MASTER THESIS

Title: The Clark LDF method: study and comparison with other claims reserving methods

Author: Raí Daniel Rivera Rubio

Advisors: Teresa Costa and Eva Boj

Academic year: 2022-2023

UNIVERSITAT DE BARCELONA
Facultat d'Economia i Empresa

Màster
de Ciències
Actuarials
i Financeres

Faculty of Economics and Business

Universitat de Barcelona

Master thesis

Master in Actuarial and Financial Sciences

THE CLARK LDF METHOD: STUDY AND COMPARISON WITH OTHER CLAIMS RESERVING METHODS

Author: Raí Daniel Rivera Rubio

Advisors: Teresa Costa and Eva Boj

The content of this document is the sole responsibility of the author, who declares that he has not incurred plagiarism and that all references to other authors have been expressed in the text.

ABSTRACT

Claims reserving are of utmost importance for an insurance company because they are the ones that will allow the company to face the liabilities acquired in the following periods. For this reason, implementing an accurate calculation of reserves is very important, since this amount will have an impact on the financial statements, besides being one of the important points to measure the solvency of such entity. For this purpose, a widely used model has been the deterministic Chain Ladder model, however, in the last decades, more sophisticated models including stochastic processes have been developed, thus, the calculation can be more accurate. The objective of this work is to use different models for the calculation of reserves, such as Clark LDF and Clark Cape Cod, in order to evaluate them and choose a reserve model that is accurate. Among the results found in this analysis, the Clark LDF model has shown less variability in the reserve projections, as opposed to the Chain Ladder model.

Keywords: Claims reserving, Chain Ladder model, Clark LDF, Cape Cod Model.

CONTENTS

1. INTRODUCTION

The projections of claims incurred and claims paid are of utmost importance in an insurance company, this is because the solvency of that company will depend on its ability to assume its liabilities in terms of claims, therefore, determining reserves amount that is capable of facing the claims that will be reported will be crucial to be solvent in the long term.

The objective of this work is to briefly study and apply the Clark method and compare it with other methods used when estimating reserves, to determine how efficient and accurate the Clark method could be compared to other methods. It is interesting to review the Clark method, since it has not been reviewed in this master's degree in actuarial and financial sciences, and it is important to know new methodologies for the calculation of reserves.

In recent decades various models have been developed to estimate the reserves of an insurance company, due to the importance of these reserves, what has been sought is greater precision at the time of estimating these values, since they directly affect the income statements of insurance undertakings. There can be deterministic models, such as the classic Chain Ladder model, or stochastic models such as Mack Chain Ladder model, multivariate Chain Ladder model, Munich Chain Ladder model, the Clark method, etc. In this work we will briefly review Chain Ladder and Mack Chain Ladder models, and, additionally, we will give more emphasis on Clark's method, which employs the LDF and Cape Cod models. The objective will be to replicate that methodology and apply it to claims data to compare the results of the projected reserves with some other classical models and to be able to draw conclusions about it.

This work is structured as follows: First, we will present a brief theoretical review of reserving methods, and then we will review the mathematical development of the method proposed by Clark (2003). Once we have analyzed the model in depth, we will proceed to use it on a claims dataset and compare the results with other models we have applied. Additionally, we have developed a Shiny application to show the obtained results. Finally, we will present the conclusions of this master thesis and the bibliography.

2. BACKGROUND

Over the last decades, the calculation of insurance reserves has presented a great development, due to the search for a better precision in its estimation, which will influence the solvency of an insurance company. An adequate reserve will determine if it is capable of facing the future obligations that may arise from the provision of insurance services.

As we have mentioned, several models have been developed with different methodologies, ranging from classical and deterministic models to models that use more statistics, such as stochastic models.

In general, the reserves have been represented in a matrix format, with the upper left triangle representing the losses amount at the time of the analysis and the lower right triangle the loss values for the following periods to be estimated. This matrix is structured as follows: in the rows i is the year in which claims occurred, while in the columns j is the year of development or payment of loss, i.e., the first cell on the upper left is the losses amount of claims that occurred in year i, which is paid in the same year, with j being

equal to 0. Therefore, to give a more practical example, if we have a matrix of losses and reserves, which has 10 years of occurrence, starting in 2000, that is, 10 rows, and this matrix has 9 years of development, 10 columns, taking into account that the first column is the year of development 0, the cell on the upper right is the losses amount of claims occurred in year 2000, that is paid after 9 years of its occurrence, in 2009. For this reason, we can say that the last year of the claim's occurrence in the matrix, known as origin period or accident year, is called I, while the last year of development would be $J = I - 1$.

Accordingly, let us introduce some notation: we start with a random variable $c_{i,j}$, which refers to the incremental amounts, i.e., it represents the incremental amount paid for claims occurred in the accident year i in the development period j. It is important to mention that these amounts have been paid in calendar year $i + j$.

Now, to represent these amounts cumulatively, we will introduce the notation $C_{i,j}$, which refers to the total amount of claims occurred in period i that have been paid up to the development period j, thus, in the last development period J, the cumulative amount $C_{i,J}$, of all claims that occurred in i, will have been paid.

By way of summary, in the upper left triangle we have the observed values of losses, it is known as *run-off* triangle in literature, and the values to be predicted are in the lower right triangle, therefore, the time at which they are evaluated is in period I, for this reason, if cell $i + j < I$, that cell has a real loss amount, on the contrary, if cell $i + j \geq I$, that cell should be predicted.

In **Table 1**, we can see a matrix in which the above mentioned is visualized.

Accident year i	Development year j								
	0	**1**	**2**	**...**	**j**	**...**	**$J-2$**	**$J-1$**	**$J = I-1$**
1									...
2								...	
3							...		
⋮		Observations				...			
i					...				
⋮				...			To predict		
$I-2$			...						
$I-1$		...							
I	...								

Table 1. Losses and reserves matrix

Source: Own elaboration

Continuing with the explanation of the matrix, we are now interested in identifying the current values, future values, and ultimate values in the matrix containing cumulative losses. The current values would be the counter diagonal of the matrix, that is, we do not place them in period I, and they would be all the accumulated amounts of claims paid per year of origin until the period I. The next diagonal (from left to right) would be the projected values for the following period per year of origin; it is very common to work with annual periods, in this way, each diagonal that continues after the counter diagonal (from left to right), are the future values for the following years. Finally, to know the ultimate value to be paid up to period $I + J$, by year of origin, the last column of the matrix must be taken.

Therefore, if we wish to know the future value, or in other words, the reserve needed to meet the future payments of the period of origin i, it could be calculated as follows:

$$R_i = \sum_{\ell=I-i+1}^{I-1} \hat{c}_{i\ell} = \hat{C}_{i,I} - C_{i,I-i}. \tag{1}$$

Finally, the total reserve is:

$$R = \sum_{i=2}^{I} R_i. \tag{2}$$

2.1. CHAIN LADDER MODEL

Once we have outlined the structure of the matrix for predicting reserve amounts, we will proceed to a summary of certain models that will help us to predict these future loss amounts.

To begin this brief review, we will start with the model that has been used the most, and one of the pioneers, is the deterministic Chain Ladder method. This method has been widely used due to its simplicity, since it does not require any advanced software and does not use any statistical tool that generates difficulties at the time of its use. In such a way that, as a summary, the method consists of calculating the cumulative losses value of the following development year by means of the value of the preceding period multiplied by a development factor.

The formula for the value of the cumulative loss for the following period according to the Chain Ladder method is as follows:

$$C_{i,j+1} = f_j \cdot C_{i,j} \tag{3}$$

where:

f_j is the development factor.

Therefore, we are interested in knowing how to calculate the development factor of the classical Chain Ladder method and the rest of the projected values, which are given using the following formulas. For more details, see Amin *et al.* (2018):

$$\hat{f}_j = \frac{\sum_{i=1}^{I-j-1} C_{i,j+1}}{\sum_{i=1}^{I-j-1} C_{ij}} \tag{4}$$

$$\hat{C}_{ij} = C_{i,I-i} \cdot \prod_{l=I-i}^{j-1} \hat{f}_l. \tag{5}$$

2.2. MACK CHAIN LADDER MODEL

The Chain Ladder model only allows us to estimate the reserves amount but not the variability around this amount. To measure the uncertainty related to the calculation of the reserve, some stochastic methods have been developed.

Mack (1993) describes a free distribution model that reproduces the same results than the Chain Ladder method and introduces the calculation of the standard errors of the reserve estimation. For this, Mack (1993) begins by describing the assumptions made by the classical Chain Ladder model and the theorems that support these assumptions, to finally estimate the standard errors.

The implicit assumptions of the Chain Ladder model are:

$$E\big(C_{i,k+1}\big|C_{i1}, \dots, C_{ik}\big) = C_{ik} \cdot f_k, 1 \leq i \leq I, 1 \leq k \leq I-1 \tag{6}$$

$$\{C_{il}, \dots, C_{iI}\}, \{C_{jl}, \dots, C_{jI}\}, i \neq j, \text{ are independent.} \tag{7}$$

Regarding the theorems that support the assumptions, we have:

Theorem 1: Let $D = \{C_{ik}|i + k \leq I + 1\}$ be the set of all data observed so far.

Under the assumptions (6) and (7) we have

$$E(C_{iI}|D) = C_{i,I+1-i} \cdot f_{I+1-i} \cdot \dots \cdot f_{I-1}. \tag{8}$$

Theorem 2: Under the assumptions (6) and (7) the estimators $\hat{f}_k, 1 \leq k \leq I-1$, are unbiased and uncorrelated.

Now, to calculate the mean squared error and standard errors, Mack (1993) begins by defining the mean squared error formula:

$$mse(\hat{C}_{iI}) = E\left(\big(\hat{C}_{iI} - C_{iI}\big)^2\big|D\right) \tag{9}$$

9

where $D = \{C_{ik}|i + k \leq I + 1\}$ is the set of all observed data.

$$mse(\hat{R}_i) = E\left(\left(\hat{R}_i - R_i\right)^2\Big|D\right) = E\left(\left(\hat{C}_{iI} - C_{iI}\right)^2\Big|D\right) = mse(\hat{C}_{iI}). \qquad (10)$$

Now, applying the general rule, we have that:

$$E(X - a)^2 = Var(X) + (E(X) - a)^2. \qquad (11)$$

Based on the above equation, we can obtain the following expression:

$$mse(\hat{C}_{iI}) = Var(C_{iI}|D) + \left(E(C_{iI}|D) - \hat{C}_{iI}\right)^2. \qquad (12)$$

Now, we define the variance of C:

$$Var\left(C_{i,k+1}\big|C_{iI}, \ldots, C_{ik}\right) = C_{ik} \cdot \sigma_k^2, 1 \leq i \leq I, 1 \leq k \leq I - 1 \qquad (13)$$

with σ_k^2 being an unknown parameter, which is defined in the following equation:

$$\hat{\sigma}_k^2 = \frac{1}{I - k - 1}\sum_{i=1}^{I-k} C_{ik}\left(\frac{C_{i,k+1}}{C_{ik}} - \hat{f}_k\right)^2, 1 \leq k \leq I - 2. \qquad (14)$$

Based on the above, the following theorem is obtained:

Theorem 3: Under the assumptions (6), (7) and (13) the mean squared error $mse(\hat{R}_i)$ can be estimated by:

$$mse(\hat{R}_i) = \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2}\left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}}\right). \qquad (15)$$

For a better understanding about the model that has been developed by Mack, it is recommended to review his article (Mack, 1993).

In addition, more stochastic reserve estimation models are available, such as the Bornhuetter-Ferguson, Merz and Wüthrich, GLM and Bootstrap models, etc., for more details of these models, see the following articles: Amin *et al*. (2018), Schmidt and Zocher (2008) and England and Verrall (2002).

2.3. CLARK METHOD – LDF AND CAPE COD MODELS

As mentioned, for the estimation of insurance reserves, we will use the methodology proposed by David Clark (Clark, 2003), therefore, this article will serve as the main guide

to develop and apply the LDF and Cape Cod models to our data. To summarize, this methodology consists first on estimating a pattern of emerging losses, in order to estimate the value of the loss per year. In such a way that, when we have obtained this emerging loss pattern, the main models to develop for the estimation of reserves will be the LDF model and the Cape Cod model. To use these models, we must identify the distribution of the occurrence of real losses and, once we have identified this distribution, we will use the models mentioned in this distribution.

Based on what was mentioned in the previous paragraph, to estimate our model, we start by estimating the pattern of emerging losses, to estimate this pattern we will use two curves, the Weibull and the Log-logistic curve. These curves are widely used in the actuarial field; therefore, we will estimate the parameters of these curves and we will use the maximum likelihood function to select the appropriate parameters. Next, we detail the functions of the curves that will be used to estimate this pattern.

The Log-logistic curve has the form:

$$G(t|\omega,\theta) = \frac{t^{\omega}}{t^{\omega} + \theta^{\omega}}. \tag{16}$$

The Weibull curve has the form:

$$G(t|\omega,\theta) = 1 - exp(-(t/\theta)^{\omega}) \tag{17}$$

where θ is a scale parameter and ω is a shape parameter.

It is important to mention that these curves are used because the behavioral pattern presented by these curves is an increasing pattern; therefore, the methodology proposed by Clark (2003) seeks to find a curve in which the amount of the loss moves as fast as possible from 0% to 100% in an increasing pattern.

