Volume 3, Issue 2, ISSN: 2948-4316
DOI: https://doi.org/10.5281/zenodo.14928662
Published on: 31 December 2024

Journal of Social Sciences and Business
Quest International University
https://qiu.edu.my/journal/index.php/jssb/index

COMPARISON OF GENERAL INSURANCE CLAIMS RESERVING METHODS IN MALAYSIA: BASIC CHAIN LADDER METHOD VS. OVER-DISPERSED POISSON MODEL

Carmen Lim Jia En[1*], Tan Kok Kiang[2] and Rahimah Abdul Rafar[3]

[1,3]Asia Pacific University of Technology and Innovation, Kuala Lumpur, Malaysia

[2]Quest International University, Perak, Malaysia

**Corresponding author: limje.carmen@gmail.com[1]*

Abstract: The adoption of stochastic claims reserving methods has become increasingly prevalent in the General Insurance ("GI") business to quantify uncertainty. The research study will compare two methodologies: the Basic Chain Ladder ("BCL") method, which is a deterministic approach, and the Over-Dispersed Poisson ("ODP") model, which is a stochastic approach. These methodologies provide varying levels of accuracy in estimating claims reserves. Nonetheless, the absence of a thorough and comparative examination regarding implementing these methods within Malaysia's GI sector, encompassing both conventional and Takaful insurance, provide a clear understanding of which model is suitable for distinct operational contexts and objectives within the compliance framework of GI companies. With a data analysis between a general conventional insurer and a general Takaful operator in Malaysia, the study has found that the hypotheses tested in this study have significant differences. Both methods can easily approach the prediction reserves. However, the ODP model provides statistical information, for example, prediction error (also known as root mean square error, which is also discussed in this research), an advantage the BCL does not have. This research will give decision-makers the information to make educated choices in selecting claim reserves prediction methods.

Keywords: General Insurance, Claims Reserves, Basic Chain Ladder Method, Over-Dispersed Poisson Model.

INTRODUCTION

General Insurance ("GI") is also called "non-life insurance" and "property and casualty insurance". The GI industry offers a range of coverage options to individuals and businesses based on the following lines of business: motor/car insurance, property insurance, liability insurance, accident insurance, health insurance, marine insurance, etc.

Reserves, also referred to as technical provisions, constitute a fundamental component of an insurance company's liabilities on its balance sheet (Dina Manolache, 2020). Soomro (2019) outlined several reasons for estimating reserves, including covering insurance liabilities for policies already written, evaluating business division performance, calculating ultimate claim expenses, assessing the fairness of previous claims reserves, and appraising an insurer's liabilities in mergers or acquisitions. The accuracy and dependability of their estimation significantly impact the insurer's financial stability (Dina Manolache, 2020). Two main categories of reserves in insurance are claims reserves and premium reserves. This research will focus on claims reserves due to their close relation to future liabilities arising from past events. Accurate prediction of claims reserves is critical for an insurer's solvency. Moreover, claims reserves is essential for evaluating reserve sufficiency and monitoring an insurer's financial health, making it a highly relevant focus area in today's insurance industry.

According to Björkwall (2011), each GI business must establish a reserve fund to provide future payments to policyholders for claim events that have already happened. The amount is sometimes denoted as the provision for outstanding claims or the claims reserve (Björkwall, 2011). The accurate estimation of the claims reserve holds significant importance. Underestimating the reserve would impede the insurance company's ability to fulfill its liabilities, while overestimating it would result in the unnecessary retention of excess capital instead of utilizing it for alternative purposes, such as investments with higher risk and potentially higher returns (Björkwall, 2011). Based on historical data, the actuary can determine estimations, or more accurately, predictions, regarding the expected amount of outstanding claims (Björkwall, 2011).

In the past, the predominant approach for estimating claims was deterministic computations to determine the point estimate of losses (Mann, 2011). Nevertheless, in recent times, alternative approaches have been adopted that involve the application of statistical or stochastic models incorporating random variables aimed at generating an appropriate range of claims estimates (Mann, 2011). The uses of each method depend on the available information and data. Examples of deterministic methods are the Expected Loss Ratio method, BCL method, Bornhuetter-Ferguson method, Cape Cod method, etc. Examples of stochastic models are the Mack method, the ODP model, the Negative Binomial model, etc.

Carrato et al. (2016) stated that, one of the advantages of deterministic methods is their ability to generate a single outcome, i.e., the Actuarial Best Estimate. This characteristic helps comprehension and effective communication with firm management. Furthermore, it is common practice for these evaluations to incorporate external judgment and knowledge transparently, without stochastic assumptions (Carrato et al., 2016). However, there is a notable limitation of deterministic methods, that is, their limited ability to quantify the degree of uncertainty associated with the Actuarial Best Estimate (Carrato et al., 2016). Using scenario and sensitivity testing can offer valuable insights into the unpredictability of reserves (Carrato et al., 2016). However, it is important to note that achieving a comprehensive understanding of the volatility is unattainable (Carrato et al., 2016). In order to address this issue, the application of a stochastic model is necessary.

Stochastic reserving methods produce a point estimate, commonly called the Actuarial Best Estimate, for future claims, along with an estimated related variability (Hindley, 2017). Certain methodologies also offer an approximation of the full distribution of future claims and their corresponding cash flows (Hindley, 2017). Stochastic reserving methods are commonly applied in evaluating the level of uncertainty associated with a predicted reserve estimate, and the evaluation of reserving risk within the framework of capital requirement determination (Hindley, 2017).

According to the survey findings conducted by Francis (2016), the BCL method emerged as the most commonly used technique for point estimation, with the Bornhuetter-Ferguson method ranking second in terms of application. The loss ratio method is commonly applied, although average cost and Cape Cod methods are widely used. There is limited application of alternative methodologies, including generalized linear models ("GLMs"), which are grounded in statistical principles, within the actuarial profession to reserve.

The adoption of stochastic claims reserving methods has become increasingly prevalent in the GI business to quantify uncertainty (England & Verrall, 2002). This research study will compare two distinct methodologies: the BCL method, which is a deterministic approach, and the ODP model, which is a stochastic approach. The BCL method is one of the most widely used deterministic models in the insurance industry due to its simplicity and ease of application. The ODP model is a foundational stochastic method that provides a probabilistic framework for quantifying the uncertainty in claims reserve estimates. These methodologies provide varying levels of accuracy in estimating claims reserves. Nonetheless, the absence of a thorough and comparative examination regarding implementing these approaches within Malaysia's GI sector, encompassing both conventional and Takaful insurance, hinders a clear understanding of which model is most suitable for distinct operational contexts and objectives within the compliance framework of GI companies.

Despite the increasing significance of employing stochastic techniques in the GI sector (England & Verrall, 2002), a dearth of scholarly research systematically compares the BCL method and the ODP model in the GI industry, encompassing both conventional and Takaful practices. Although there is evidence indicating the potential effectiveness of each technique in specific circumstances, it is necessary to conduct a comprehensive analysis of their respective strengths and limits within the framework of claims reserving in the GI business (both conventional and Takaful). This gap hinders the decision-making processes of industry professionals who strive to adopt the most appropriate model for predicting a more precise claims reserve. Additionally, it poses challenges for organizational management in pursuing financial stability. Hence, conducting a comprehensive comparative analysis of the BCL method, the ODP model, in the GI industry (including both conventional and Takaful sectors) is imperative. This study aims to offer empirically derived insights into these methods' performance, applicability, and adaptability within the distinct context of GI claims reserving. Moreover, this study aims to provide decision-makers with the required information to make educated choices in selecting a modelling technique through the evaluation of prediction accuracy.

