	1,767,720	2,584,773
Copula	973,049	1,283,755	2,057,028	1,440,719	1,765,789	2,567,583
Silo	683,914	968,232	1,521,631	1,060,896	1,299,698	1,976,361

6. Practical Issues for Implementation

Note that a loss reserving dataset from a single company was utilized in this article because a primary insurer will be more interested in analysis of dependence among the lines of business in the same company for effective enterprise risk management, rather than whole industry. However, this methodology can be applied for analysis of reserving trend in a dataset of several companies as well. For example, analyzing dependencies among the loss development profiles of several companies can be of interest for the regulatory authorities and reinsurance companies.

One can be also concerned about the predictive ability of these models, which can be distorted by catastrophic claims. In this regard, insurance companies often decompose the loss development profiles into two layers, primary and excess layers, to minimize the distortion impact from the presence of catastrophic claims (Dew and Hedges 1998), and our proposed methodology can be applied to analyze dependencies among the loss development for the primary layer. Note that such distinction of layers is usually internal and may not be disclosed to the public, which also applies to the dataset used in this article.

7. Conclusions

In this article, we explored and introduced a novel approach which considered possible dependence among the multiple lines of business via vine copula. Use of a vine copula is very flexible and it allows us to model high-dimensional dependence among the business lines, not only bivariate dependence. In the case of traditional copula models, one needs to calibrate many copula models with different families, and go through model validation procedure to choose the best one among the candidates. On the other hand, our proposed vine copula structure enables us to explore the optimal copula structure for the given data in a unified manner.

For model selection, a stepwise approach can be applied to choose the marginal distributions, copula structure, and family for each pair of copulas subsequently. Our empirical analysis on a

synthetic insurance portfolio consisting of nine lines of business showed weak associations among the lines.

Further, it was also shown that the naive implementation of a cross-classified model may result in a counterintuitive pattern of loss development, so one might consider a constrained cross-classified model to mitigate this issue. In our work, constraints on the development lag parameters are naturally incorporated via prior elicitation in a Bayesian framework. We expect that a more thorough discussion of the constrained cross-classified model in terms of variable selection could be one of the future directions of this work.

Supplementary Materials: The dataset and code for data analysis are available at https://github.com/ssauljin/vine_copula_reserving.

Author Contributions: Both authors worked on the development of the methodology and proofreading. Data preparation, empirical analysis, and draft preparation was done by H.J. Both authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Society of Actuaries’ James C. Hickman doctoral scholarship.

Acknowledgments: The authors thank anonymous referees and editors for helpful comments that improved this article.

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

Abbreviations

The following abbreviations are used in this manuscript:

MCMC	Markov chain Monte Carlo
DIC	Deviance information criterion
LPML	Logarithm of the pseudo marginal likelihood
CPO	Conditional predictive ordinate
IFM	Inference by margin
ERM	Enterprise risk management
AY	Accident year
RMSE	Root mean squared error
MAE	Mean absolute error
VaR	Value at risk
CTE	Conditional tail expectation

Appendix A

Table A1. \hat{R} values for marginal models.

		Unconstrained Lognormal Model								
		T1	T2	T3	T4	T5	T6	T7	T8	T9
η_2		1.000	1.001	1.000	1.000	1.001	1.002	1.002	1.000	1.003
η_3		1.000	0.999	1.000	1.002	1.000	0.999	1.000	1.003	1.002
η_4		1.000	1.001	0.999	1.000	1.002	0.999	1.001	1.000	1.002
η_5		1.001	1.001	1.002	1.000	1.003	0.999	1.001	1.001	1.000
η_6		1.006	1.001	0.999	1.003	0.999	1.000	1.002	1.000	1.000
η_7		1.000	1.001	1.002	1.000	1.000	1.003	0.999	1.000	1.001
η_8		0.999	1.002	1.000	1.006	0.999	1.001	1.000	1.000	1.000
η_9		1.000	0.999	0.999	1.001	1.002	1.002	1.002	1.002	1.001

Table A1. Cont.