Now that we have established the two curves that will be used to estimate the increasing pattern of losses, we will define the two models that will be used to estimate the reserves amount. The models proposed by Clark (2003) are the LDF model and the Cape Cod model. Basically, the difference between these two models is that the LDF model considers that the estimation of the loss development for each year is independent, therefore, a parameter must be estimated for each year of estimation, while the Cape Cod model considers that there is a relationship between the expected loss values for each year in the period studied, and this existing relationship is given by an exposure base, which can be the level premium or known in the insurance sector as the earned premium, therefore, this model will only have 3 parameters to be estimated.

In the following, we set out the models that have been mentioned, for this purpose we establish that the expected incremental loss value, $\hat{\mu}_{i,j}$, with the two models is:

Model #1 "Cape Cod"

The expected loss is estimated based on the growth curve and the expected loss ratio:

$$\hat{\mu}_{i,j} = \text{Premium}_i \cdot ELR \cdot [G(j|\omega,\theta) - G(j-1|\omega,\theta)] \tag{18}$$

with 3 parameters to estimate:

$$ELR, \omega, \theta. \tag{19}$$

Model #2 "LDF"

The expected loss is estimated based on the growth curve and an estimated ultimate loss for each accident year:

$$\hat{\mu}_{i,j} = Ult_i \cdot [G(j|\omega,\theta) - G(j-1|\omega,\theta)] \tag{20}$$

with $n+2$ parameters to estimate: n accident year (one ultimate loss for each accident year), ω, θ.

Based on the above, it is important to take into account that the LDF model presents an over parameterization, as mentioned in Clark (2003), if we have a triangle in which 10 years are considered, there would be a total of 55 data, therefore, for the LDF model, 12 parameters should be estimated, while for the Cape Cod model only 3. It is important to highlight that the Cape Cod method could present a higher variance of the estimated process, however, the total variance could be somewhat lower due to the fact that the variance of the parameters is larger in the LDF model. It is worth mentioning that what is sought is greater precision and less variability in the estimation of reserves, since this will allow us to deal with claims.

Now, the next step is to define the distribution of the real losses, in order to be able to estimate the variance around the estimated value. To do this, Clark (2003) proposed to use the Poisson distribution, and once we have considered this distribution, by means of the maximum likelihood function, we will select the best parameters.

To estimate the variance around the estimated value, it is important to keep in mind that the variance is divided into two parts, we have the process variance and the parameters variance. The variance process refers to the randomness that the stock estimates can take, while the parameters variance refers to the uncertainty in our parameters.

Process variance

For this section, it is important to note that, in any period analyzed, the loss has a constant ratio of variance mean, which is given by the following expression:

$$\frac{Variance}{Mean} = \sigma^2 \approx \frac{1}{n-p} \cdot \sum_{i,j}^{n} \frac{(c_{i,j} - \hat{\mu}_{i,j})^2}{\hat{\mu}_{i,j}} \tag{21}$$

where:

p is the number of parameters

$c_{i,j}$ is the actual incremental loss

$\hat{\mu}_{i,j}$ is the expected incremental loss

(This is recognized as being equivalent to a chi-square error term).

Furthermore, it should be noted that the previously defined variable, $G(t|\omega,\theta)$, is the pattern of the occurrence of losses, therefore, the actual loss will present a distribution around this variable.

As mentioned in previous paragraphs, the variable $c_{i,j}$ is the variable that we call the incremental real loss, therefore, we are interested in assuming a distribution to this variable in order to model it, for this reason, Clark (2003) proposes that $c_{i,j}$ follows an over dispersed Poisson distribution. One of the main advantages of assuming a discrete distribution in a continuous variable is that by having a discrete variable, small changes in the variable will not be taken into account, since being a discrete distribution and Poisson, we will have a large mass in zeros. These small incremental values do not necessarily represent a loss of precision, since generally the mean is larger than the σ^2 scale factor. Additionally, Clark (2003) mentions that the maximum likelihood method aims to estimate the mean and variance of the reserve's distribution, therefore, once these two parameters are estimated by means of the maximum likelihood function, we will be able to adopt some other distribution.

To estimate the second part of the variance, the variance of the parameters, we first proceed to establish the maximum likelihood function for the estimation of the best parameters.

Best Parameters

$$\text{Likelihood} = \prod_i \Pr(c_i) = \prod_i \frac{\lambda_i{}^{c_i/\sigma^2} \cdot e^{-\lambda_i}}{(c_i/\sigma^2)!} = \prod_i \frac{(\mu_i/\sigma^2)^{c_i/\sigma^2} \cdot e^{-\mu_i/\sigma^2}}{(c_i/\sigma^2)!} \quad (22)$$

$$\text{LogLikelihood} = \sum_i (c_i/\sigma^2) \cdot \ln(\mu_i/\sigma^2) - \mu_i/\sigma^2 - \ln\big((c_i/\sigma^2)!\big) \quad (23)$$

$$l = \sum_i c_i \cdot \ln(\mu_i) - \mu_i \quad \text{if } \sigma^2 \text{ is assumed to be known.} \quad (24)$$

Based on what has been developed, we now apply it to the first model, the Cape Cod model.

$$l = \sum_{i,j} \big(c_{i,j} \cdot \ln(ELR. P_i[G(j) - G(j-1)]) - ELR \cdot P_i[G(j) - G(j-1)]\big) \quad (25)$$

where:

$c_{i,j}$ is the actual loss in accident year i and development period j

13

P_i is the premium for accident year i

$$For \; \frac{\delta l}{\delta ELR} = 0, \qquad \widehat{ELR} = \frac{\sum_{i,j} c_{i,j}}{\sum_{i,j} P_i \cdot [G(j) - G(j-1)]}. \tag{26}$$

Applied to the second model, LDF model:

$$l = \sum_{i,j} (c_{i,j} \cdot \ln(Ult_i \cdot [G(j) - G(j-1)]) - Ult_i \cdot [G(j) - G(j-1)]) \tag{27}$$

$$\frac{\delta l}{\delta Ult_i} = \sum_j \left(\frac{c_{i,j}}{Ult_i} - [G(j) - G(j-1)] \right) \tag{28}$$

$$For \; \frac{\delta l}{\delta Ult_i} = 0, \quad \widehat{Ult_i} = \frac{\sum_j c_{i,j}}{\sum_j [G(j) - G(j-1)]}. \tag{29}$$

Now that we have obtained the best parameters, we can calculate the second part of the variance, the variance of the parameters.

Parameter's variance

To calculate the variance of the estimated parameters, it is necessary to calculate the information matrix of the second derivative (I), a calculation known as Delta Method. And once we have obtained this matrix, we proceed to calculate the variance of the reserves, which considers the process variance and the parameters variance. In this master thesis the development of the variance of the parameters will be reviewed superficially, if more detail is required, it is recommended to review the main article by Clark (2003).

In the case of the Cape Cod model, the information matrix of the second derivative is 3×3, since, as mentioned above, for this method it is only necessary to estimate three parameters. This matrix is shown below.

$$I = \begin{bmatrix} \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial ELR^2} & \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial ELR \partial \omega} & \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial ELR \partial \theta} \\ \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial \omega \partial ELR} & \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial \omega^2} & \sum_{y,j} \frac{\partial^2 y,j}{\partial w \partial \theta} \\ \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial \theta \partial ELR} & \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial \theta \partial \omega} & \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial \theta^2} \end{bmatrix}. \tag{30}$$

Once we have calculated the information matrix of the second derivative, to calculate the covariance matrix, we simply calculate it using the inverse of the I-matrix.

$$\sum = \begin{bmatrix} Var(ELR) & Cov(ELR,\omega) & Cov(ELR,\theta) \\ Cov(\omega,ELR) & Var(\omega) & Cov(\omega,\theta) \\ Cov(\theta,ELR) & Cov(\theta,\omega) & Var(\theta) \end{bmatrix} \geq -\sigma^2 \cdot I^{-1}. \tag{31}$$

For the case of the LDF model, since we have mentioned that this model uses $n + 2$ parameters, where n is the number of periods to be used in the model, we have a matrix of size $(n + 2) \times (n + 2)$. Like the previous matrix, the covariance matrix is calculated with the inverse of the I matrix calculated for the LDF model.

$$\begin{bmatrix} \sum_j \frac{\partial^2 l_{1,j}}{\partial Ult_1^2} & 0 & \cdots & 0 & \sum_j \frac{\partial^2 l_{1,j}}{\partial Ult_1 \partial \omega} & \sum_j \frac{\partial^2 l_{1,j}}{\partial Ult_1 \partial \theta} \\ 0 & \sum_j \frac{\partial^2 l_{2,j}}{\partial Ult_2^2} & \cdots & 0 & \sum_j \frac{\partial^2 l_{2j}}{\partial Ult_2 \partial \omega} & \sum_j \frac{\partial^2 l_{2,j}}{\partial Ult_2 \partial \theta} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sum_j \frac{\partial^2 l_{n,j}}{\partial Ult_n^2} & \sum_j \frac{\partial^2 l_{n,j}}{\partial Ult_n \partial \omega} & \sum_j \frac{\partial^2 l_{n,j}}{\partial Ult_n \partial \theta} \\ \sum_j \frac{\partial^2 l_{1,j}}{\partial \omega \partial Ult_1} & \sum_j \frac{\partial^2 l_{2,j}}{\partial \omega \partial Ult_2} & \cdots & \sum_j \frac{\partial^2 l_{n,j}}{\partial \omega \partial Ult_n} & \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial \omega^2} & \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial \omega \partial \theta} \\ \sum_j \frac{\partial^2 l_{1,j}}{\partial \theta \partial Ult_1} & \sum_j \frac{\partial^2 l_{2,j}}{\partial \theta \partial Ult_2} & \cdots & \sum_t \frac{\partial^2 l_{n,t}}{\partial \theta \partial Ult_n} & \sum_{y,j} \frac{\partial^2 l_{yj}}{\partial \theta \partial \omega} & \sum_{y,j} \frac{\partial^2 l_{y,j}}{\partial \theta^2} \end{bmatrix}. \tag{32}$$

Finally, to calculate the variance of the reserves, the variance process and the variance of the parameters are taken into account.

The process variance of R:

$$\sigma^2. \sum \mu_{i,j}. \tag{33}$$

And for the parameter's variance, we use the following formula, with the covariance matrix Σ:

$$Var(E[R]) = (\partial R)' \cdot \Sigma(\partial R). \tag{34}$$

15

Once we have developed the entire method proposed by Clark (2003), i.e., the calculation of reserves by means of the LDF model and the Cape Cod model, it is important to highlight all the assumptions of this models:

- incremental losses are independent and identically distributed (idd)
- the variance/mean scale parameter σ^2 is fixed and known
- variance estimates are based on an approximation to the Rao-Cramer lower bound.

3. NUMERICAL APPLICATION

To start whit these, numerical examples we make a brief presentation of the data that we will use, and then, the following steps to estimate the models.

The claims data we will use is the ABC database included in the ChainLadder package (Gesmann *et al,*. 2023) in R. This database is an accumulated *run- off* triangle of a company's 11-year portfolio of workers' compensation claims. Therefore, we will make a brief visual description of the used data, with the first year of origin losses, we will plot the growth function discussed by Clark (2003). In this way, we will adjust the Weibull and the Log-logistic distributions with these data and using the information criteria AIC and BIC, we will determine which curve is the most appropriate to model the used data, and thus have an idea of which curve to use in the models developed by Clark in his analysis.

Once we have reached this point, we will proceed to estimate the reserves with the different methods. We will start with the deterministic Chain Ladder model, then, we will use the Mack Chain Ladder model, this last one presents the same reserve values of the deterministic Chain Ladder model, however, as we mentioned in the previous section, it allows us to obtain the standard errors, so we can evaluate the estimation we have made and compare its uncertainty with that of the Clark methods. Finally, we will estimate the Clark LDF and Clark Cape Cod models with the two distributions mentioned, to observe the differences that the results of the model present when using these distributions.

We begin by presenting the following *run-off* triangle to visualize the incremental claims amount according to the accident year and their respective development.