LITERATURE REVIEW

Claims Development Triangles

Dina Manolache (2020) stated that, two time axes have been included in the claims developments of the triangle data: the horizontal axis represents the development periods, and the vertical axis represents the accident periods. Accident year is defined as the year in which an uncertainty event happens or an uncertainty event is reported (Mann, 2011). Development year is defined as the time that has elapsed since the claim happened (Mann, 2011).

A claims triangle is a triangular dataset that consists of the claim amount. Claims triangles can consist of paid or incurred claims in incremental or cumulative form (Mann, 2011). An example of a cumulative form claims triangle is shown in TABLE I, and an incremental form claims triangle is shown in TABLE II.

TABLE I: EXAMPLE OF A CUMULATIVE FORM CLAIMS TRIANGLE

Accident Year	Development Year			
	0	1	2	3
2019	100	120	130	145
2020	105	115	125	
2021	115	130		
2022	125			

TABLE II: EXAMPLE OF AN INCREMENTAL FORM CLAIMS TRIANGLE

Accident Year	Development Year			
	0	1	2	3
2019	100	20	10	15
2020	105	10	10	
2021	115	15		
2022	125			

The data sets of the claims reserves are combined into so-called run-off triangles (as shown in TABLE III) in the chain claims reserving methods (Dina Manolache, 2020). The claims amount is in the upper part of the triangle (this is the claims triangle), and the predicted amount for claims reserves is in the lower part (Dina Manolache, 2020).

TABLE III: EXAMPLE OF AN INCREMENTAL FORM CLAIMS TRIANGLE

Accident Year	Development Year			
	0	1	2	3
2019				
2020	Claims Amount			
2021			Predicted Amount for Claims Reserves	
2022				

Types of Errors

According to Hindley (2017), from the point of view of insurance organization management or the Board of Directors, the main type of uncertainty that is of concern in the context of reserving is the variation between the current estimate of future claims outflow and the actual yet presently undisclosed future outflow. A straightforward approach to evaluating the total uncertainty involves examining the extent to which past estimates deviated from the actual claims outflow that occurred after that (Hindley, 2017). In the majority of circumstances, the estimation of the total uncertainty requires the application of a statistical model or another appropriate process to the data that is accessible (Hindley, 2017). If the model possesses inherent stochasticity, it can be fitted to the data and afterward applied to generate an estimation of at least a portion of the total uncertainty (Hindley, 2017).

There are three different components of the total uncertainty as Hindley (2017) stated:

1. Parameter error: It is also referred to as the estimation error. This denotes the level of uncertainty associated with the estimates of the parameters under the assumption that the model is accurately stated (Carrato et al., 2016). The chosen parameters during the model fitting process may be inaccurate due to the necessity for estimation, potentially resulting in differences from the underlying, unobservable "true" values.
2. Process error: It is also referred to as "Stochastic error". This phenomenon demonstrates the inherent unpredictability of future developments (Carrato et al., 2016). Despite the accurate specification of the model and precise estimation of parameters, the inherent volatility inherent in the insurance process is predicted to lead to deviations from the expected results (Carrato et al., 2016). In alternative terminology, it is frequently referred to as the variance of a random variable (Carrato et al., 2016).
3. Model error: This error refers to the difference between the assumed parametric form of a model and the actual yet unknown parametric form (Carrato et al., 2016). It represents the error that arises due to the selection of a certain model. The applied model might show inaccuracies in its representation of the claims process under analysis, hence deviating from the actual dynamics observed in the real world.

Note that, in certain instances, the terms "risk" or "variance" are used instead of the term "error" (Hindley, 2017). Therefore, while organization management may like to prioritize the total uncertainty, in practical terms, it is typically essential to assess it by splitting it into different components (Hindley, 2017), as previously explained. According to Carrato et al. (2016), the two elements of independent risk, namely parameter error and process error, refer to the risks that emerge from the intrinsic unpredictability inside the insurance process.

This research paper will focus on parameter error and process error; model error is omitted in this paper. The inclusion of model error is typically not explicitly or implicitly accounted for in the outcomes obtained by applying most stochastic reserving techniques (Hindley, 2017). Assessing model error is challenging due to the inherent uncertainty of the genuine unknown underlying distribution (Carrato et al., 2016). The estimation of the technique cannot be determined just by relying on the data used in the modelling process (Carrato et al., 2016).

When a specific statistical model is applied to a particular dataset or triangle, it generates fitted data points corresponding to the historical and known data (Hindley, 2017). The model is subsequently applied to predict future outcomes that extend beyond the scope, or "out of the sample," of the existing dataset (Hindley, 2017). This involves estimating the future expenses associated with claims (Hindley, 2017). The prediction error, which includes the parameter error and process error, contributes to the total uncertainty in the forecast (Hindley, 2017).

The Mean-Squared Error of Prediction ("MSEP") is a commonly applied measure in statistical applications for evaluating prediction variance (Hindley, 2017). Another simple term that may be applied in this situation is the Root Mean Squared Error of Prediction ("RMSEP") (Hindley, 2017). This metric is obtained by taking the square root of the MSEP, representing the prediction error.

Formulas for calculating MSEP and RMSEP extracted from England and Verrall (2002).
Consider a random variable X and a predicted value $\hat{X}$. The MSEP is:

$$E\left[\left(X-\hat{X}\right)^2\right]=E\left[\left((X-E[X])-\left(\hat{X}-E[X]\right)\right)^2\right]$$

$$E\left[\left(X-\hat{X}\right)^2\right]\approx E[(X-E[X])^2]-2E\left[(X-E[X])\left(\hat{X}-E\left[\hat{X}\right]\right)\right]+E[\left(\hat{X}-E\left[\hat{X}\right]\right)^2]$$

Assume that the future observations are independent of past observations,

$$E\left[\left(X-\hat{X}\right)^2\right]\approx E[(X-E[X])^2]+E[\left(\hat{X}-E\left[\hat{X}\right]\right)^2]$$

From the equation above can conclude that:

$$Prediction\ error\ =\ Process\ error\ +\ Estimation\ error$$

When examining variability, the primary focus lies on the RMSEP, which is alternatively referred to as the prediction error.

$$RMSEP\ =\sqrt{MSEP}$$

According to England and Verrall (2002), when evaluating the prediction error of future payments and reserve estimates by traditional statistical methods, the procedure can be simplified to calculate two key elements: the process error and the estimation error. In an alternative approach, it is possible to directly derive the RMSEP by calculating the standard deviation of the full predictive distribution (England & Verrall, 2002). The variation between prediction error and standard error, i.e., the standard error can be defined as the square root of the estimation error, while the prediction error pertains to the fluctuation in a forecast, considering both the uncertainty in estimating parameters and the inherent variability in the forecasted data (England & Verrall, 2002).