Unconstrained Gamma Model									
	T1	T2	T3	T4	T5	T6	T7	T8	T9
η_2	0.999	1.000	1.000	1.000	1.006	1.000	0.999	1.000	0.999
η_3	1.000	1.000	0.999	1.002	1.001	0.999	1.004	1.000	1.003
η_4	1.002	1.001	1.004	1.000	1.001	1.000	1.004	0.999	1.001
η_5	1.004	1.000	1.002	1.002	0.999	1.001	1.000	1.003	1.002
η_6	1.003	1.002	1.002	1.002	1.000	0.999	0.999	1.000	0.999
η_7	1.000	1.004	1.001	1.001	1.002	0.999	1.001	1.000	1.000
η_8	1.003	1.000	1.003	1.001	1.000	1.003	1.002	1.003	1.000
η_9	1.000	1.002	1.000	1.002	0.999	1.004	1.011	1.000	1.001
Constrained Lognormal Model									
	T1	T2	T3	T4	T5	T6	T7	T8	T9
η_2	0.999	0.999	0.999	1.000	1.001	0.999	1.000	1.000	0.999
η_3	1.002	1.002	1.001	1.001	1.001	1.002	1.003	1.000	1.002
η_4	0.998	1.001	1.002	1.000	1.001	0.999	1.000	0.998	1.002
η_5	1.003	1.000	0.999	0.999	0.999	1.001	1.001	1.000	1.003
η_6	1.002	0.999	0.999	0.999	1.002	1.000	0.999	0.999	1.003
η_7	0.999	1.001	1.000	1.000	0.999	0.999	0.999	0.999	1.000
η_8	1.001	1.002	1.000	1.000	1.001	0.999	1.000	1.001	0.998
η_9	1.001	1.000	1.000	1.001	1.000	0.999	1.001	1.000	0.999
Constrained Gamma Model									
	T1	T2	T3	T4	T5	T6	T7	T8	T9
η_2	0.999	1.003	0.999	1.002	1.000	1.002	0.999	0.999	1.000
η_3	0.999	1.000	0.999	1.004	1.003	1.001	0.999	1.000	0.999
η_4	1.000	0.999	1.003	1.001	1.004	1.002	0.999	0.999	1.004
η_5	0.999	1.001	0.999	1.001	1.001	1.002	0.999	0.999	1.000
η_6	0.999	1.001	1.000	1.000	1.000	0.998	0.999	1.000	1.000
η_7	1.000	1.001	0.999	1.000	0.999	0.999	1.000	1.001	1.000
η_8	1.000	0.999	1.000	1.000	0.998	1.000	1.000	1.001	0.998
η_9	1.001	1.000	1.001	1.000	1.000	1.001	1.000	1.002	1.000