Accident year i	Incremental amount by development year j										
	1	2	3	4	5	6	7	8	9	10	11
1977	153,638	188,412	134,534	87,456	60,348	42,404	31,238	21,252	16,622	14,440	12,200
1978	178,536	226,412	158,894	104,686	71,448	47,990	35,576	24,818	22,662	18,000	NA
1979	210,172	259,168	188,388	123,074	83,380	56,086	38,496	33,768	27,400	NA	NA
1980	211,448	253,482	183,370	131,040	78,994	60,232	45,568	38,000	NA	NA	NA
1981	219,810	266,304	194,650	120,098	87,582	62,750	51,000	NA	NA	NA	NA
1982	205,654	252,746	177,506	129,522	96,786	82,400	NA	NA	NA	NA	NA
1983	197,716	255,408	194,648	142,328	105,600	NA	NA	NA	NA	NA	NA
1984	239,784	329,242	264,802	190,400	NA	NA	NA	NA	NA	NA	NA
1985	326,304	471,744	375,400	NA	NA	NA	NA	NA	NA	NA	NA
1986	420,778	590,400	NA	NA	NA	NA	NA	NA	NA	NA	NA
1987	496,200	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 2. Incremental amount by origin year i and development year j

Source: Own elaboration

To estimate the models, we need the accumulated claims losses, therefore, the following triangle presents the above one in a cumulative way:

Accident year i	Accumulated amount by development year j										
	1	2	3	4	5	6	7	8	9	10	11
1977	153,638	342,050	476,584	564,040	624,388	666,792	698,030	719,282	735,904	750,344	762,544
1978	178,536	404,948	563,842	668,528	739,976	787,966	823,542	848,360	871,022	889,022	NA
1979	210,172	469,340	657,728	780,802	864,182	920,268	958,764	992,532	1,019,932	NA	NA
1980	211,448	464,930	648,300	779,340	858,334	918,566	964,134	1,002,134	NA	NA	NA
1981	219,810	486,114	680,764	800,862	888,444	951,194	1,002,194	NA	NA	NA	NA
1982	205,654	458,400	635,906	765,428	862,214	944,614	NA	NA	NA	NA	NA
1983	197,716	453,124	647,772	790,100	895,700	NA	NA	NA	NA	NA	NA
1984	239,784	569,026	833,828	1,024,228	NA	NA	NA	NA	NA	NA	NA
1985	326,304	798,048	1,173,448	NA	NA	NA	NA	NA	NA	NA	NA
1986	420,778	1,011,178	NA	NA	NA	NA	NA	NA	NA	NA	NA
1987	496,200	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 3. Accumulated amount by origin year i and development year j
Source: Own elaboration

In a more visual way, we proceed to graph the development of each accident year in an accumulated way:
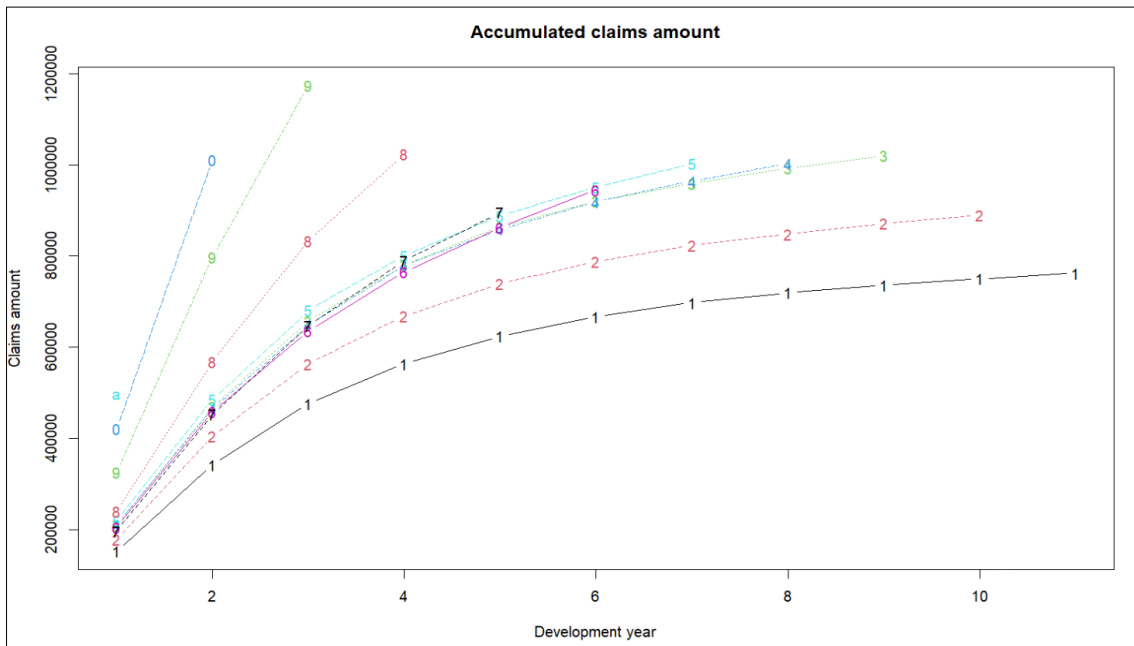

Figure 1. Accumulated claims amount
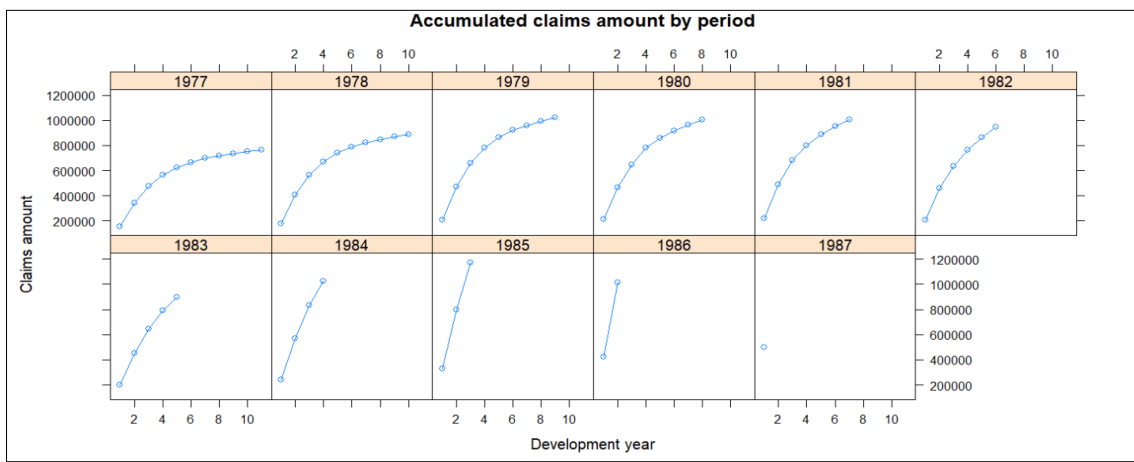Source: Own elaboration

Figure 2. Accumulated claims amount by period
Source: Own elaboration

In the graphs presented, we can observe the pattern of behavior of the accumulated losses, this pattern is the one we seek to model, that is, this pattern is called a growth function in the article by Clark (2003), and, therefore, we will use two curves (Weibull and Log-logistic) to find out which distribution best fits the data.

In the following graph we can observe the development of the claims amount in a non-cumulative way, that is, the claims amount that develops period after period according to the origin period or accident year. As we can see, the claims amount is decreasing as the periods go by, highlighting the first 3 periods in which more than 50% of the claims amount have been disbursed according to the accident year.
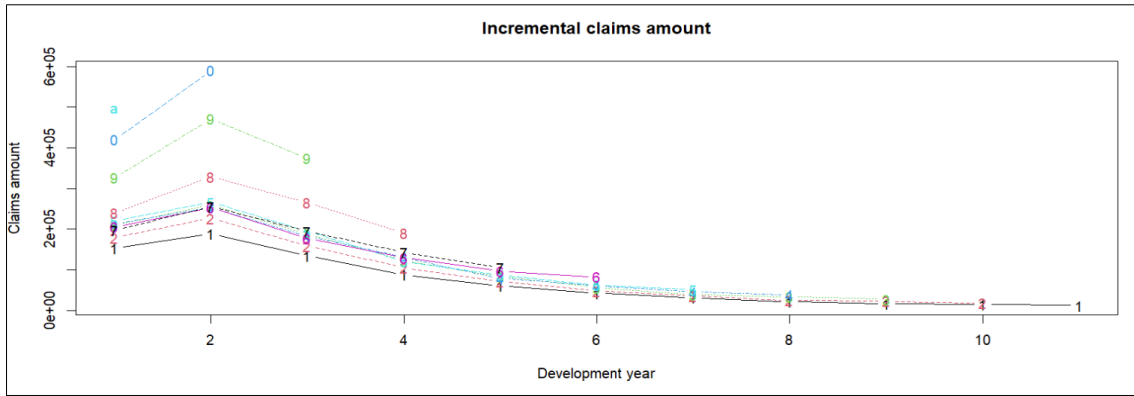
19

Figure 3. Incremental claims amount

Source: Own elaboration

Now, we proceed to fit the Weibull and Log-logistic distribution functions to our data, to have an idea of which curve we should use to estimate the Clark LDF and Clark Cape Cod models. Applying the AIC and BIC information criteria, we present the following results:

Information criteria		
Distribution	**AIC**	**BIC**
Weibull	3.770	4.565
Log-logistic	8.763	9.559

Table 4. Information criteria

Source: Own elaboration

Based on the information criteria presented (AIC and BIC), we can see that our data best fit the Weibull distribution.

Next, we can observe the development of the claims amounts for the period 1977, which are the data we have fitted.

Figure 4. Claims amount - 1977

Source: Own elaboration

20

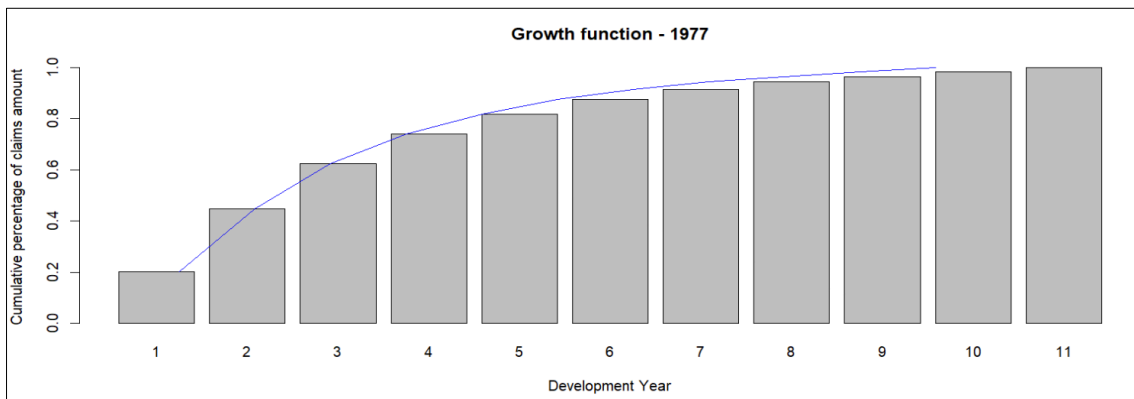The following graph shows the growth curve for the 1977 period.

Figure 5. Growth function - 1977

Source: Own elaboration

CHAIN LADDER

Once we have performed a brief analysis of our data, we proceed to estimate the proposed models. We start with the deterministic Chain Ladder method since its estimation is quite simple. For this, using the formula for the development factors described in the previous section, we obtain the following factors:

Development factors	2.309	1.421	1.200	1.113	1.073	1.048	1.034	1.026	1.020	1.016

Table 5. Development factors Chain Ladder

Source: Own elaboration

Applying these development factors, we obtain the following matrix:

Accident year i	Accumulated amount by development year j										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1977	153,638	342,050	476,584	564,040	624,388	666,792	698,030	719,282	735,904	750,344	762,544
1978	178,536	404,948	563,842	668,528	739,976	787,966	823,542	848,360	871,022	889,022	903,477
1979	210,172	469,340	657,728	780,802	864,182	920,268	958,764	992,532	1,019,932	1,040,522	1,057,440
1980	211,448	464,930	648,300	779,340	858,334	918,566	964,134	1,002,134	1,028,236	1,048,994	1,066,050
1981	219,810	486,114	680,764	800,862	888,444	951,194	1,002,194	1,036,480	1,063,477	1,084,946	1,102,586
1982	205,654	458,400	635,906	765,428	862,214	944,614	989,539	1,023,392	1,050,048	1,071,246	1,088,663
1983	197,716	453,124	647,772	790,100	895,700	960,849	1,006,547	1,040,981	1,068,095	1,089,658	1,107,375
1984	239,784	569,026	833,828	1,024,228	1,140,421	1,223,371	1,281,553	1,325,396	1,359,918	1,387,372	1,409,929
1985	326,304	798,048	1,173,448	1,408,060	1,567,797	1,681,832	1,761,818	1,822,091	1,869,551	1,907,292	1,938,303
1986	420,778	1,011,178	1,436,983	1,724,284	1,919,895	2,059,540	2,157,490	2,231,299	2,289,417	2,335,635	2,373,611
1987	496,200	1,145,527	1,627,905	1,953,379	2,174,979	2,333,178	2,444,142	2,527,757	2,593,597	2,645,956	2,688,977

Table 6. Accumulated amount by origin year i and development year j

Source: Own elaboration

In summary, we present below the values of this matrix developed, separating the amounts of current losses and the amounts of future losses by year of origin:

Years	Current Value	Future Value	Ultimate Value
1977	762,544	-	762,544
1978	889,022	14,455	903,477
1979	1,019,932	37,508	1,057,440
1980	1,002,134	63,916	1,066,050
1981	1,002,194	100,392	1,102,586
1982	944,614	144,049	1,088,663
1983	895,700	211,675	1,107,375
1984	1,024,228	385,701	1,409,929
1985	1,173,448	764,855	1,938,303
1986	1,011,178	1,362,433	2,373,611
1987	496,200	2,192,777	2,688,977
Total	10,221,194	5,277,760	15,498,955

Table 7. Chain Ladder model results

Source: Own elaboration

As we can see, we have a total of 10.2 million of present value of claims amounts, applying the deterministic Chain Ladder model, we obtain that 5.3 million remains to be developed, in other words, having 11 origin period, at the end of 11 periods of development for each origin period, we obtain a claims amount of 15.5 million. This estimated value (5.3 million) is 52% of the current claims amounts.

MACK CHAIN LADDER

Continuing with the model estimations, we proceed to use the Mack Chain Ladder model, which will present the same values of future losses, however, this model will now estimate the standard errors, that is, the possible deviations from what we estimated with the deterministic Chain Ladder model. This is the added value of this model and the reason for using this model in this analysis, since, in this way, it will allow us to compare the classic Chain Ladder model with the models developed by Clark (2003).