After selecting the model, the variability of the claims reserve can be determined by either analytical or simulation methods (Björkwall, 2011). The estimators for reserves are frequently complicated functions of the observations, making it challenging to develop analytical expressions. Consequently, bootstrapping gained significant popularity in obtaining the variability of the claims reserve (Björkwall, 2011). Nevertheless, the current bootstrap methods are designed based on statistical assumptions and primarily aim to measure the accuracy of the actuary's best estimate while lacking the capability to modify the estimation (Björkwall, 2011).

Basic Chain Ladder ("BCL") Method

According to Mann (2011) and Dina Manolache (2020), the BCL method is a highly recognized method of claims prediction in actuarial fields. This method establishes a fundamental basis upon which subsequent methodologies can be constructed. The method was created during a period when computers were not easily accessible, necessitating the use of straightforward closed-form equations (Dina Manolache, 2020). Historically, the traditional actuarial literature has characterized the BCL method as a computational procedure only for predicting claims reserves (Dina Manolache, 2020). The BCL method assumes that all external factors, such as inflation of claim expenses, changes in the mix of business, and variations in the rate of settlement of claims, may be effectively disregarded (Weke, 2008).

Formulas for BCL method extracted from England and Verrall (2002) and Mann (2011).

Assume to have the following set of incremental claims data:

$$\{C_{ij}: i = 1, \dots, n; j = 1, \dots, n - i + 1\}$$

The suffix i refers to the row and could indicate accident year. The suffix j refers to the column and could indicate development year. Let C_{ij} be the incremental paid or incurred claims for the i^{th} accident year and j^{th} development period.

The cumulative claims for the i^{th} accident year and j^{th} development period are defined as:

$$D_{ij} = \sum_{k=1}^{j} C_{ik}$$

The development factors of the BCL method are denoted by $\lambda_j, j = 2, \dots, n$

The BCL method estimates the development factor for the j^{th} development period is defined as:

$$\widehat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} D_{ij}}{\sum_{i=1}^{n-j+1} D_{i,j-1}}$$

These are then applied to the latest cumulative claims in each row ($D_{i,n-i+1}$) to produce prediction of future values of cumulative claims:

$$\widehat{D}_{i,n-i+2} = D_{i,n-i+1}\hat{\lambda}_{n-i+2}$$
$$\widehat{D}_{i,k} = D_{i,k-1}\hat{\lambda}_k, \qquad k = n - i + 3, n - i + 4, \dots, n$$

In its most basic iteration, the BCL method provides a method for generating predictions solely for ultimate claims (England & Verrall, 2002). In this context, the term "ultimate" refers to the most recent development year observed, excluding any further considerations (England & Verrall, 2002). From a statistical perspective, after a point estimate is obtained, the subsequent logical progression involves generating estimates of the probable variability in the outcome (England & Verrall, 2002). This allows for assessments, such as determining whether additional reserves should be maintained in addition to the projected values for cautionary purposes (England & Verrall, 2002). The measure of variability usually applied in this context is the prediction error, defined as the standard deviation of the distribution of expected reserve outcomes (England & Verrall, 2002). Additionally, considering other variables, such as the likelihood of unforeseen occurrences that could amplify uncertainty, is advantageous yet challenging to represent accurately through modelling (England & Verrall, 2002).

According to England and Verrall (2002), the initial stage in acquiring the prediction error involves the development of an underlying statistical model, wherein assumptions are made on the data. If the objective is to establish a stochastic model that mirrors the BCL method, the projected values must align with those generated by the BCL method (England & Verrall, 2002). Two options have been used to address this issue: one involves providing distributions for the data, while the other involves only specifying the first two moments (England & Verrall, 2002). Consequently, a stochastic model is required.

Over-Dispersed Poisson ("ODP") Model

The ODP model is a statistical model applied in GI claims reserving. According to Hindley (2017), it is specifically designed to be fitted to the incremental claims considered within a data triangle. The Poisson distribution possesses certain constraints, including that it exclusively admits non- negative integer values and shows a variance consistently equivalent to the mean (Carrato et al., 2016). Nevertheless, the ODP model generalizes the Poisson model, addressing certain drawbacks while maintaining the fundamental structure and advantageous characteristic of producing reserve estimations that align with those derived from the BCL method (Carrato et al., 2016; Hindley, 2017).

The ODP distribution is defined by a variance that can deviate from the mean, and it accepts non-integer values (Hindley, 2017). The requirement for non-negative values is less stringent because it only applies to the column totals within the triangle of incremental claim amounts (Hindley, 2017). However, according to Hindley (2017), it is important to acknowledge that this condition implies that the expected value of the incremental claim amounts cannot be negative, which may be undesirable in specific scenarios. The fundamental framework of the ODP model closely resembles that of the original Poisson model, wherein the average of the incremental claims amounts is proposed as the outcome of multiplying an origin period parameter by a development period parameter (Hindley, 2017). In order to account for the potential difference between the variance and the mean, it is necessary to assume the variance (Hindley, 2017).

The ODP model presumes the existence of a scale parameter, which establishes a relationship between the variance of the incremental claims amounts and the product of the mean and the scale parameter (Hindley, 2017). The scale parameter can be either constant or permitted to differ based on the development time (Hindley, 2017). The outcomes generated by this method are comparable to those obtained using the deterministic BCL method, where the sum of the columns weighs the average (Hindley, 2017). The provided estimation contains the average reserve value and the corresponding variation, although it does not include the full distribution (Hindley, 2017).

Formulas for ODP model extracted from England and Verrall (2002) and Mann (2011).

In claims reserving, the ODP model assumes that the incremental claims C_{ij} are distributed as independent over-dispersed Poisson random variable. The model is parameterized as follows:

$$\text{Mean: } E\left[C_{ij}\right\} = m_{ij}$$

$$\text{Variance: } Var\left[C_{ij}\right\} = \phi m_{ij}$$

The first framework is as follows:

$$m_{ij} = x_i y_j$$

The first framework of the model shows the non-linear in its parameters, necessitating the use of non-linear modeling approaches to provide parameter estimates (England & Verrall, 2002). Obtaining maximum likelihood estimates requires developing the likelihood function and optimizing it concerning the parameters, a task that is only sometimes straightforward (England & Verrall, 2002). Nevertheless, to estimate, it is frequently advantageous to re-parameterize the model so that the mean has a linear form (Ogutu, 2011). The log link function is applied in the context of GLMs (Ogutu, 2011).

The second framework is as follows:

$$\log\left(m_{ij}\right) = c + \alpha_i + \beta_j$$

According to England and Verrall (2002), the second framework presents a GLM, wherein the results C_{ij} are represented as Poisson random variables with a logarithmic link function and a linear predictor $\eta_{ij} = c + \alpha_i + \beta_j$. The consideration of over-dispersion is incorporated into the fitting approach by estimating the scale parameter ϕ, which is unknown (England & Verrall, 2002). The application of the log link function keeps the model linear to the parameters (England & Verrall, 2002). Parameter estimates can be obtained using maximum likelihood using commonly available statistical software capable of fitting GLMs (England & Verrall, 2002). Nevertheless, applying the log link function presents a constraint regarding the ease of interpreting parameter estimates (England & Verrall, 2002). The parameters can be connected to the chain-ladder link ratios or the non-linear model y_j parameters, as demonstrated by Verrall (1991). However, adding additional calculations can be perceived as a drawback when applying the ODP model (England & Verrall, 2002). Additionally, it should be noted that fitted values will consistently show a positive, leading to equivalent link ratios that are always greater than 1 (England & Verrall, 2002).