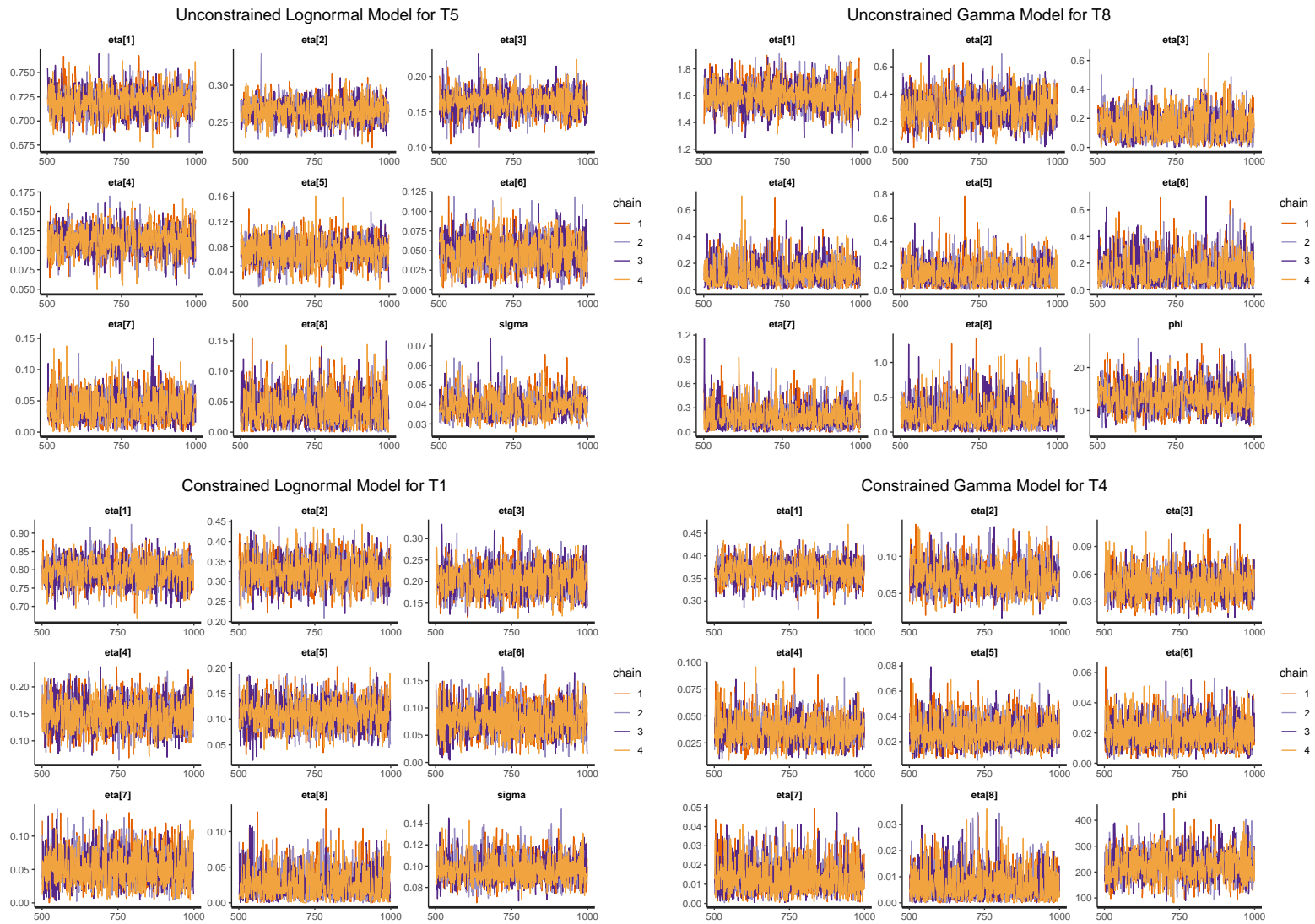


Figure A1. Randomly chosen traceplots for marginal models.