Using the *MackChainLadder* function of the *ChainLadder* package of R, we proceed to obtain the following results for the eleven development periods we are calculating:

			Mack Chain Ladder			
Years	**Current Value**	**Dev.To. Date**	**Ultimate Value**	**Future Value**	**Mack.S. E**	**CV%**
1977	762,544	1.00	762,544	-	-	
1978	889,022	0.98	903,477	14,455	285	2.0
1979	1,019,932	0.97	1,057,440	37,508	923	2.5
1980	1,002,134	0.94	1,066,050	63,916	2,758	4.3
1981	1,002,194	0.91	1,102,586	100,392	5,715	5.7
1982	944,614	0.87	1,088,663	144,049	7,613	5.3
1983	895,700	0.81	1,107,375	211,675	14,854	7.0
1984	1,024,228	0.73	1,409,929	385,701	22,419	5.8
1985	1,173,448	0.61	1,938,303	764,855	37,293	4.9
1986	1,011,178	0.43	2,373,610	1,362,432	62,244	4.6
1987	496,200	0.19	2,688,977	2,192,777	107,919	4.9
Total	10,221,194	0.66	15,498,954	5,277,760	152,283	3.0

Table 8. Mack Chain Ladder model results

Source: Own elaboration

The first column of the table above shows the current values, the second column refers to the percentage that the current values represent compared to the ultimate values. The third and fourth columns show the ultimate value of the losses and the projected value of the reserve by accident year, respectively. Finally, in the fifth and sixth columns of the table there are the standard errors of the estimate, i.e., in the sixth column we can see the percentage by which the reserve estimate is dispersed, based on the calculated standard error (fifth column), being this value 3% overall.
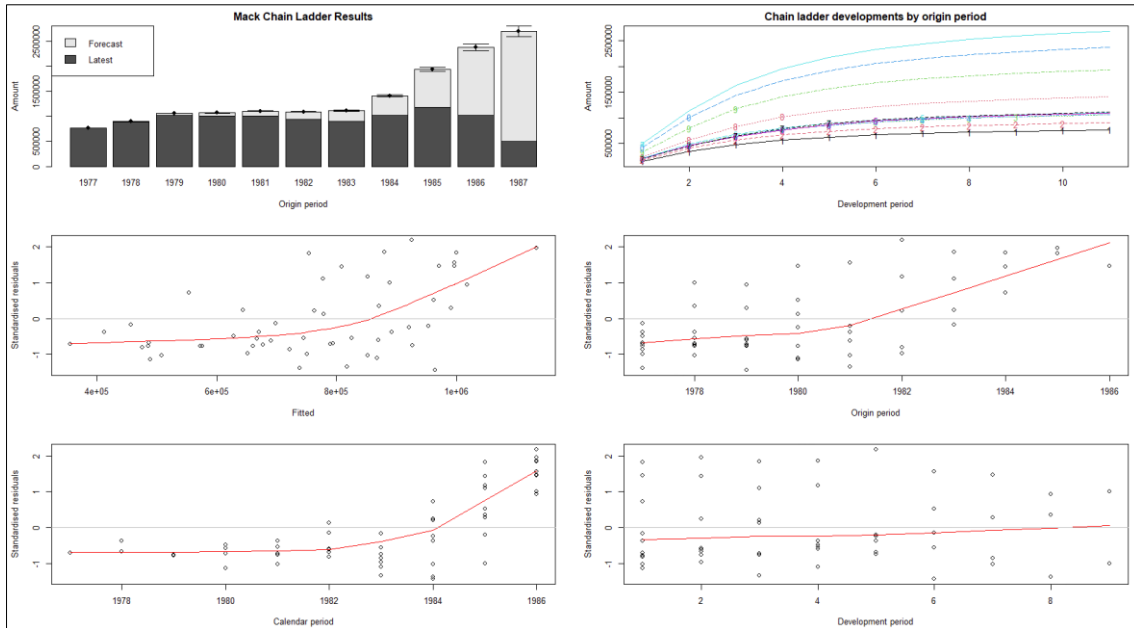
Next, we present some graphs of the estimated model:

Figure 6. Mack Chain Ladder model graphs

Source: Own elaboration

To summarize, what we expect from the graph of the standardized residuals is that they are random, i.e., they do not show a clear trend and are close to zero, so that the residuals of the model used are showing a slight trend. Additionally, we can also visualize that the residuals increase and are greater than zero as the years of development pass with respect to the origin period.

CLARK LDF MODEL

As for the models proposed by Clark (2003), we start with the Clark LDF model. For this, in the same way, we will use the *ClarkLDF* function of the R package *ChainLadder*. It is important to mention that, a priori, based on the fit of our data that we made in previous paragraphs, we have obtained that our data fit a Weibull distribution, therefore, we proceed to use this distribution to estimate the growth function of the Clark LDF model. Also, it should be noted that the information criteria do not have to be an absolute truth, that is, it is also important to take into account the results of the models considering the other options that did not present the best AIC or BIC, in order to make a decision, considering the practical or theoretical knowledge of the analyst.

Next, we present the results of the estimated model. It is important to note that the model we will estimate, like the previous one, only estimates eleven development periods for each source period.

	Clark LDF model (Weibull)					
Origin	**Current Value**	**Ldf**	**Ultimate Value**	**Future Value**	**St. Error**	**CV%**
1977	762,544	1.00	762,544	-	-	
1978	889,022	1.01	894,574	5,552	3,970	71.5
1979	1,019,932	1.02	1,037,148	17,216	7,262	42.2
1980	1,002,134	1.04	1,037,045	34,911	10,697	30.6
1981	1,002,194	1.07	1,067,393	65,199	15,132	23.2
1982	944,614	1.12	1,054,336	109,722	20,221	18.4
1983	895,700	1.20	1,078,598	182,898	27,054	14.8
1984	1,024,228	1.36	1,395,497	371,269	41,378	11.1
1985	1,173,448	1.67	1,962,910	789,462	67,379	8.5
1986	1,011,178	2.40	2,426,726	1,415,548	107,371	7.6
1987	496,200	5.06	2,508,246	2,012,046	184,120	9.2
Total	10,221,194		15,225,019	5,003,825	276,442	5.5

Table 9. Clark LDF model (Weibull)

Source: Own elaboration

Based on the results presented, it can be extracted that the estimated reserve amount is 5 million, thus obtaining a total loss value of 15.23 million. As with the previous results, columns 5 and 6 show the calculation of the standard errors of the estimates and the percentage that these represent in relation to the estimated reserve value, thus obtaining an overall value of 5.5% standard error in relation to the estimated future value. The maximum value of the growth function of this model is 0.9916

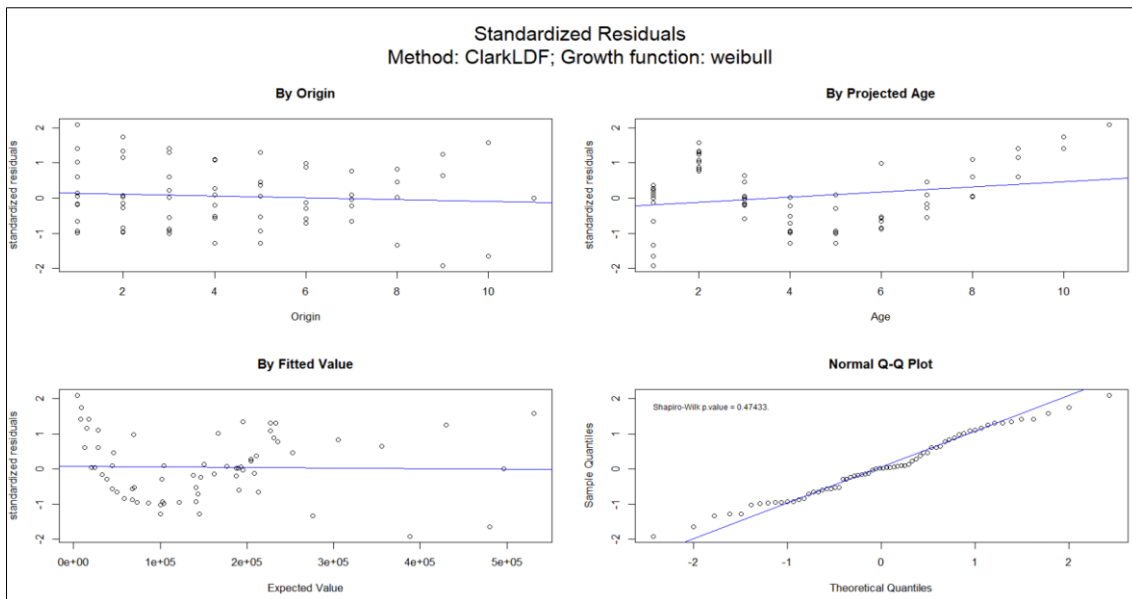Next, we present a graph of the results obtained from the Clark LDF model with Weibull distribution:

Figure 7. Standardized residuals method: CLarkLDF; Growth function: Weibull

Source: Own elaboration

As previously mentioned, we expect the standardized residuals to present a random behavior around 0, in a certain way, the standardized residuals have the desired behavior, taking into account that there is a very slight tendency.

Now, to compare the results of the model estimation with the different distributions proposed by Clark (2003), we proceed to estimate the model, using the Log-logistic distribution.

	Clark LDF model (Log-logistic)					
Origin	Current Value	Ldf	Ultimate Value	Future Value	Std. Error	CV%
1977	762,544	1.00	762,544	-	-	
1978	889,022	1.02	902,124	13,102	3,114	23.8
1979	1,019,932	1.03	1,054,564	34,632	5,188	15.0
1980	1,002,134	1.06	1,062,035	59,901	7,006	11.7
1981	1,002,194	1.10	1,098,179	95,985	9,163	9.5
1982	944,614	1.15	1,084,851	140,237	11,487	8.2
1983	895,700	1.23	1,102,352	206,652	14,612	7.1
1984	1,024,228	1.37	1,404,703	380,475	21,496	5.6
1985	1,173,448	1.65	1,934,626	761,178	34,263	4.5
1986	1,011,178	2.35	2,371,924	1,360,746	54,203	4.0
1987	496,200	5.39	2,674,852	2,178,652	101,540	4.7
Total	10,221,194		15,452,754	5,231,560	147,530	2.8

Table 10. Clark LDF model (Log-Logistic)

Source: Own elaboration

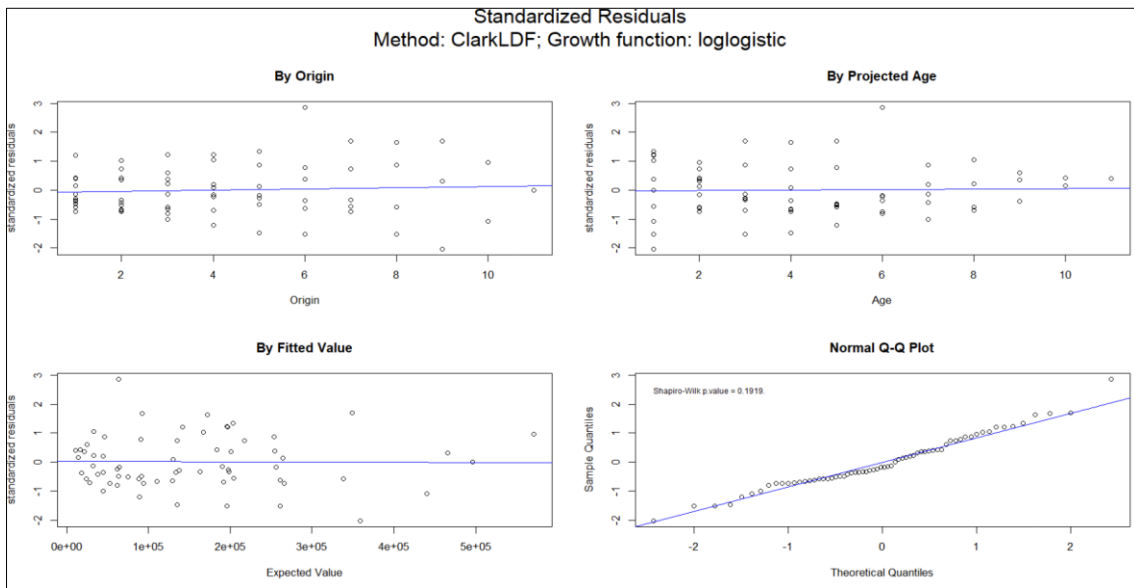The graphs of the estimated model are presented below:

Figure 8. Standardized residuals method: CLarkLDF; Growth function: Log-logistic

Source: Own elaboration

According to the above results, the future reserve value is 5.23 million, thus obtaining a total of 15.45 million, and the coefficient of variation is 2.8%. On the other hand, if we observe the graphs of the last estimated model, we can see that the standardized residuals present a behavior more in line with expectations. The maximum value of the growth function of this model is 0.9131.