The ODP was first introduced in the publication of (Renshaw & Verrall, 1998). They introduced the model as a GLM. The GLM framework is widely recognized and renowned for its flexibility in statistical modelling (McCullagh & Nelder, 1989). The application of GLMs offers numerous associated alternatives and extensions that can be explored if the fundamental ODP model becomes insufficient (Hindley, 2017). One potential approach to incorporate calendar period trends, such as inflation, into the linear predictor is introducing calendar period parameters (Hindley, 2017). This would enable explicit modelling of these trends (Hindley, 2017). It was, alternatively, assuming a different distribution, such as the gamma distribution from the Tweedie family, to account for variations in the distribution of incremental claims (Hindley, 2017).

A widely accepted practice in academic literature is applying an ODP distribution in conjunction with a logarithmic link function when dealing with incremental data (Björkwall, 2011). According to Björkwall (2011), one implication of this specific assumption is that the expected claims estimates derived through maximum likelihood estimation of the parameters in the GLM are equivalent to those produced by the BCL method, provided that the column sums of the triangle are positive (Renshaw & Verrall, 1998). Hence, the estimated expected values can be derived by either maximum likelihood estimation or the BCL method (Björkwall, 2011). Additionally, the estimated variances acquired from the GLM assumption can be utilized to calculate or simulate a measure of variability (Björkwall, 2011).

The ODP distribution can be distinguished from the Poisson distribution because the variance is not equivalent to the mean (England & Verrall, 2002). However, rather, it is directly proportional to the mean (England & Verrall, 2002). The Poisson distribution upon which the model in this section is built does not imply that it is limited to positive integer data (England & Verrall, 2002). A strategy known as "quasi-likelihood" can be used to get around this restriction and can be applied to both positive and negative non-integer data (McCullagh & Nelder, 1989). Up to a constant of proportionality, the likelihood with quasi-likelihood in this situation is equivalent to a Poisson likelihood (England & Verrall, 2002). The complete or quasi-likelihood estimates the same parameters with data that only contains positive integers (England & Verrall, 2002). The fundamental presumption in modeling is that the variance is inversely proportional to the mean and that the data are not constrained to positive integers (England & Verrall, 2002).

It is important to acknowledge that the model has limitations, as highlighted by Mann (2011) and Renshaw and Verrall (1998):

1. The over-parameterization of the model, resulting in $\alpha_1 = \beta_1 = 0$.
2. Claims must be represented as integer values. Typically, attaining this objective is easy as financial statements commonly round claims to the nearest dollar.
3. Each $C_{ij} \geq 0$. According to this limitation, it is required that claims in each period have positive values. However, as previously discussed, the concepts of salvage and subrogation allow for the occurrence of negative claims within a given time frame. Nevertheless, the utilization of quasi-likelihood maximization presents a viable approach for estimating negative and non-integer claims, effectively addressing the second and third limitations mentioned earlier. According to (Renshaw & Verrall, 1998), when there are negative incremental losses, it is advisable to employ quasi-likelihood estimation. Additionally, for assessing the goodness of fit in modelling, the Pearson χ^2 statistic is recommended as a suitable alternative to deviances.
4. Requires that $\sum_i C_{ij} \geq 0$ (Hindley, 2017). While the ODP model remains applicable when there are just a few negative incremental values in the triangle, however, $\sum_i C_{ij} \geq 0$ refers to the sum of each column and row in the incremental triangle, which must not be negative. This limitation must not be infringed upon for the model to function properly.

Steps in order to acquire the prediction error extracted from England and Verrall (2002) as below shown.

Future payment estimates can be derived by substituting the parameter estimations into the second framework and using the exponential function: $\hat{C}_{ij} = \hat{m}_{ij} = \exp{(\hat{\eta}_{ij})}$

The summing process can determine the year of origin and the total estimated reserves. In addition, the inclusion of prediction errors is necessary for the analysis. Below work will introduce the examination by focusing on one single incremental payment, C_{ij} .

Consider origin year i and claim payments in development year j. The MSEP is as below:

$$[1]: MSEP[\hat{C}_{ij}] = E[(C_{ij} - \hat{C}_{ij})^2 \approx Var[C_{ij}] + Var[\hat{C}_{ij}]$$

$$[2]: Var[C_{ij}] = \phi m_{ij}$$

$$[3]: Var[\hat{C}_{ij}] \approx \left|\frac{\partial m_{ij}}{\partial \eta_{ij}}\right|^2 Var[\hat{\eta}_{ij}]$$

Substitute [2] and [3] into [1], hence: $MSEP[\hat{C}_{ij}] \approx \phi \hat{m}_{ij} + \hat{m}_{ij}^2 Var[\hat{\eta}_{ij}]$

DATA ANALYSIS

The main aim of this paper is to present claims reserving prediction and evaluate the accuracy of claims reserving prediction for selected Malaysia general insurers (both conventional and Takaful), i.e., Allianz General Insurance Company (Malaysia) Berhad ("Allianz") and Etiqa General Takaful Berhad ("Etiqa"), based on BCL method and ODP model on their claims paid data.

Based on the list of insurers licensed by Bank Negara Malaysia, there are 19 insurers for the GI business and 4 operators for the General Takaful business. As of 2022, Allianz had the highest total assets (RM 7,276,189,000) compared to other GI businesses, and Etiqa had the highest total assets (RM 4,913,802,000) compared to other General Takaful businesses.

The dataset used in this research was secondary data of net claims paid triangles (cumulative form) taken from the 2022 Financial Statements of both companies. England and Verrall (2002) state that using paid claims rather than incurred claims is better since negative values are less likely to appear in the paid claims. This is because the appearance of these negative incremental values in the data may cause problems when applying some claims reserving methods.

Data Preparation

To illustrate the methodology, consider the net claims paid triangle data in TABLE IV and TABLE V, shown in cumulative form. The data is analysed using Microsoft Excel and R Studio.

TABLE IV: ALLIANZ NET CLAIMS PAID IN CUMULATIVE FORM

Accident Year	Development Year						
	0	1	2	3	4	5	6
2016	468300	817863	925817	972070	988580	1001058	1007421
2017	518300	896008	998910	1034851	1053901	1071913	
2018	507250	888891	983920	1030774	1066274		
2019	496380	839564	939217	1017622			
2020	416786	736854	858851				
2021	416655	843445					
2022	649947						

Note. Extracted from Allianz General Insurance Company (Malaysia) Berhad Financial statements for the year ended 31 December 2022 (p 133).