References

- Aas, Kjersti, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics* 44: 182–98. [\[CrossRef\]](#)
- Abdallah, Anas, Jean-Philippe Boucher, and Helene Cossette. 2015. Modeling dependence between loss triangles with hierarchical archimedean copulas. *ASTIN Bulletin: The Journal of the IAA* 45: 577–99. [\[CrossRef\]](#)
- Arreola Hernandez, Jose, Shawkat Hammoudeh, Duc Khuong Nguyen, Mazin A. M. Al Janabi, and Juan Carlos Reboredo. 2017. Global financial crisis and dependence risk analysis of sector portfolios: A vine copula approach. *Applied Economics* 49: 2409–27. [\[CrossRef\]](#)
- Bedford, Tim, and Roger M. Cooke. 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence* 32: 245–68. [\[CrossRef\]](#)
- Bedford, Tim, and Roger M. Cooke. 2002. Vines—A new graphical model for dependent random variables. *The Annals of Statistics* 30: 1031–68. [\[CrossRef\]](#)
- Beran, Rudolf. 1977. Minimum hellinger distance estimates for parametric models. *The Annals of Statistics* 5: 445–63. [\[CrossRef\]](#)
- Braun, Christian. 2004. The prediction error of the chain ladder method applied to correlated run-off triangles. *ASTIN Bulletin: The Journal of the IAA* 34: 399–423. [\[CrossRef\]](#)
- Czado, Claudia. 2010. Pair-copula constructions of multivariate copulas. In *Copula Theory and Its Applications*. Berlin/Heidelberg: Springer, pp. 93–109.
- Czado, Claudia. 2019. *Analyzing Dependent Data with Vine Copulas*. Berlin/Heidelberg: Springer.
- Dew, Edward, and Barton Hedges. 1998. Reserving for excess layers: A guide to practical reserving applications. *Insurance: Mathematics and Economics* 22: 178. [\[CrossRef\]](#)
- Dissmann, Jeffrey, Eike Christian Brechmann, Claudia Czado, and Dorota Kurowicka. 2013. Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis* 59: 52–69.
- Embrechts, Paul, Filip Lindskog, and Alexander McNeil. 2001. *Modelling Dependence with Copulas*. Rapport technique. Zurich: Département de mathématiques, Institut Fédéral de Technologie de Zurich.
- England, Peter D., and Richard J. Verrall. 2002. Stochastic claims reserving in general insurance. *British Actuarial Journal* 8: 443–518. [\[CrossRef\]](#)
- Gao, Guangyuan. 2018. *Bayesian Claims Reserving Methods in Non-life Insurance with Stan*. Berlin/Heidelberg: Springer.
- Geisser, Seymour. 2017. *Predictive Inference*. Abingdon-on-Thames: Routledge.
- Gelfand, Alan E., and Dipak Kumar Dey. 1994. Bayesian model choice: Asymptotics and exact calculations. *Journal of the Royal Statistical Society: Series B (Methodological)* 56: 501–14.
- Gelfand, Alan E., Dipak Kumar Dey, and Hong Chang. 1992. *Model Determination Using Predictive Distributions with Implementation via Sampling-Based Methods*. Technical Report. Stanford: Department of Statistics, Stanford University.
- Gelman, Andrew, and Donald Bruce Rubin. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* 7: 457–72. [\[CrossRef\]](#)
- Genest, Christian, and Anne-Catherine Favre. 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* 12: 347–68. [\[CrossRef\]](#)
- Goh, Gyuhyeong, and Dipak Kumar Dey. 2014. Bayesian model diagnostics using functional bregman divergence. *Journal of Multivariate Analysis* 124: 371–83. [\[CrossRef\]](#)
- Gruber, Lutz Fabian, and Claudia Czado. 2015. Sequential bayesian model selection of regular vine copulas. *Bayesian Analysis* 10: 937–63. [\[CrossRef\]](#)
- Gruber, Lutz Fabian, and Claudia Czado. 2018. Bayesian model selection of regular vine copulas. *Bayesian Analysis* 13: 1107–31. [\[CrossRef\]](#)
- Haff, Ingrid Hobæk, Kjersti Aas, and Arnoldo Frigessi. 2010. On the simplified pair-copula construction—Simply useful or too simplistic? *Journal of Multivariate Analysis* 101: 1296–310. [\[CrossRef\]](#)
- Hellinger, Ernst. 1909. Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen. *Journal für die reine und angewandte Mathematik* 136: 210–71. [\[CrossRef\]](#)
- Hua, Lei, and Harry Joe. 2011. Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis* 102: 1454–71. [\[CrossRef\]](#)
- Ibrahim, Joseph G., Ming-Hui Chen, and Debajyoti Sinha. 2014. *Bayesian Survival Analysis*. Berlin/Heidelberg: Springer.