CLARK CAPE COD MODEL

Finally, to finish the estimation of the reserve models, we proceed to estimate the Clark Cape Cod model, using the *ClarkCapeCod* function of the *ChainLadder* package of R. As with the estimation of the previous model, we first estimate the model with the Weibull distribution. It is important to note that, for the model to estimate future reserve values, it considers the premium and the expected loss ratio. The results are shown below:

				Clark Cape Cod model (Weibull)				
Origin	**Current Value**	**Premium**	**ELR**	**Future Growth Factor**	**Future Value**	**Ultimate Value**	**Std. Error**	**CV%**
1977	762,544	1,920,000	0.58	-	-	762,544	-	
1978	889,022	1,940,000	0.58	0.00	4,971	893,993	9,934	199.8
1979	1,019,932	1,960,000	0.58	0.01	13,665	1,033,597	16,886	123.6
1980	1,002,134	1,980,000	0.58	0.02	28,584	1,030,718	24,999	87.5
1981	1,002,194	2,000,000	0.58	0.05	53,684	1,055,878	34,891	65.0
1982	944,614	2,020,000	0.58	0.08	95,004	1,039,618	46,881	49.3
1983	895,700	2,040,000	0.58	0.14	161,368	1,057,068	61,095	37.9
1984	1,024,228	2,060,000	0.58	0.22	264,872	1,289,100	77,580	29.3
1985	1,173,448	2,080,000	0.58	0.35	420,456	1,593,904	96,537	23.0
1986	1,011,178	2,100,000	0.58	0.53	642,799	1,653,977	118,319	18.4
1987	496,200	2,120,000	0.58	0.76	935,499	1,431,699	142,135	15.2
Total	10,221,194	22,220,000			2,620,901	12,842,095	308,643	11.8

Table 11. Clark Cape Cod model (Weibull)

Source: Own elaboration

In this model there are new columns, among these columns we have the premium and the expected loss ratio (ELR), since, to estimate future values, the Clark Cape Cod model is based on this information, as mentioned in the theoretical framework section. As for the results of this model, we have a projected reserve amount of 2.62 million, which gives us an approximate total amount of claims of 12.8 million, with a variation coefficient of 11.8%. The maximum value of the growth function of this model is 0.9943. The respective graphs of the standardized residuals are presented below:
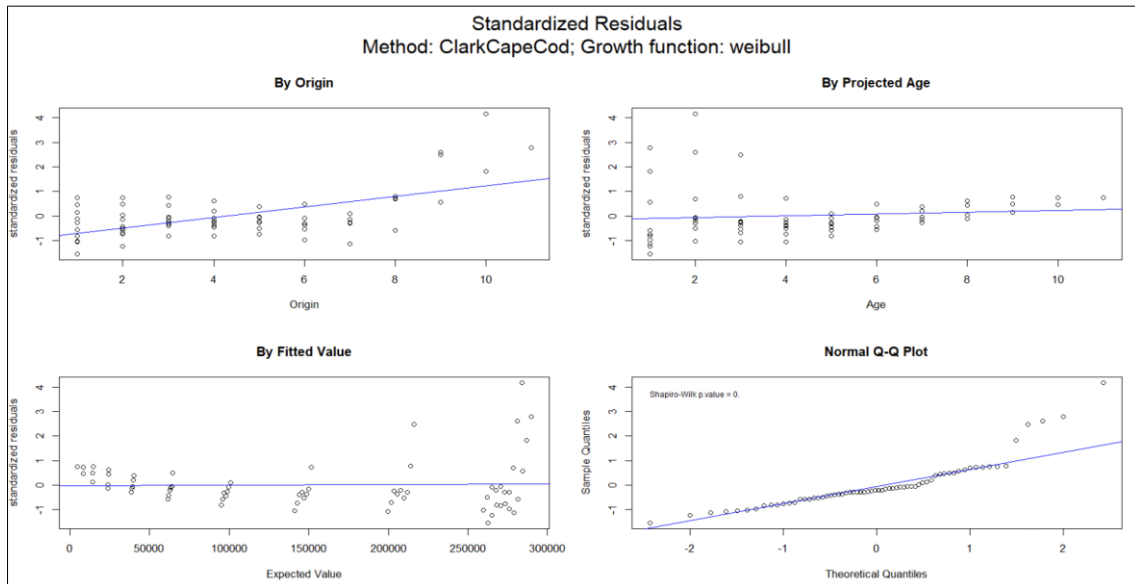
Figure 9. Standardized residuals method: CLarkCapeCod; Growth function: Weibull

Source: Own elaboration

To conclude this section, we estimate the Clark Cape Cod model, now with the Log-logistic distribution.

	Clark Cape Cod model (Log-logistic)							
Origin	Current Value	Premium	ELR	Future Growth Factor	Future Value	Ultimate Value	Std. Error	CV%
1977	762,544	1,920,000	0.63	-	-	762,544	-	
1978	889,022	1,940,000	0.63	0.01	12,988	902,010	14,740	113.5
1979	1,019,932	1,960,000	0.63	0.02	29,810	1,049,742	22,614	75.9
1980	1,002,134	1,980,000	0.63	0.04	52,027	1,054,161	30,283	58.2
1981	1,002,194	2,000,000	0.63	0.07	82,069	1,084,263	38,574	47.0
1982	944,614	2,020,000	0.63	0.10	123,857	1,068,471	48,049	38.8
1983	895,700	2,040,000	0.63	0.14	183,972	1,079,672	59,276	32.2
1984	1,024,228	2,060,000	0.63	0.21	273,855	1,298,083	72,886	26.6
1985	1,173,448	2,080,000	0.63	0.32	413,644	1,587,092	89,564	21.7
1986	1,011,178	2,100,000	0.63	0.48	635,930	1,647,108	110,012	17.3
1987	496,200	2,120,000	0.63	0.73	966,360	1,462,560	134,194	13.9
Total	10,221,194	22,220,000			2,774,511	12,995,705	295,056	10.6

Table 12. Clark Cape Cod model (Log-logistic)

Source: Own elaboration

In **Table 12**, we observe that the future value of reserves is 2.78 million, which means an approximate total amount of claims of 13 million, with a standard error amount of 295 thousand, which means a coefficient of variation of 10.6%. The maximum value of the growth function of this model is 0.9323.

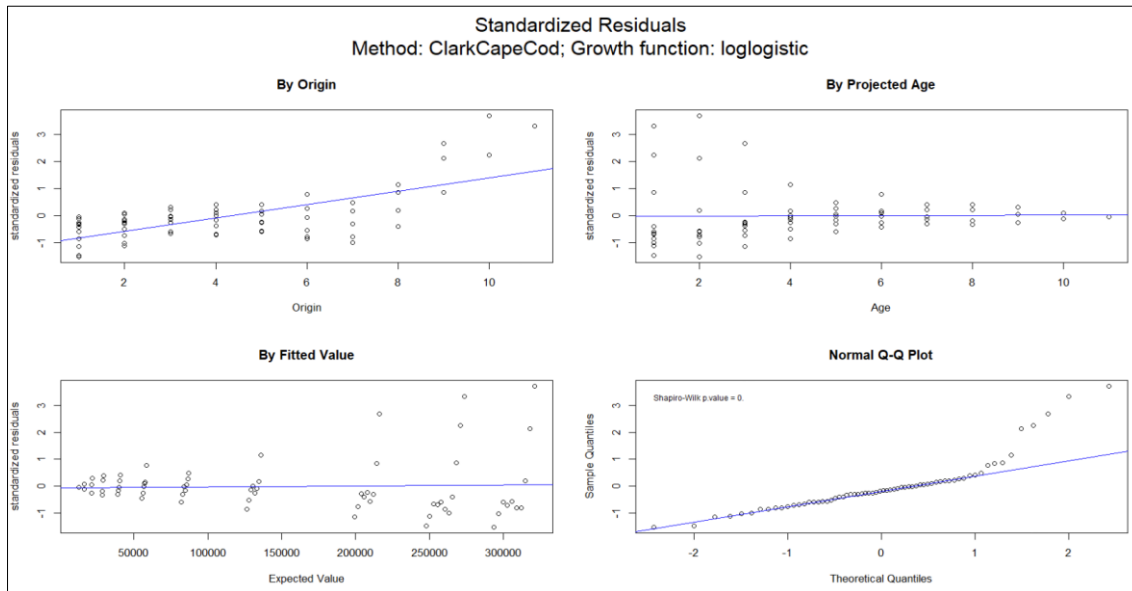Below are the graphs of the standardized residuals of the estimated model:

Figure 10. Standardized residuals method: CLarkCapeCod; Growth function: Log-logistic

Source: Own elaboration

The above graphs show that there is a slight upward trend in the graph of standardized waste and period of origin.

4. SHINY APPLICATION

To visualize the results of the models developed in R, an application has been developed in Shiny, which will be presented below.

Shiny is an R package that allows to easily create interactive web apps for anyone can use it (Wickham, 2021), it allows the development of interactive data visualizations and analysis.

The structure of the Shiny application in R is divided into 2 main parts, the UI part and the server part. Regarding the UI section, this one is in charge of the interface, that is to say, in this section we design how we want the app to be visualized for the interaction with the user, while in the server part, in this section all the logic and the computational part of the application is programmed. These two sections are connected by the function *shinyApp* so that the application can be executed.

The application consists of 4 tabs: Data, Exploratory Data Analysis (EDA), Models, Comparative results

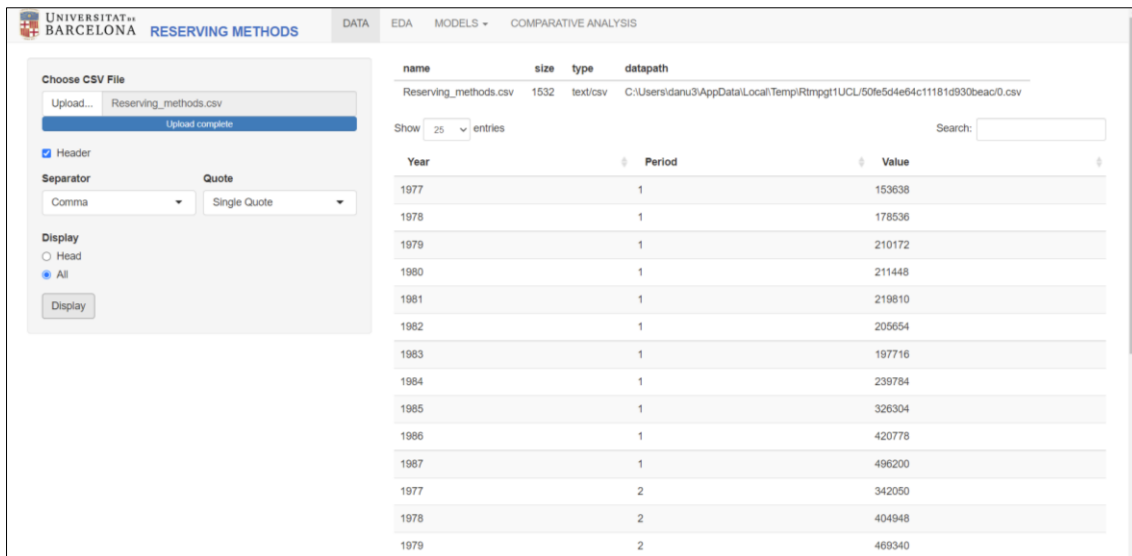In the Data tab, we can load our data and visualize if the loading has been correct.

Figure 11. Data

Source: Own elaboration

In the Exploratory Data Analysis (EDA) tab, we will be able to visualize 5 graphs of our data. The purpose of these graphs is to present the loaded data in a visual way.
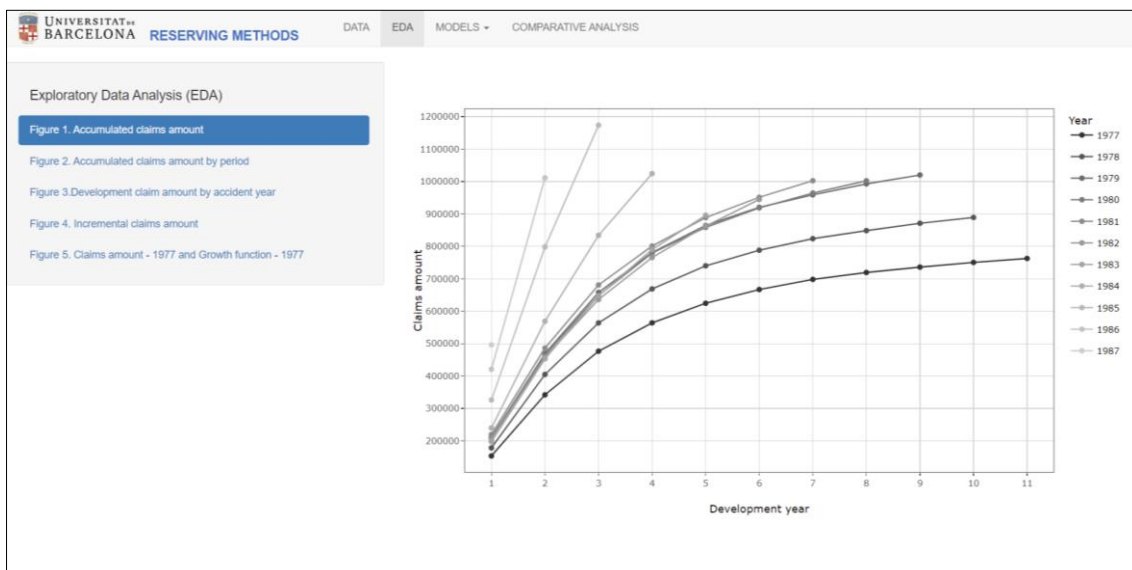