TABLE V: ETIQA NET CLAIMS PAID IN CUMULATIVE FORM

Accident Year	Development Year						
	0	1	2	3	4	5	6
2016	307415	521478	599827	628770	637303	648086	652966
2017	340963	559277	632054	657164	667774	677032	
2018	340369	538219	596401	623065	638128		
2019	450253	681490	750921	805694			
2020	395860	602965	614597				
2021	271665	532664					
2022	491813						

Note. Extracted from Etiqa General Takaful Berhad Directors' Report and Audited Financial Statements 31 December 2022 (p 157).

Two Sample Independent T-Test

TABLE VI and TABLE VII show the result of obtaining two-sample independent t-tests, both assumed variances equal and unequal, on net claims paid by Allianz and Etiqa (using Microsoft Excel).

TABLE VI: T-TEST, TWO SAMPLE ASSUMING EQUAL VARIANCES

	Allianz	Etiqa
Mean	837407.9286	566579.0357
Variance	47964675300	17952884772
Observations	28	28
Pooled Variance	32958780036	
Hypothesized Mean Difference	0	
df	54	
t Stat	5.581789421	
P(T<=t) one-tail	3.95588E-07	
t Critical one-tail	1.673564906	
P(T<=t) two-tail	7.91176E-07	
t Critical two-tail	2.004879288	

TABLE VII: T-TEST, TWO SAMPLE ASSUMING UNEQUAL VARIANCES

	Allianz	Etiqa
Mean	837407.9286	566579.0357
Variance	47964675300	17952884772
Observations	28	28
Hypothesized Mean Difference	0	
df	45	
t Stat	5.581789421	
P(T<=t) one-tail	6.49198E-07	
t Critical one-tail	1.679427393	
P(T<=t) two-tail	1.2984E-06	
t Critical two-tail	2.014103389	

H0: There is no significant difference between the means of the net claims paid of general conventional insurer and general Takaful operator.
H1: There is a significant difference between the means of the net claims paid of general conventional insurer and general Takaful operator.

Difference between the means of the net claims paid of Allianz and Etiqa is
837407.9286 – 566579.0357 = 270828.8929

Significance level = 0.05
p-value (assuming equal variances) = 7.91176E–07
p-value (assuming unequal variances) = 1.2984E–06

Since p-values for both assuming equal variances and unequal variances are lower than the significance level of 0.05. Hence, reject the null hypothesis and conclude that there is a significant difference (270828.8929) between the means of general conventional insurer and general Takaful operator.

Parameters Estimation

TABLE VIII and TABLE IX show the development factors constructed by applying the BCL method (using Microsoft Excel). The BCL method is constructed by using historical cumulative data based on development factors. Development factors are factors that take the development of the claims from one development year to the next. Development year is the time taken for the claim to develop from its accident year.

TABLE VIII: DEVELOPMENT FACTORS OF ALLIANZ

Years of Development	0 to 1	1 to 2	2 to 3	3 to 4	4 to 5	5 to 6
Development Factors	1.7788	1.1262	1.0539	1.0234	1.0149	1.0064

TABLE IX: DEVELOPMENT FACTORS OF ETIQA

Years of Development	0 to 1	1 to 2	2 to 3	3 to 4	4 to 5	5 to 6
Development Factors	1.6312	1.1000	1.0525	1.0179	1.0154	1.0075

The claims reserve estimation using the ODP model is formulated with a single scale parameter and with separate parameters for each accident year and development year. TABLE X and TABLE XI show the estimated coefficients of each parameter (i.e., all development years and accident years) constructed by fitting the ODP model (using R Studio).

TABLE X: ESTIMATE COEFFICIENTS OF ALLIANZ

	Estimate
(intercept)	13.0313
Accident Year 2017	0.0684
Accident Year 2018	0.0779
Accident Year 2019	0.0544
Accident Year 2020	-0.0628
Accident Year 2021	0.038
Accident Year 2022	0.3533
Development Year 1	-0.2501
Development Year 2	-1.4937
Development Year 3	-2.2256
Development Year 4	-3.008
Development Year 5	-3.4341
Development Year 6	-4.2731

TABLE XI: ESTIMATE COEFFICIENTS OF ETIQA

	Estimate
(intercept)	12.712967
Accident Year 2017	0.043695
Accident Year 2018	-0.000245
Accident Year 2019	0.25068
Accident Year 2020	0.031141
Accident Year 2021	-0.016616
Accident Year 2022	0.392887
Development Year 1	-0.460186
Development Year 2	-1.813193
Development Year 3	-2.361724
Development Year 4	-3.38612
Development Year 5	-3.522665
Development Year 6	-4.220066

Prediction of Claims Reserves

A cumulative form data triangle is needed to predict the claims reserves by applying the BCL method. TABLE XII and TABLE XIII show the prediction of claims reserves construct by applying the BCL method (using Microsoft Excel).

TABLE XII: ALLIANZ PREDICTION OF CLAIMS RESERVES IN CUMULATIVE FORM, CONSTRUCT BY APPLYING BCL METHOD

Accident Year	Development Year						
	0	1	2	3	4	5	6
2016	468300	817863	925817	972070	988580	1001058	1007421
2017	518300	896008	998910	1034851	1053901	1071913	1078726.3739
2018	507250	888891	983920	1030774	1066274	1082191.2566	1089069.9619
2019	496380	839564	939217	1017622	1041426.9637	1056973.3057	1063691.7188
2020	416786	736854	858851	905154.9277	926328.9788	940157.1469	946133.0443
2021	416655	843445	949912.4788	1001125.8776	1024544.9519	1039839.2805	1046448.7850
2022	649947	1156097.8779	1302031.3131	1372228.7790	1404328.9659	1425292.6812	1434352.2335

TABLE XIII: ETIQA PREDICTION OF CLAIMS RESERVES IN CUMULATIVE FORM, CONSTRUCT BY APPLYING BCL METHOD

Accident Year	Development Year						
	0	1	2	3	4	5	6
2016	307415	521478	599827	628770	637303	648086	652966
2017	340963	559277	632054	657164	667774	677032	682129.9595
2018	340369	538219	596401	623065	638128	647927.2097	652806.0141
2019	450253	681490	750921	805694	820130.6597	832724.7354	838995.0402
2020	395860	602965	614597	646882.8447	658473.8799	668585.5247	673619.8833
2021	271665	532664	585935.5552	616715.7646	627766.2573	637406.3502	642205.9339
2022	491813	802228.8872	882459.5401	928816.6679	945459.4754	959978.1232	967206.6288

An incremental form data triangle is needed to predict the claims reserves by applying the ODP model. TABLE XIV and TABLE XV show the prediction of claims reserves construct by fitting the ODP model (using Microsoft Excel).