- Jeong, Himchan. 2020. Testing for random effects in compound risk models via bregman divergence. *ASTIN Bulletin: The Journal of the IAA* 50: 777–98. [CrossRef]
- Joe, Harry, and Dorota Kurowicka. 2011. *Dependence Modeling: Vine Copula Handbook*. Singapore: World Scientific.
- Joe, Harry, and James Jianmeng Xu. 1996. *The Estimation Method of Inference Functions for Margins for Multivariate Models*. Technical Report. Vancouver: University of British Columbia.
- Kreuzer, Alexander, and Claudia Czado. 2019. Bayesian inference for dynamic vine copulas in higher dimensions. *arXiv* arXiv:1911.00702.
- Loaiza, Maya Ruben Albeiro, Jose Eduardo Gomez-Gonzalez, and Luis Fernando Melo Velandia. 2015. Latin american exchange rate dependencies: A regular vine copula approach. *Contemporary Economic Policy* 33: 535–49. [CrossRef]
- Louzada, Francisco, and P. H. Ferreira. 2016. Modified inference function for margins for the bivariate clayton copula-based sun tobit model. *Journal of Applied Statistics* 43: 2956–76. [CrossRef]
- Mack, Thomas. 1993. Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin: The Journal of the IAA* 23: 213–25. [CrossRef]
- Merz, Michael, and Mario V. Wüthrich. 2008. Prediction error of the multivariate chain ladder reserving method. *North American Actuarial Journal* 12: 175–97. [CrossRef]
- Merz, Michael, and Mario V. Wüthrich. 2010. Paid–incurred chain claims reserving method. *Insurance: Mathematics and Economics* 46: 568–79. [CrossRef]
- Nagler, Thomas, Christian Bumann, and Claudia Czado. 2019. Model selection in sparse high-dimensional vine copula models with an application to portfolio risk. *Journal of Multivariate Analysis* 172: 180–92. [CrossRef]
- Nelsen, Roger B. 1999. *An Introduction to Copulas*. Berlin/Heidelberg: Springer.
- Peters, Gareth W., Alice X. D. Dong, and Robert Kohn. 2014. A copula based bayesian approach for paid–incurred claims models for non-life insurance reserving. *Insurance: Mathematics and Economics* 59: 258–78. [CrossRef]
- Reboredo, Juan C. , and Andrea Ugolini. 2016. Systemic risk of spanish listed banks: A vine copula covar approach. *Spanish Journal of Finance and Accounting/Revista Española de Financiación y Contabilidad* 45: 1–31. [CrossRef]
- Schamberger, Benedikt, Lutz Fabian Gruber, and Claudia Czado. 2017. Bayesian inference for latent factor copulas and application to financial risk forecasting. *Econometrics* 5: 21. [CrossRef]
- Schmidt, Klaus D. 2006. Optimal and additive loss reserving for dependent lines of business. In *Proceeding of the Casualty Actuarial Society Forum*, pp. 319–51. Available online: <https://www.casact.org/pubs/forum/06fforum/323.pdf> (accessed on 10 October 2020).
- Shi, Peng, Sanjib Basu, and Glenn G. Meyers. 2012. A bayesian log-normal model for multivariate loss reserving. *North American Actuarial Journal* 16: 29–51. [CrossRef]
- Shi, Peng, and Edward W. Frees. 2011. Dependent loss reserving using copulas. *ASTIN Bulletin: The Journal of the IAA* 41: 449–86.
- Shi, Peng, and Lu Yang. 2018. Pair copula constructions for insurance experience rating. *Journal of the American Statistical Association* 113: 122–33. [CrossRef]
- Sklar, Abe. 1959. Fonctions de repartition a n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Universite de Paris* 8: 229–31.
- Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika Van Der Linde. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Methodological)* 64: 583–639. [CrossRef]
- Taylor, Greg, and Gráinne McGuire. 2016. *Stochastic Loss Reserving Using Generalized Linear Models*. Arlington County: Casualty Actuarial Society.
- Truc, Carlos íos, Aviral Kumar Tiwari, and Faisal Alqahtani. 2020. Value-at-risk and expected shortfall in cryptocurrencies' portfolio: A vine copula–based approach. *Applied Economics* 52: 2580–93.
- Zhang, Yanwei. 2010. A general multivariate chain ladder model. *Insurance: Mathematics and Economics* 46: 588–99. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).