Figure 12. EDA

Source: Own elaboration

In the next tab Models, you can choose the model of which you want to visualize the results, the models to choose are Chain Ladder, Mack Chain Ladder, Clark LDF and Clark Cape Cod.
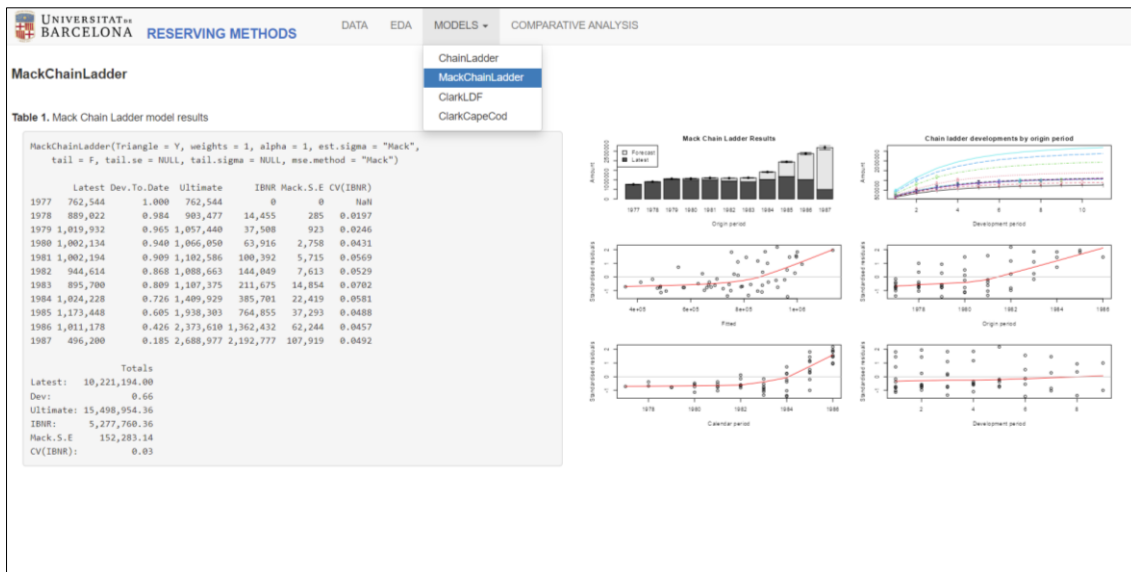
Figure 13. Models – Mack Chain Ladder

Source: Own elaboration

For the Clark LDF and Clark Cape Cod models, two models are calculated, since the first model is calculated with the Log-logistic distribution and the second with the Weibull distribution.
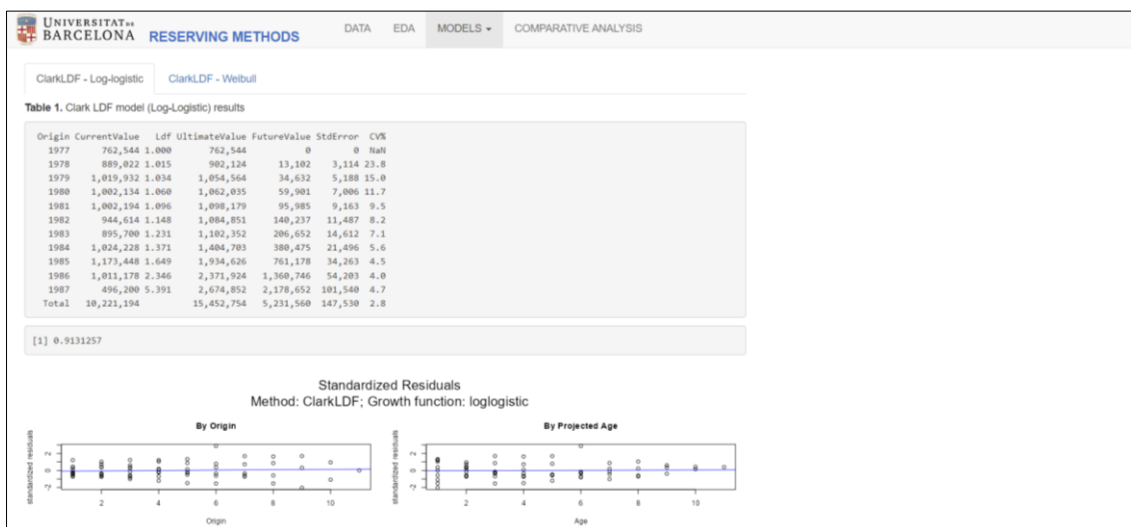

Figure 14. Models – Clark LDF

Source: Own elaboration

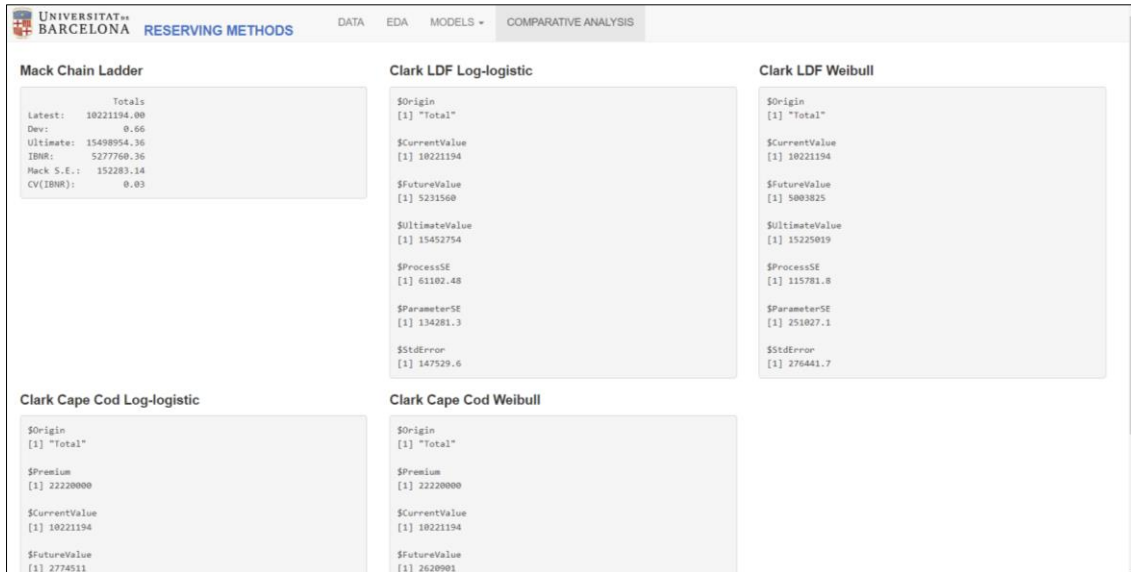Finally, the last tab shows the total results of all the models used.

Figure 15. Comparative analysis

Source: Own elaboration

The application can be viewed at the following link:

https://raidanielriverarubio.shinyapps.io/Reserving_methods/

5. COMPARISON OF RESULTS

Once we have developed all the models that we had proposed at the beginning, we proceed to present a summary table of the most relevant of these models, so that we can compare them and see an appropriate model that fits our values, in order to make a decision about what future values amount would be the most appropriate to reserve.

	Mack Chain Ladder	**Clark LDF (Weibull)**	**Clark LDF (Log-logistic)**	**Clark Cape Cod (Weibull)**	**Clark Cape Cod (Log-logistic)**
Current Value	10,221,194	10,221,194	10,221,194	10,221,194	10,221,194
Future Value	5,277,760	5,003,825	5,231,560	2,620,901	2,774,511
Ultimate Value	15,498,954	15,225,019	15,452,754	12,842,095	12,995,705
Dev.To. Date	0.660	0.671	0.661	0.796	0.787
Std. Error	152,283	276,442	147,530	308,643	295,056
CV%	3.0	5.5	2.8	11.8	10.6

Table 13. Models resume

Source: Own elaboration

Table 13 shows the results of the estimated models, from which we can mention some relevant points: the Clark Cape Cod models are those with the lowest reserve amount, this may be due to several factors, one of which is that their estimation method differs slightly

from the rest, since it considers a premium amount that varies over time and a level of loss ratio that is maintained in the periods of analysis (ELR), therefore, the amount of reserves will depend directly on the amount of premium expected in the following periods. In addition, the growth function of the Log-logistic model presented a value of 0.91 for the Clark LDF model, while with the Weibull distribution it was 0.99, therefore, in the case of the model with Weibull distribution, the reserves pending for the development periods that exceed 11 periods will be almost nil, while for the model with Log-logistic distribution, these reserves will be greater, since this distribution presents a longer tail.

In addition, with respect to the Clark Cape Cod models, it should also be noted that they are the models with the highest level of coefficient of variation and the highest standard errors amount

Another point to highlight is that the models with Weibull distribution were not the best estimated models, based on the fact that the coefficient and standard errors were lower with the models that used the Log-logistic distribution. At this point, it is worth noting that, previously, we had verified that our data fit better to the Weibull distribution when we want to model the growth curve, this based on the information criteria (AIC and BIC). However, as we have mentioned, the information criteria are not an absolute truth, it is also important to take into account the estimation results and to understand the reasons for the possible disadvantages of estimations with one distribution or another. In this case, the Weibull distributions fit our data because they present a maximum value in the growth function of 0.99, however, they present greater variability in the estimates.

Among the models with the lowest coefficient of variation are the Clark LDF model with Log-logistic distribution and the Chain Ladder model, since the first one has a coefficient of variation of 2.8%, and the second one a coefficient of 3.0%. In addition, their standard errors are 147.5 thousand and 152.3 thousand, respectively. The difference in the projected reserves amount of these two models is 273.9 thousand, being the Mack Chain Ladder model the one with higher reserves.

6. CONCLUSIONS

To conclude this analysis, based on all that has been explained in the previous sections, we can conclude that the projection of the reserves amount is very important for insurance companies, since, apart from having an impact on their income statement, it will also allow them to be able to face the claims that will be reported in the future, in other words, the solvency of an insurance company depends a lot on its estimated reserves. Therefore, it is of utmost importance to employ an efficient and effective model, so that the reserves amount is reliable.

As we have mentioned, there are several models that have been developed in the last decades, from deterministic models to stochastic and more sophisticated models. However, it is also important to take into account that the model must also be easy to use and simple to explain, for this reason, the deterministic Chain Ladder model has been used in the last decades. However, the objective of this analysis was to use a model that, in a certain way, can also be quite easy to implement, thanks to technological progress, since the model is programmed in free software (R). And, in addition, the results obtained are not very complex to understand. Therefore, in view of all these factors, we can conclude that the Clark LDF model with Log-logistic distribution has been the model with the highest accuracy and the lowest variability with respect to the rest of the models

for our database, including Chain Ladder; however, it must be taken into account, as mentioned by the author in his article, that this model has a disadvantage, and that is the over parameterization of the model.

7. BIBLIOGRAPHY

Amin, Z., Antonio, K., Beirlant, J., Castro, C., Dean, G., Frees, E., Gan, G., Gao, L., Garrido, J., Hua, L., Ismail, N., Kim, J., Okine, N., Shyamalkumar, N., Su, J., Tan, C., Verdonck, T., Viswanathan, K. (2018). *Loss Data Analytics: An open text authored by the Actuarial Community*. Retrieved from https://openacttexts.github.io/Loss-Data-Analytics/

Clark, D. (2003). LDF curve-fitting and stochastic reserving: a maximum likelihood approach. *CAS Forum*, *3, 4*, 41-92. Retrieved from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=b890889afe9cb578fdc1d92618f5648c22aabdda#page=45

England, P., Verrall, R. (2002). Stochastic Claims Reserving in General Insurance. *British Actuarial Journal*, 8, 443-518.

Gesmann, M., Murphy, D., Zhang, Y., Carrato, A., Crupi, G., Dutang, C., Lacoume, A., Charpentier, A., Wüthrich, M., Concina, F., Dal Moro, E., Krvavych, Y., Goulet, v., De Virgilis, M., Spina, M. (2023). *ChainLadder: Statistical methods and models for claims reserving in general insurance*. R package v. 0.2.17. Retrieved from http://cran.r-project.org/package=ChainLadder

Mack, T. (1993). Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *ASTIN Bulletin: The Journal of the IAA*, 23(2), 213-225. doi:https://doi.org/10.2143/AST.23.2.2005092

R Development Core Team. (2023). *R: a language and environment for statistical computing*. Vienna, Austria. Retrieved from http://www.R-project.org/

Raço, E., Haxhi, K., Llagami, E., Zaçaj, O. (2022). Comparison of Statistical Methods for Claims Reserve Estimation Using R Language. *WSEAS Transactions on Mathematics*, *21*, 547-552. doi:10.37394/23206.2022.21.61

Schmidt, K., Zocher, M. (2008). The Bornhuetter-Ferguson Principle. *Casualty Actuarial Society*, 2(1), 85-110. doi:https://doi.org/10.1007/978-3-319-30056-6_4

Wickham, H. (2021). Mastering Shiny: Build Interactive Apps, Reports, and Dashboards Powered by R. United States of America.: *O'Reilly Media, Inc.*