TABLE XIV: ALLIANZ PREDICTION OF CLAIMS RESERVES IN INCREMENTAL FORM, CONSTRUCT BY FITTING ODP MODEL

Accident Year	Development Year						
	0	1	2	3	4	5	6
2016	468300	349563	107954	46253	16510	12478	6363
2017	518300	377708	102902	35941	19050	18012	6813.0829
2018	507250	381641	95029	46854	35500	15916.3156	6878.1156
2019	496380	343184	99653	78405	23806.1683	15546.6428	6718.3643
2020	416786	320068	121997	46300.1278	21173.3798	13827.2975	5975.3622
2021	416655	426790	106468.0267	51210.5068	23418.9312	15293.7572	6609.0817
2022	649947	506105.2851	145932.5781	70192.7284	32099.6370	20962.7011	9058.8730

TABLE XV: ETIQA PREDICTION OF CLAIMS RESERVES IN INCREMENTAL FORM, CONSTRUCT BY FITTING ODP MODEL

Accident Year	Development Year						
	0	1	2	3	4	5	6
2016	307415	214063	78349	28943	8533	10783	4880
2017	340963	218314	72777	25110	10610	9258	5097.9613
2018	340369	197850	58182	26664	15063	9799.2095	4878.8070
2019	450253	231237	69431	54773	14436.6732	12594.0782	6270.3095
2020	395860	207105	11632	32285.8393	11591.0439	10111.6449	5034.3616
2021	271665	260999	53271.5662	30780.2029	11050.5005	9640.0927	4799.5863
2022	491813	310415.8099	80230.6588	46357.1120	16642.8172	14518.6455	7228.5085

No claims reserves prediction can be made for the first accident year (i.e., the accident year 2016) because it is impossible to predict the highest development year.

Model Comparison

TABLE XVI shows the comparison of the predicted claims reserves for each accident year by applying the BCL method and the ODP model for Allianz.

TABLE XVI: THE COMPARISON OF PREDICTED CLAIMS RESERVES OF BCL METHOD AND ODP MODEL FOR ALLIANZ

Accident Year	Claims Reserves	
	BCL Method	ODP Model
2017	6813.3739	6813.0829
2018	22795.9619	22794.4312
2019	46069.7188	46071.1755
2020	87282.0443	87276.1673
2021	203003.7850	203000.3036
2022	784405.2335	784351.8027
Total	1150370.1174	1150306.9632

H0: There is no significant difference in predicting claims reserves of general conventional insurer between the BCL method and the ODP model.

H1: There is a significant difference in predicting claims reserves of general conventional insurer between the BCL method and the ODP model.

Difference of total predicted claims reserves of Allianz between BCL method and ODP model is
11503370.1174 – 1150306.9632 = 63.1542

Hence, reject the null hypothesis and conclude that, there is a significant difference (63.1542) in predicting general conventional insurer claims reserves between the BCL method and the ODP model.

TABLE XVII shows the comparison of the predicted claims reserves for each accident year by applying the BCL method and the ODP model for Etiqa.

TABLE XVII: THE COMPARISON OF PREDICTED CLAIMS RESERVES OF BCL METHOD AND ODP MODEL FOR ETIQA

Accident Year	Claims Reserves	
	BCL Method	ODP Model
2017	5097.9595	5097.9613
2018	14678.0141	14678.0165
2019	33301.0402	33301.0610
2020	59022.8833	59022.8896
2021	109541.9339	109541.9485
2022	475393.6288	475393.5518
Total	697035.4597	697035.4288

H0: There is no significant difference in predicting claims reserves of general Takaful operator between the BCL method and the ODP model.
H1: There is a significant difference in predicting claims reserves of general Takaful operator between the BCL method and the ODP model.

Difference of total predicted claims reserves of Etiqa between BCL method and ODP model is
697035.4597 – 697035.4288 = 0.0308

Hence, reject the null hypothesis and conclude that, there is a significant difference (0.0308) in predicting general Takaful operator claims reserves between the BCL method and the ODP model.

Root of Mean Square Error of Prediction ("RMSEP")

TABLE XVIII shows the total RMSEP of applying the ODP model in predicting claims reserves for Allianz and Etiqa (using R Studio).

TABLE XVIII: TOTAL RMSEP OF THE ODP MODEL IN PREDICTING CLAIMS RESERVES FOR ALLIANZ AND ETIQA

Companies	Total RMSEP
Allianz	118770.5
Etiqa	117906.7

H0: There is no significant difference in RMSEP of the ODP model in predicting claims reserves for general conventional insurer and general Takaful operator.
H1: There is a significant difference in RMSEP of the ODP model in predicting claims reserves for general conventional insurer and general Takaful operator.

Difference of RMSEP = 118770.5 – 117906.7 = 863.8

Hence, reject the null hypothesis and conclude that, there is a significant difference (863.8) in RMSEP of the ODP model in predicting claims reserves for general conventional insurer and general Takaful operator.

CONCLUSION

Summary of Findings and Discussion

A two-sample independent t-test (for both assumed equal variances and unequal variances) was conducted to determine if its mean differs across two independent groups. As a result, there is a significant difference (270828.8929) between the means of the general conventional insurer and the general Takaful operator.

TABLE XVI shows the comparison of predicted claims reserves of the BCL method and the ODP model for Allianz. The total predicted claims reserves using the BCL method (11503370.1174) is higher than the ODP model (1150306.9632). As a result, there is a significant difference (63.1542) in predicting general conventional insurer claims reserves between the BCL method and the ODP model. TABLE XVII shows the comparison of predicted claims reserves of the BCL method and the ODP model for Etiqa. The total predicted claims reserves using the BCL method (697035.4597) is higher than the ODP model (697035.4288). As a result, there is a significant difference (0.0308) in predicting general conventional insurer claims reserves between the BCL method and the ODP model. The difference in using both claims reserves predicting methods in predicting Allianz's claims reserves is large. However, in predicting Etiqa's claims reserves is small.

In TABLE XVIII, the ODP model suggests that Allianz's claims reserves are with 118770.5 of RMSEP, and Etiqa's claims reserves are with 117906.7 of RMSEP. As a result, there is a significant difference (863.8) in RMSEP of the ODP model in predicting claims reserves for general conventional insurer and general Takaful operator. It is known that the lower the RMSEP value, the better the predictive accuracy. As a result, Etiqa has the lowest RMSEP value, and we conclude that the ODP model performs better for the general Takaful operator.

According to the survey findings conducted by ROC/GIRO Working Party (2007), the ODP model is a primary method used by practitioners to evaluate reserving uncertainty. Therefore, it is recommended to conduct the ODP model for predicting claims reserves. ODP model provides statistical tests on the prediction (i.e., it gives the prediction error); this advantage makes it preferable to the BCL method.

Conclusion

This study mainly focused on determining and comparing the claims reserves prediction by applying the BCL method and the ODP model in GI (conventional and Takaful) claims reserves in Malaysia. Claims reserve prediction is crucial as payment delays often occur in insurance companies, and this is due to delays in reporting, claims processing delays, and legal proceedings. Hence, predicting claims reserves is important to settle known and unknown future claims. With a data analysis between a general conventional insurer and a general Takaful operator in Malaysia, the study has found that the hypotheses tested in this study have significant differences. Therefore, all four hypotheses in this research have been rejected based on the analysis and findings. Given the relatively recent development of this field of study and the lack of existing research, it might take a lot of work to determine the exact explanations of the test results.

Limitation of Research

Due to the limited access to data, this study only analysed the selected companies' overall paid claims but was not split into lines of business. Each line of business has unique characteristics, short-tail business and long-tail business, which may only have specific claims reserves prediction methods that can apply. Short-tail business refers to lines for which claims are typically reported shortly, with the most extended claim delays at most around five years from the loss event (Hardy, 2022). Long-tail business refers to lines that take well over five years to settle all the claims arising in a year (Hardy, 2022). Besides that, the limitation of this study is that the data analysis of ODP does not consider assumptions, for example, inflation, interests, and expenses. As a result, this may limit the accuracy of the findings.

Suggestions for Further or Additional Research

In addition to the above conclusions, it is recommended that further study is done to check how the claims reserves prediction methods are conducted when extending them to include inflation, interests, and expenses. As one of the advantages of stochastic methods is their ability to generate substantial statistical information to get more accurate results, it is recommended to conduct stochastic methods other than the ODP model for future development. Lastly, it is recommended that further study is done to check on the prediction of claims reserves for each line of business in GI companies. This is because different lines of business have different delays in claims settlement. It should consider the characteristics of each line of business when conducting different predicting claims reserves methods.

APPENDICES

R Code Working

1.0 Install Packages

```
library(ChainLadder)
library(dplyr)
```

2.0 Data Preparation

2.1 Import Data

Cumulative form net paid-claims triangle data from:

General Conventional Insurer: Allianz General Insurance Company (Malaysia) Berhad ("AGIC")

General Takaful Operator: Etiqa General Takaful Berhad ("EGTB")

Source: Companies' Year 2022 Financial Statement

```
agic_cp = read.csv("/cloud/project/Allianz_ncp_2022.csv", header = FALSE)
egtb_cp = read.csv("/cloud/project/Etiqa_ncp_2022.csv", header = FALSE)
```

2.2 Data Manipulation

AGIC

```
#Convert original dataset into triangular format
agic_cp_t = as.triangle(as.matrix(agic_cp))

# Convert dataset into matrix form and incremental form
agic_cp_inc = as.matrix(cum2incr(agic_cp_t))

agic_claims = as.vector(agic_cp_inc)
agic_n.origin = nrow(agic_cp_inc)
agic_n.dev = ncol(agic_cp_inc)
agic_origin = factor(rep(1:agic_n.origin, agic_n.dev))
agic_dev = factor(rep(1:agic_n.dev, each = agic_n.origin))

# Convert dataset into data frame format
agic_dataframe = data.frame(claims = agic_claims, origin = agic_origin,
                            dev = agic_dev)
```

EGTB

```
#Convert original dataset into triangular format
egtb_cp_t = as.triangle(as.matrix(egtb_cp))

# Convert dataset into matrix form and incremental form
egtb_cp_inc = as.matrix(cum2incr(egtb_cp_t))

egtb_claims = as.vector(egtb_cp_inc)
egtb_n.origin = nrow(egtb_cp_inc)
egtb_n.dev = ncol(egtb_cp_inc)
egtb_origin = factor(rep(1:egtb_n.origin, egtb_n.dev))
egtb_dev = factor(rep(1:egtb_n.dev, each = egtb_n.origin))
```

```r
# Convert dataset into data frame format
egtb_dataframe = data.frame(claims = egtb_claims, origin = egtb_origin,
                            dev = egtb_dev)
```

3.0 Data Analysis

3.1 Over-Dispersed Poisson (ODP) Model

3.1.1 Modelling: New Quasi-Poisson Family

```r
quasipoisson = function(link = "log") {
  linktemp = substitute(link)
  if (!is.character(linktemp)) {
    linktemp = deparse(linktemp)
    if (linktemp == "link")
      linktemp = eval(link)
}

  if (any(linktemp == c("log", "identity", "sqrt")))
    stats = make.link(linktemp)
  else stop(paste(linktemp, "link not available for poisson",
    "family; available links are", "\"identity\",
    \"log\" and \"sqrt\""))
  variance = function(mu) mu
  validmu = function(mu) all(mu > 0)
  dev.resids = function(y, mu, wt) wt * (y - mu)^2 / mu
  aic = function(y, n, mu, wt, dev) NA
  initialize = expression({
    n = rep(1, nobs)
    mustart = rep(mean(y), length(y))
})

  structure(list(family = "quasipoisson", link = linktemp,
                 linkfun = stats$linkfun, linkinv = stats$linkinv,
                 variance = variance, dev.resids = dev.resids,
                 aic = aic, mu.eta = stats$mu.eta,
                 initialize = initialize, validmu = validmu,
                 valideta = stats$valideta),
            class = "family")
}
```

3.1.2 Fit Model

AGIC

```r
agic_model = glm(agic_claims ~ agic_origin + agic_dev, family = quasipoisson(),
                 subset = !is.na(claims), data = agic_dataframe)
summary(agic_model)
```

```
##
## Call:
## glm(formula = agic_claims ~ agic_origin + agic_dev, family = quasipoisson(),
```

```
##     data = agic_dataframe, subset = !is.na(claims))
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  13.03133    0.07027 185.443  < 2e-16 ***
## agic_origin2  0.06839    0.08883   0.770 0.453355
## agic_origin3  0.07793    0.08911   0.875 0.395627
## agic_origin4  0.05435    0.09031   0.602 0.556250
## agic_origin5 -0.06277    0.09461  -0.663 0.517141
## agic_origin6  0.03801    0.09564   0.397 0.696674
## agic_origin7  0.35332    0.10594   3.335 0.004521 **
## agic_dev2    -0.25006    0.05748  -4.350 0.000571 ***
## agic_dev3    -1.49374    0.09672 -15.444 1.28e-10 ***
## agic_dev4    -2.22558    0.14661 -15.180 1.64e-10 ***
## agic_dev5    -3.00803    0.24419 -12.318 3.02e-09 ***
## agic_dev6    -3.43410    0.36987  -9.285 1.31e-07 ***
## agic_dev7    -4.27307    0.80432  -5.313 8.69e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 4084.986)
##
##     Null deviance: 4849304  on 27  degrees of freedom
## Residual deviance:   61275  on 15  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 8
```

EGTB

```
egtb_model = glm(egtb_claims ~ egtb_origin + egtb_dev, family = quasipoisson(),
                 subset = !is.na(claims), data = egtb_dataframe)
summary(egtb_model)
```

```
##
## Call:
## glm(formula = egtb_claims ~ egtb_origin + egtb_dev, family = quasipoisson(),
##     data = egtb_dataframe, subset = !is.na(claims))
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  12.712967   0.111126 114.401  < 2e-16 ***
## egtb_origin2  0.043695   0.144015   0.303 0.765747
## egtb_origin3 -0.000245   0.146418  -0.002 0.998687
## egtb_origin4  0.250680   0.138792   1.806 0.090997 .
## egtb_origin5  0.031141   0.148480   0.210 0.836700
## egtb_origin6 -0.016616   0.154856  -0.107 0.915973
## egtb_origin7  0.392887   0.162212   2.422 0.028565 *
## egtb_dev2    -0.460186   0.091790  -5.013 0.000154 ***
## egtb_dev3    -1.813193   0.165159 -10.978 1.44e-08 ***
## egtb_dev4    -2.361724   0.234149 -10.086 4.46e-08 ***
## egtb_dev5    -3.386120   0.453846  -7.461 2.01e-06 ***
## egtb_dev6    -3.522665   0.591277  -5.958 2.63e-05 ***
## egtb_dev7    -4.220066   1.191258  -3.543 0.002954 **
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 6867.53)
##
##     Null deviance: 3990678  on 27  degrees of freedom
## Residual deviance:  103013  on 15  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 7
```