8. ANNEX

8.1.R CODE

```
#### TFM ####

library(ChainLadder)

library(dplyr)

library(tidyverse)

library(reshape2)

library(univariateML)

#### Data & graphs ####

data("ABC")

force(ABC)

Y <- force(ABC)

Y <- as.triangle(Y)

Y2<- cum2incr(Y)

Y2<-as.matrix(Y2)

Y2 <- as.triangle(Y2)

# Graficos a presentar

plot(Y, main = "Accumulated claims amount", xlab = "Development year", ylab =
"Claims amount")

plot(Y, lattice = TRUE, main = "Accumulated claims amount by period", xlab =
"Development year", ylab = "Claims amount")

plot(Y2, main = "Incremental claims amount", xlab = "Development year", ylab =
"Claims amount")

barplot(Y2,

    legend.text                                                              =
c("1977","1978","1979","1980","1981","1982","1983","1984","1985","1986","1987"),

    main = "Incremental claims amount by period", xlab = "Development year", ylab =
"Claims amount"

    )

# Graficos incrementales

plot(Y2)
```

```
plot(Y2, lattice = TRUE)

#### Growth function ####

v1 <- Y2[1,]

GF <- cumsum(v1) / sum(v1)

plot(GF, main = "Growth function", xlab = "Development Year")

lines(GF, col = "blue")

#### Distribution adjustment  ####

barplot(v1, main = "Claims amount - 1977", xlab = "Development Year", ylab = "Claims
amount")

barplot(GF, main = "Growth function - 1977", xlab = "Development Year", ylab =
"Cumulative percentage of claims amount")

lines(GF, col = "blue")

comparacion_aic <- AIC(

  mlllogis(GF),

  mlweibull(GF)

)

AIC <- comparacion_aic %>% rownames_to_column(var = "Distribution") %>%
arrange(AIC)

comparacion_bic <- BIC(

  mlllogis(GF),

  mlweibull(GF)

)

BIC <- comparacion_bic %>% rownames_to_column(var = "Distribution") %>%
arrange(BIC)

Comparison <- cbind(AIC, BIC)

#### Reserving modelsL####

#### Chain Ladder deterministic model ####

n <- ncol(Y)

f <- sapply(1:(n-1),

        function(i){

          sum(Y[c(1:(n-i)),i+1])/sum(Y[c(1:(n-i)),i])
```

```r
        }
)
CT<- cbind(Y)
ChainLadder<- function(mat, factor) {
 n<- ncol(mat)
 for(k in 1:(ncol(mat)-1)) {
   mat[(n-k+1):n, k+1] <- mat[(n-k+1):n,k]*factor[k]}
 return(mat)
}
Model1<-ChainLadder(CT,f); Model1
Current_Value <- CT[cbind(1:nrow(CT), ncol(CT):1)]; Current_Value
Current_Value <-as.vector(Current_Value); Current_Value
Ultimate_Value <- Model1[,11]
Ultimate_Value <- as.vector(Ultimate_Value); Ultimate_Value
Future_Value <- as.vector(round(Ultimate_Value - Current_Value, 2) )  ; Future_Value
TotalCV <- round(sum(Current_Value),2); TotalCV
TotalFV <- round(sum(Future_Value),2); TotalFV
TotalUV <- round(sum(Ultimate_Value),2); TotalUV
Years <- 1977:1987; Years
CLV <- cbind(Years, Current_Value,Future_Value, Ultimate_Value); CLV
CLV <- as.data.frame(CLV); CLV
Total_CV <- as.vector(TotalCV); Total_CV
Total_FV <- as.vector(TotalFV); Total_FV
Total_UV <- as.vector(TotalUV); Total_UV
Total_CV<-format(Total_CV, big.mark = ",", scientific = FALSE)
Total_FV<-format(Total_FV, big.mark = ",", scientific = FALSE)
Total_UV<-format(Total_UV, big.mark = ",", scientific = FALSE)
Totales <- cbind(Total_CV, Total_FV, Total_UV); Totales
CLV; Totales
```

```
#### Mack Chain Ladder model ####

#MackChainLadder(Y, weights = 1, alpha=1, est.sigma="log-linear",

#           tail=T, tail.se=NULL, tail.sigma=NULL, mse.method="Mack")


MCL <- MackChainLadder(Y, weights = 1, alpha=1, est.sigma="Mack",

              tail=F, tail.se=NULL, tail.sigma=NULL, mse.method="Mack")

MCL

plot(MCL)

#### Clarck LDF model ####

### Loglogistic LDF ###

cldfll <- ClarkLDF(Y, cumulative = TRUE, maxage = 11,

           adol = FALSE, adol.age = NULL, origin.width = NULL,

           G = "loglogistic")

cldfll

plot(cldfll)

### Weibull LDF ###

cldfw <- ClarkLDF(Y, cumulative = TRUE, maxage = 11,

           adol = FALSE, adol.age = NULL, origin.width = NULL,

           G = "weibull")

cldfw

plot(cldfw)

#### Clarck Cape Cod model ####

### Loglogistic Cape Cod ###

cccll <-  ClarkCapeCod(Y,Premium=1900000+20000*1:11,  cumulative  =  TRUE,
maxage = 11,

            adol = FALSE, adol.age = NULL, origin.width = NULL,

            G = "loglogistic")

cccll

plot(cccll)

### Weibull Cape Cod ###
```

```r
cccw <- ClarkCapeCod(Y,Premium=1900000+20000*1:11, cumulative = TRUE, maxage = 11,

                 adol = FALSE, adol.age = NULL, origin.width = NULL,

                 G = "weibull")

cccw

plot(cccw)
```

8.2.SHINNY APP

```r
#Library

library(ChainLadder)

library(RColorBrewer)

library(ggplot2)

library(plotly)

library(readr)

library(tidyr)

library(shiny)

library(xtable)

library(paletteer)

# Define UI

ui <- navbarPage(

 tags$style(type='text/css', '#txt_out {white-space: pre-wrap;}'),

   title=span(img(src = "Logo.png", align = 'center',height = 60,style="margin-top: -14px;padding-right:10px;

   padding-bottom:10px"), strong(span("RESERVING METHODS", style = "color:#3A6FC0"))),

   tabPanel("DATA",

   sidebarPanel(

     fileInput("upload","Choose CSV File",

     buttonLabel = "Upload...",
```

```r
      accept=c('text/csv',

      'text/comma-separated-values,text/plain',

      '.csv','xlsx'),multiple = FALSE),

    checkboxInput('header', 'Header', TRUE),

    fluidRow(

      column(6,style=list("padding-right: 5px;"),

          selectInput('sep', 'Separator',

           choices=c(Comma=",",

           Semicolon=';',

           Tabulation='\t'),selected = ",")),

      column(6,style=list("padding-left: 5px;"),

          selectInput("quote", "Quote",

          choices = c("Double" = ","Double Quote" = "","Single Quote" = "'"),selected =
"'"))
    ),

        radioButtons("disp", "Display",

            choices = c(Head = "head",

                    All = "all"),

            selected = "all"),

      actionButton("Display", "Display"),

    ),

      mainPanel(

        tableOutput("files"),

        dataTableOutput("contents")

        )),

    tabPanel("EDA",

    navlistPanel(
```

"Exploratory Data Analysis (EDA)",

tabPanel("Figure 1. Accumulated claims amount",plotlyOutput("Figure1",width = "100%", height = "630px")),

tabPanel("Figure 2. Accumulated claims amount by period",plotlyOutput("Figure2",width = "100%", height = "630px")),

tabPanel("Figure 3.Development claim amount by accident year",plotlyOutput("Figure3",width = "100%", height = "450px")),

tabPanel("Figure 4. Incremental claims amount",plotlyOutput("Figure4"),width = "100%", height = "450px"),

tabPanel("Figure 5. Claims amount - 1977 and Growth function - 1977",h4(strong("Claims amount - 1977")),plotlyOutput("Figure5"),h4(strong("Growth function - 1977")),plotlyOutput("Figure6")),

)),

navbarMenu("MODELS",

tabPanel("ChainLadder",h4(strong("ChainLadder")),br(),h5(strong("Table 1."), ("Accumulated amount by origin year i and development year j")),tableOutput("Model1"),h5(strong("Table 2."), ("Chain Ladder model results")),br(),tableOutput("Model1A"),tableOutput("Model1B")),
tabPanel("MackChainLadder",h4(strong("MackChainLadder")),br(),h5(strong("Table 1."),("Mack Chain Ladder model results")),

column(6,verbatimTextOutput("Model2")),

column(6,plotOutput("Model2A"))),

tabPanel("ClarkLDF",

mainPanel(

tabsetPanel(type = "tabs",

tabPanel("ClarkLDF - Log-logistic",h5(strong("Table 1."),("Clark LDF model (Log-Logistic) results")),verbatimTextOutput("Model3"),verbatimTextOutput("Model3A"),plotOutput("Model3B")),

tabPanel("ClarkLDF - Weibull",h5(strong("Table 1."),("Clark LDF model (Weibull) results")),verbatimTextOutput("Model3C"),verbatimTextOutput("Model3D"),plotOutput("Model3E"))

```r
            ))),

        tabPanel("ClarkCapeCod",

        mainPanel(

        tabsetPanel(type = "tabs",

        tabPanel("ClarkCapeCod - Log-logistic",h5(strong("Table 1."),("Clark Cape
Cod model (Log-logistic)
results")),verbatimTextOutput("Model4"),verbatimTextOutput("Model4A"),plotOutput(
"Model4B")),

        tabPanel("ClarkCapeCod - Weibull",h5(strong("Table 1."),("Clark Cape Cod
model (Weibull)
results")),verbatimTextOutput("Model4C"),verbatimTextOutput("Model4D"),plotOutpu
t("Model4E")))))),

    tabPanel("COMPARATIVE ANALYSIS",

        column(4,h4(strong("Mack Chain Ladder")),verbatimTextOutput("RMCL")),

        column(4,h4(strong("Clark LDF Log-
logistic")),verbatimTextOutput("RCLOG")),

        column(4,h4(strong("Clark LDF Weibull")),verbatimTextOutput("RCWEI")),

        column(4,h4(strong("Clark Cape Cod Log-
logistic")),verbatimTextOutput("RCCLOG")),

        column(4,h4(strong("Clark Cape Cod
Weibull")),verbatimTextOutput("RCCWEI")))

)
# Define server logic required to draw

server <- function(input, output) {

  output$files <- renderTable(input$upload)

  mydata<-reactive({

    req(input$Display,input$upload,input$header,input$sep,
input$quote,file.exists(input$upload$datapath))

  tryCatch(

  {
```

```r
    df <- read.csv(input$upload$datapath,

    header = input$header,

    sep = input$sep,

    quote = input$quote)

        },

      error = function(e) {

      stop(safeError(e))

    }

    )

    if(input$disp == "head") {

      return(head(df))

    }

    else {

      return(df)

    }

  })

  output$contents <- renderDataTable({

    req(mydata())

    mydata()

  })

  output$Figure1<- renderPlotly({

   Data <- mydata()

    Data <- Data[!is.na(Data$Value),]

    Figure1 <- Data%>%

    ggplot(aes(Period,Value,group = Year,colour=factor(Year)))+

    geom_point()+

    geom_line()+
```