3.2 Determine uncertainty statistics - root mean squared error of prediction (RMSEP)

AGIC

```
# Get coefficients
agic_coef = agic_model$coefficients

# Get dispersion parameter
agic_disp = summary(agic_model)$dispersion

# Get covariance matrix of parameters
agic_cov.param = agic_disp * summary(agic_model)$cov.unscaled

agic_n.fut.points = length(agic_claims[is.na(agic_claims)])
agic_fut.design = matrix(0, nrow = agic_n.fut.points, ncol = length(agic_coef))
agic_fut.points = agic_claims
agic_fut.points[!is.na(agic_claims)] = 0
agic_fut.points[is.na(agic_claims)] = 1:agic_n.fut.points

for (p in 1:agic_n.fut.points){
  # All points and a constant in the predictor
  agic_fut.design[p, 1] = 1
  # Row factor
  agic_fut.design[p, 1 +
    as.numeric(agic_origin[match(p, agic_fut.points)]) - 1] = 1
  # Column factor
  agic_fut.design[p, 1 + (agic_n.origin - 1) +
    as.numeric(agic_dev[match(p, agic_fut.points)]) - 1] = 1
}

agic_fitted.values = diag(as.vector(exp(agic_fut.design %*% agic_coef)))
agic_total.reserve = sum(agic_fitted.values)

# Determine covariance matrix of linear predictors
agic_cov.pred = agic_fut.design %*% agic_cov.param %*% t(agic_fut.design)

# Determine covariance matrix of fitted values
agic_cov.fitted = agic_fitted.values %*% agic_cov.pred %*% agic_fitted.values

# Determine uncertainty statistics
agic_total.rmsep = sqrt(agic_disp * agic_total.reserve + sum(agic_cov.fitted))
```

```
agic_total.rmsep
```

```
## [1] 118770.5
```

EGTB

```
# Get coefficients
egtb_coef = egtb_model$coefficients

# Get dispersion parameter
egtb_disp = summary(egtb_model)$dispersion

# Get covariance matrix of parameters
egtb_cov.param = egtb_disp * summary(egtb_model)$cov.unscaled

egtb_n.fut.points = length(egtb_claims[is.na(egtb_claims)])
egtb_fut.design = matrix(0, nrow = egtb_n.fut.points, ncol = length(egtb_coef))
egtb_fut.points = egtb_claims
egtb_fut.points[!is.na(egtb_claims)] = 0
egtb_fut.points[is.na(egtb_claims)] = 1:egtb_n.fut.points

for (p in 1:egtb_n.fut.points){
  # All points and a constant in the predictor
  egtb_fut.design[p, 1] = 1
  # Row factor
  egtb_fut.design[p, 1 +
    as.numeric(egtb_origin[match(p, egtb_fut.points)]) - 1] = 1
  # Column factor
  egtb_fut.design[p, 1 + (egtb_n.origin - 1) +
    as.numeric(egtb_dev[match(p, egtb_fut.points)]) - 1] = 1
}

egtb_fitted.values = diag(as.vector(exp(egtb_fut.design %*% egtb_coef)))
egtb_total.reserve = sum(egtb_fitted.values)

# Determine covariance matrix of linear predictors
egtb_cov.pred = egtb_fut.design %*% egtb_cov.param %*% t(egtb_fut.design)

# Determine covariance matrix of fitted values
egtb_cov.fitted = egtb_fitted.values %*% egtb_cov.pred %*% egtb_fitted.values

# Determine uncertainty statistics
egtb_total.rmsep = sqrt(egtb_disp * egtb_total.reserve + sum(egtb_cov.fitted))
egtb_total.rmsep
```

```
## [1] 117906.7
```

REFERENCES

Björkwall, S. (2010). *Stochastic claims reserving in non-life insurance*. http://su.diva-portal.org/smash/get/diva2:406884/FULLTEXT01.pdf

Carrato, A., McGuire, G., & Scarth, R. (2016). *A practitioner's introduction to stochastic reserving*. https://www.actuaries.org.uk/documents/practitioners-introduction-stochastic-reserving

Dina Manolache, A. E. (2020). Chain claims reserving methods in non-life insurance. *Proceedings of the International Conference on Applied Statistics*, *1*(1), 216–225. https://doi.org/10.2478/icas-2019-0019

England, P. D., & Verrall, R. J. (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal*, *8*(3), 443–518. https://doi.org/10.1017/s1357321700003809

Francis, L. (2016). ASTIN working party releases report on reserving practices for general insurance worldwide. In *Actuarial Review Magazine CAS*. Casualty Actuarial Society (CAS). https://ar.casact.org/astin-working-party-releases-report-on-reserving-practices-for-general-insurance-worldwide/

Hardy, M. R. (2022). *Advanced short-term actuarial mathematics study note*. Education Committee of the Society of Actuaries. https://www.soa.org/4a18d7/globalassets/assets/files/edu/2023/2023-exam-astam-21-23.pdf

Hindley, D. (2017). *Claims reserving in general insurance*. https://doi.org/10.1017/9781139924696

Mann, E. M. (2011). *A comparison of stochastic claim reserving methods* [Report, Kansas State University]. https://krex.k-state.edu/server/api/core/bitstreams/f1cccf94-3a04-4cb5-a860-17209739acd3/content

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). Chapman & Hall/CRC. https://doi.org/10.1201/9780203753736

Ogutu, J. A. (2011). *Claims reserving using overdispersed poisson model* [Thesis]. http://erepository.uonbi.ac.ke/bitstream/handle/11295/14228/Ogutu%20Julie%20Amolo_Claims%20Reserving%20using%20Overdispersed%20Poisson%20Model.pdf?sequence=3&isAllowed=y

Renshaw, A. E., & Verrall, R. J. (1998). A stochastic model underlying the chain-ladder technique. *British Actuarial Journal*, *4*(4), 903–923. https://doi.org/10.1017/S1357321700000222

ROC/GIRO Working Party. (2007). Best Estimates and Reserving Uncertainty, Paper for GIRO 2007. https://www.actuaries.org.uk/system/files/documents/pdf/bhprizegibson.pdf

Soomro, F. S. H. (2019). *Technical workshop: Actuarial reserving for non-life business*. SHMA Consulting. https://shmaconsulting.com/wp-content/uploads/2019/11/Actuarial-Reserving-Falak-Sher-Haider-Soomro.pdf

Verrall, R. J. (1991). Chain ladder and maximum likelihood. *Journal of the Institute of Actuaries*, *118*(3), 489–499. https://doi.org/10.1017/s0020268100019545

Weke, P. G. O. (2008). *Deterministic and stochastic modelling of technical reserves in short-term insurance contracts*[Working Paper]. http://erepository.uonbi.ac.ke/bitstream/handle/11295/39399/Weke_Deterministic%20and%20stochastic%20modelling?sequence=1&isAllowed=y