47

```
scale_x_continuous(n.breaks = 11)+

scale_y_continuous(n.breaks = 10)+

labs(x= "Development year", y="Claims amount",color="Year")+

theme_bw()+

theme(plot.margin=margin(25,25,25,25),text = element_text(size =
11),axis.text=element_text(size=9),legend.text=element_text(size=9),panel.grid.major =
element_line(color='lightgrey'))+

scale_color_grey()

Figure1<-ggplotly(Figure1)%>% layout(dragmode = "pan")

})
output$Figure2 <- renderPlotly({

Data <- mydata()

Data <- Data[!is.na(Data$Value),]

Figure2 <- Data%>%

ggplot(aes(Period,Value,group = Year))+

geom_point(alpha=0.30)+

geom_line(size=0.24)+

facet_wrap(~Year,ncol = 4)+

scale_x_continuous(n.breaks = 6)+

scale_y_continuous(n.breaks = 6)+

theme_bw()+

theme(text = element_text(size
=10),axis.text=element_text(size=8),legend.text=element_text(size=8),

panel.grid.major = element_line(color='lightgrey'),

strip.background=element_rect(fill="#EEEEEE"),

strip.text.x = element_text(size = 10,

color = "black",

face = "bold",
```

```r
                          margin = unit(c(4,10,4,10), "points")),

        strip.text.y = element_text(size = 10,

                          color = "black",

                          face = "bold",

                          margin = unit(c(4,10,4,10), "points")))+

    labs(x= "Development year", y="Claims amount",color="Year")+

    scale_colour_grey(start = 0.6, end = 0)

    Figure2<-ggplotly(Figure2)%>% layout(dragmode = "pan")

    })

output$Figure3 <- renderPlotly({

      Data <- mydata()

    Data <- Data[!is.na(Data$Value),]

    Figure3 <- Data%>%

    ggplot(aes(Period,Value,fill = Year))+ geom_bar(stat = "identity")+

    scale_x_continuous(n.breaks = 11)+theme_bw()+ scale_colour_grey(start = 0.6, end
= 0)+

      labs(x= "Development year", y="Claims amount",color="Year")

    Figure3<-ggplotly(Figure3)%>% layout(dragmode = "pan")

})

output$Figure4 <- renderPlotly({

    Data <- mydata()

  Y  <- as.triangle(Data, origin = "Year", dev = "Period", "Value")

  Y2 <- cum2incr(Y)

  Y2 <- as.data.frame(Y2)

  Y2 <- Y2[!is.na(Y2[["value"]]),]

  Figure4 <- Y2%>%

    ggplot(aes(Period,value,group = Year,colour=factor(Year)))+
```

```r
  geom_point()+

  geom_line()+

  scale_x_continuous(n.breaks = 11)+

  scale_y_continuous(n.breaks = 10)+

  labs(x= "Development year", y="Claims amount",color="Year")+

  theme_bw()+

  theme(plot.margin=margin(25,25,25,25),text = element_text(size =
11),axis.text=element_text(size=9),legend.text=element_text(size=9),panel.grid.major =
element_line(color='lightgrey'))+

  scale_color_grey()

 Figure4<-ggplotly(Figure4)%>% layout(dragmode = "pan")

})

output$Figure5 <- renderPlotly({

  Data <- mydata()

 Y  <- as.triangle(Data, origin = "Year", dev = "Period", "Value")

 Y2 <- cum2incr(Y)

 n <- 1:ncol(Y2)

 V1 <- cbind(n,Y2[1,])%>%

  as.data.frame()%>%

  rename(Year=n,Amount=V2)

 Figure5 <- ggplot(V1,aes(x=Year,y=Amount))+

  geom_bar(stat = "identity")+

  scale_x_continuous(n.breaks = ncol(Y2))+

  scale_y_continuous(n.breaks = (ncol(Y2)-1))+

  labs(x= "Development year", y="Claims amount",color="Year")+

  theme_bw()+

  theme(plot.margin=margin(25,25,25,25),text = element_text(size =
11),axis.text=element_text(size=9),legend.text=element_text(size=9),panel.grid.major =
element_line(color='lightgrey'))+
```

```
    scale_color_grey()

  Figure5<-ggplotly(Figure5)%>% layout(dragmode = "pan")

})

output$Figure6 <- renderPlotly({

   Data <- mydata()

  Y  <- as.triangle(Data, origin = "Year", dev = "Period", "Value")

  Y2 <- cum2incr(Y)

  n <- 1:ncol(Y2)

  V1 <- cbind(n,Y2[1,])%>%

    as.data.frame()%>%

    rename(Year=n,Amount=V2)

  GF <- cumsum(V1$Amount) / sum(V1$Amount)

  GF2 <- cbind(n,GF)%>%

    as.data.frame()%>%

    rename(Year=n,Percentage=GF)

  Figure6 <- ggplot(GF2,aes(x=Year,y=Percentage))+

    geom_bar(stat = "identity")+geom_line(aes(y=Percentage),size=0.5, color="blue")+

    scale_x_continuous(n.breaks = ncol(Y2))+

    scale_y_continuous(n.breaks = (ncol(Y2)-1))+

    labs(x= "Development year", y="Cumulative percentage of claims
amount",color="Year")+

    theme_bw()+

    theme(plot.margin=margin(25,25,25,25),text = element_text(size =
11),axis.text=element_text(size=9),legend.text=element_text(size=9),panel.grid.major =
element_line(color='lightgrey'))+

    scale_color_grey()

  Figure6<-ggplotly(Figure6)%>% layout(dragmode = "pan")
```

```r
})
output$Model1 <- renderTable({

  Data2 <- mydata()

 Y <- as.triangle(Data2, origin = "Year", dev = "Period", "Value")

 n <- ncol(Y)

 f <- sapply(1:(n-1),

        function(i){

          sum(Y[c(1:(n-i)),i+1])/sum(Y[c(1:(n-i)),i])

        }

 )

  CT<- cbind(Y)

 ChainLadder<- function(mat, factor) {

  n<- ncol(mat)

  for(k in 1:(ncol(mat)-1)) {

    mat[(n-k+1):n, k+1] <- mat[(n-k+1):n,k]*factor[k]}

  return(mat)

 }

 CT

 Model1<-ChainLadder(CT,f)

 })
output$Model1A <- renderTable({

 Data2 <- mydata()

 Y <- as.triangle(Data2, origin = "Year", dev = "Period", "Value")

 n <- ncol(Y)

 f <- sapply(1:(n-1),

        function(i){

          sum(Y[c(1:(n-i)),i+1])/sum(Y[c(1:(n-i)),i])
```

```
            }

  )

    CT<- cbind(Y)

   ChainLadder<- function(mat, factor) {

    n<- ncol(mat)

    for(k in 1:(ncol(mat)-1)) {

      mat[(n-k+1):n, k+1] <- mat[(n-k+1):n,k]*factor[k]}

     return(mat)

   }

   CT

    Model1<-ChainLadder(CT,f)

    Current_Value <- CT[cbind(1:nrow(CT), ncol(CT):1)]; Current_Value

    Current_Value <-as.vector(Current_Value); Current_Value

    Ultimate_Value <- Model1[,11]

    Ultimate_Value <- as.vector(Ultimate_Value); Ultimate_Value

    Future_Value <- as.vector(round(Ultimate_Value - Current_Value, 2) )  ;
  Future_Value

    TotalCV <- round(sum(Current_Value),2); TotalCV

    TotalFV <- round(sum(Future_Value),2); TotalFV

    TotalUV <- round(sum(Ultimate_Value),2); TotalUV

    Years <- 1977:1987; Years

    CLV <- cbind(Years, Current_Value,Future_Value, Ultimate_Value); CLV

    CLV <- as.data.frame(CLV); CLV

  })

  output$Model1B <- renderTable({

   Data2 <- mydata()

   Y <- as.triangle(Data2, origin = "Year", dev = "Period", "Value")
```

```r
n <- ncol(Y)

f <- sapply(1:(n-1),

        function(i){

          sum(Y[c(1:(n-i)),i+1])/sum(Y[c(1:(n-i)),i])

        }

)

  CT<- cbind(Y)

ChainLadder<- function(mat, factor) {

 n<- ncol(mat)

 for(k in 1:(ncol(mat)-1)) {

   mat[(n-k+1):n, k+1] <- mat[(n-k+1):n,k]*factor[k]}

 return(mat)

}

CT

Model1<-ChainLadder(CT,f)

Current_Value <- CT[cbind(1:nrow(CT), ncol(CT):1)]; Current_Value

Current_Value <-as.vector(Current_Value); Current_Value

Ultimate_Value <- Model1[,11]

Ultimate_Value <- as.vector(Ultimate_Value); Ultimate_Value

Future_Value <- as.vector(round(Ultimate_Value - Current_Value, 2) )  ;
Future_Value

TotalCV <- round(sum(Current_Value),2); TotalCV

TotalFV <- round(sum(Future_Value),2); TotalFV

TotalUV <- round(sum(Ultimate_Value),2); TotalUV

Years <- 1977:1987; Years

CLV <- cbind(Years, Current_Value,Future_Value, Ultimate_Value); CLV

CLV <- as.data.frame(CLV); CLV
```

```r
  Total_CV <- as.vector(TotalCV); Total_CV

  Total_FV <- as.vector(TotalFV); Total_FV

  Total_UV <- as.vector(TotalUV); Total_UV


  Total_CV<-format(Total_CV, big.mark = ",", scientific = FALSE)

  Total_FV<-format(Total_FV, big.mark = ",", scientific = FALSE)

  Total_UV<-format(Total_UV, big.mark = ",", scientific = FALSE)


  Totales <- cbind(Total_CV, Total_FV, Total_UV); Totales
})
output$Model2 <- renderPrint({

  Data3<-mydata()

 Y <- as.triangle(Data3, origin = "Year", dev = "Period", "Value")

 MCL <- MackChainLadder(Y, weights = 1, alpha=1, est.sigma="Mack",

            tail=F, tail.se=NULL, tail.sigma=NULL, mse.method="Mack")

 MCL

})
output$Model2A <- renderPlot({

  Data3<-mydata()

 Y <- as.triangle(Data3, origin = "Year", dev = "Period", "Value")

 MCL <- MackChainLadder(Y, weights = 1, alpha=1, est.sigma="Mack",

            tail=F, tail.se=NULL, tail.sigma=NULL, mse.method="Mack")

 plot(MCL)

 })
output$Model3 <- renderPrint({

  Data4<-mydata()

 Y <- as.triangle(Data4, origin = "Year", dev = "Period", "Value")
```

```r
  cldfll <- ClarkLDF(Y, cumulative = TRUE, maxage = 11,

          adol = FALSE, adol.age = NULL, origin.width = NULL,

          G = "loglogistic")

    cldfll

})

output$Model3A <- renderPrint({

   Data4<-mydata()

 Y <- as.triangle(Data4, origin = "Year", dev = "Period", "Value")

 cldfll <- ClarkLDF(Y, cumulative = TRUE, maxage = 11,

          adol = FALSE, adol.age = NULL, origin.width = NULL,

          G = "loglogistic")

 cldfll$GrowthFunctionMAXAGE

 })

output$Model3B <- renderPlot({

   Data4<-mydata()

 Y <- as.triangle(Data4, origin = "Year", dev = "Period", "Value")

 cldfll <- ClarkLDF(Y, cumulative = TRUE, maxage = 11,

          adol = FALSE, adol.age = NULL, origin.width = NULL,

          G = "loglogistic")

 plot(cldfll)

 })

output$Model3C <- renderPrint({

   Data5<-mydata()

 Y <- as.triangle(Data5, origin = "Year", dev = "Period", "Value")

 cldfw <- ClarkLDF(Y, cumulative = TRUE, maxage = 11,

          adol = FALSE, adol.age = NULL, origin.width = NULL,

          G = "weibull")
```

```r
    cldfw

})

output$Model3D <- renderPrint({

  Data5<-mydata()

 Y <- as.triangle(Data5, origin = "Year", dev = "Period", "Value")

 cldfw <- ClarkLDF(Y, cumulative = TRUE, maxage = 11,

            adol = FALSE, adol.age = NULL, origin.width = NULL,

            G = "weibull")

  cldfw$GrowthFunctionMAXAGE

})

output$Model3E <- renderPlot({

  Data5<-mydata()

 Y <- as.triangle(Data5, origin = "Year", dev = "Period", "Value")

 cldfw <- ClarkLDF(Y, cumulative = TRUE, maxage = 11,

            adol = FALSE, adol.age = NULL, origin.width = NULL,

            G = "weibull")

 plot(cldfw)

})

output$Model4 <- renderPrint({

  Data6<-mydata()

 Y <- as.triangle(Data6, origin = "Year", dev = "Period", "Value")

 cccll <- ClarkCapeCod(Y,Premium=1900000+20000*1:11, cumulative = TRUE,
maxage = 11,

              adol = FALSE, adol.age = NULL, origin.width = NULL,

              G = "loglogistic")

 cccll
```

```
})

output$Model4A <- renderPrint({

 Data6<-mydata()

 Y <- as.triangle(Data6, origin = "Year", dev = "Period", "Value")

 cccll <- ClarkCapeCod(Y,Premium=1900000+20000*1:11, cumulative = TRUE,
maxage = 11,

                 adol = FALSE, adol.age = NULL, origin.width = NULL,

                 G = "loglogistic")

 cccll$GrowthFunctionMAXAGE

 })

output$Model4B <- renderPlot({

   Data6<-mydata()

 Y <- as.triangle(Data6, origin = "Year", dev = "Period", "Value")

 cccll <- ClarkCapeCod(Y,Premium=1900000+20000*1:11, cumulative = TRUE,
maxage = 11,

                 adol = FALSE, adol.age = NULL, origin.width = NULL,

                 G = "loglogistic")

 plot(cccll)

 })

output$Model4C <- renderPrint({

 Data7<-mydata()

 Y <- as.triangle(Data7, origin = "Year", dev = "Period", "Value")

 cccw <- ClarkCapeCod(Y,Premium=1900000+20000*1:11, cumulative = TRUE,
maxage = 11,

                 adol = FALSE, adol.age = NULL, origin.width = NULL,

                 G = "weibull")

 cccw
```

```
})

output$Model4D <- renderPrint({

 Data7<-mydata()

 Y <- as.triangle(Data7, origin = "Year", dev = "Period", "Value")

 cccw <- ClarkCapeCod(Y,Premium=1900000+20000*1:11, cumulative = TRUE,
maxage = 11,

              adol = FALSE, adol.age = NULL, origin.width = NULL,

              G = "weibull")

 cccw$GrowthFunctionMAXAGE

 })

output$Model4E <- renderPlot({

 Data7<-mydata()

 Y <- as.triangle(Data7, origin = "Year", dev = "Period", "Value")

 cccw <- ClarkCapeCod(Y,Premium=1900000+20000*1:11, cumulative = TRUE,
maxage = 11,

              adol = FALSE, adol.age = NULL, origin.width = NULL,

              G = "weibull")

 plot(cccw)

})

output$RMCL <- renderPrint({

  Data3<-mydata()

 Y <- as.triangle(Data3, origin = "Year", dev = "Period", "Value")

 MCL <- MackChainLadder(Y, weights = 1, alpha=1, est.sigma="Mack",

           tail=F, tail.se=NULL, tail.sigma=NULL, mse.method="Mack")

 round(summary(MCL)$Totals, 2)

 })

output$RCLOG <- renderPrint({

 Data4<-mydata()
```

```
  Y <- as.triangle(Data4, origin = "Year", dev = "Period", "Value")

  cldfll <- ClarkLDF(Y, cumulative = TRUE, maxage = 11,

             adol = FALSE, adol.age = NULL, origin.width = NULL,

             G = "loglogistic")

    cldfll$Total

})

output$RCWEI <- renderPrint({

   Data5<-mydata()

 Y <- as.triangle(Data5, origin = "Year", dev = "Period", "Value")

  cldfw <- ClarkLDF(Y, cumulative = TRUE, maxage = 11,

             adol = FALSE, adol.age = NULL, origin.width = NULL,

             G = "weibull")

    cldfw$Total

})

output$RCCLOG <- renderPrint({

  Data6<-mydata()

 Y <- as.triangle(Data6, origin = "Year", dev = "Period", "Value")

  cccll <- ClarkCapeCod(Y,Premium=1900000+20000*1:11, cumulative = TRUE,
maxage = 11,

               adol = FALSE, adol.age = NULL, origin.width = NULL,

               G = "loglogistic")

  cccll$Total

})

output$RCCWEI <- renderPrint({

  Data7<-mydata()

 Y <- as.triangle(Data7, origin = "Year", dev = "Period", "Value")

  cccw <- ClarkCapeCod(Y,Premium=1900000+20000*1:11, cumulative = TRUE,
maxage = 11,
```

```
                adol = FALSE, adol.age = NULL, origin.width = NULL,

                G = "weibull")

  cccw$Total

})

}

    # Run the application

shinyApp(ui = ui, server = server)
```