DISSERTATIONES
MATHEMATICAE
UNIVERSITATIS
TARTUENSIS
118

LIIVIKA TEE

Stochastic Chain-Ladder Methods in Non-Life Insurance

DISSERTATIONES MATHEMATICAE UNIVERSITATIS TARTUENSIS
118

DISSERTATIONES MATHEMATICAE UNIVERSITATIS TARTUENSIS
118

LIIVIKA TEE

Stochastic Chain-Ladder Methods in Non-Life Insurance

UNIVERSITY OF TARTU
Press

Institute of Mathematics and Statistics, Faculty of Science and Technology, University of Tartu, Tartu, Estonia

Dissertation has been accepted for the commencement of the degree of Doctor of Philosophy (PhD) in mathematical statistics on June 13, 2017, by the Council of the Institute of Mathematics and Statistics, Faculty of Science and Technology, University of Tartu.

Supervisor:

Associate Professor Meelis Käärik, PhD.
University of Tartu
Tartu, Estonia

Opponents:

Associate Professor Katrien Antonio, PhD.
KU Leuven
Leuven, Belgium

Professor Remigijus Leipus, Habil. dr.
Vilnius University
Vilnius, Lithuania

Commencement will take place on August 29, 2017, at 11.00 in J. Liivi 2-404.

Publication of this dissertation has been granted by the Estonian Doctoral School of Mathematics and Statistics.

ISSN 1024-4212
ISBN 978-9949-77-484-5 (print)
ISBN 978-9949-77-485-2 (pdf)

University of Tartu Press
www.tyk.ee

Contents

Acknowledgements

First and foremost, I offer my sincerest gratitude to my supervisor Meelis Käärik for the guidance, encouragement and advice he has provided throughout my doctoral studies. The countless discussions we held were always enlightening and provided many new insights. I have been very fortunate to have a supervisor who sacrificed his sleeping hours to work jointly on problems and who responded promptly to my never-ending list of questions.

I would like to thank all my colleagues at the Institute of Mathematics and Statistics, who have contributed in one way or another to the completion of this dissertation. Special thanks go to my office mates Anne and Riho for being great colleagues and for the long conversations on both work and life aspects. I would also like to thank Annika for the memorable moments shared at several conferences around the world. I am thankful to Natalja for the many collaborations on courses we taught together and the numerous ideas we exchanged.

I want to thank Prof. Katrien Antonio for the enriching one-year research stay at KU Leuven during my doctoral studies. Thank you for sharing your expertise and providing me with a wonderful work environment. It was a privilege to be a part of your research group. I felt very welcomed at KU Leuven and I was extremely fortunate to meet so many new colleagues who turned into friends. It was a great pleasure to share the office with Roel H, Anastasios and of course Sander, who was always there to help me out with urgent issues on no matter what. Many thanks to all my nice colleagues at KU Leuven, especially Jose, Kadri and Roel V. Thank you, Roel V, for being a great inspiration and never leaving your constructive criticism untold. You were and are an invaluable part of this PhD journey.

Thanks as well to all my remarkable friends, in particular Tanja, Maarja and Liis, for the moments of happiness and after-work activities.

Minu suurim tänu kuulub minu armsale emale, isale ja vennale. Teie julgustavate sõnade toel ei ole ükski raskus tundunud ületamatuna.

Introduction

The idea of insurance is that persons can share the risk that they are exposed to. Every insured person pays a premium for being covered by the insurer, and those who suffer an accident receive compensation. The difficulty with this idea is that the future is unknown and the insurer therefore does not know the future costs. Accurate loss reserves are hence essential for insurers to maintain adequate capital, to efficiently price their insurance products and to remain solvent at any time moment in the future.

Loss reserving for non-life insurance is usually based on models with data aggregated in a run-off triangle. In practice, there is a long tradition of actuaries calculating reserve estimates according to deterministic methods without explicit reference to a stochastic model. The most common such method is the chain-ladder method. However, stochastic models are needed in order to assess the variability of the claims reserve. In the first part of the dissertation we concentrate on the prediction error of the estimate and use bootstrapping in the claims reserving context. We discuss and implement different types of residuals and show that the choice of residuals and their adjustments have a significant effect on the prediction error. We also address model validation in order to rank the competing reserving models and the assessment is conducted on data from different lines of business.

Many variations of the chain-ladder method have been proposed in the literature. One of these extensions is the continuous chain-ladder method. This method is designed for data recorded in continuous time but can also be applied to aggregated data. We investigate the continuous chain-ladder method and interpret it also for the discrete framework. This is useful in practice since insurance companies do not always keep track of insurance data in continuous time. Furthermore, we investigate whether and how much the use of different aggregation levels of claim data can improve the reserving process using both the continuous chain-ladder and the classical chain-ladder method.

The dissertation is set out as follows. In Chapter 1, we give a broader introduction to loss reserving in insurance and motivate the research problem. Chapter 2 is devoted to stochastic reserving methods on the basis of generalized linear models in the bootstrapping framework. In Section 2.1, we present a review of generalized linear models and their application to claim reserving, while in Section 2.2, we discuss some aspects linked to the bootstrap methodology. Section 2.3 describes the empirical study, where we demonstrate the application of different stochastic reserving models based on a data set provided by an Estonian insurer. In Section 2.4, the comparative analysis for model validation with the Schedule P database is carried out. Section 2.5 concludes this chapter. Chapter 3 investigates the influence of different levels of data aggregation. In Section 3.1, we discuss, describe and clarify the CCL method. Section 3.2 specifies the measure of the prediction error applied in the following simulation and empirical studies and describes the data aggregation idea. In Sections 3.3 and 3.4, the simulation studies based on different methodologies is carried out along with the comparison of the model performances. Section 3.5 contains an empirical study with a micro-level data from an insurance company. It is followed by conclusions in Section 3.6. The main body of this dissertation is followed by an outlook, where we summarize the main findings of this dissertation and conclude our work by presenting several suggestions for future research related to this topic. In the appendix, we formulate the step-by-step algorithm for the bootstrap procedure conducted in Chapter 2.

Chapter 1

Loss reserving in insurance

1.1 Background

Insurance companies bring security to society in the form of insurance products and contracts. Insurance contract sold by the insurer to the insured promises to pay the insured or a third party for a loss covered by the contract. Unlike other goods, insurance is sold when the ultimate cost is unknown. Every non-life insurance company is obligated to compensate its policyholders for claims that meet the terms of the policy. In order to meet and administer its contractual obligations to policyholders the insurance company has to set up loss reserves. Since loss events with the number and amount of claims are random, it is important to calculate the claims reserve carefully as underestimation would lead to solvency problems and overestimation unnecessarily holds the excess capital instead of using it for other purposes. Claims reserves are the largest liabilities on the balance sheet of a non-life insurance company. The claims estimation is one of the basic yet difficult and important actuarial tasks in the insurance industry, because it gives the certainty to be solvent at any time moment in the future. The process of estimating the outstanding liabilities is called loss reserving.

Figure 1.1 illustrates the development process of a regular non-life claim. The insurance premium is paid at the beginning of the insurance period. Typically, the accident date of a non-life insurance claim and its reporting date do not coincide and the claim cannot be settled immediately. Thus, when a claim occurs at some concrete time and is reported to the insurer at some time later, then one or several transactions follow to make payments for the claim until the settlement. The time gap between occurrence date and reporting

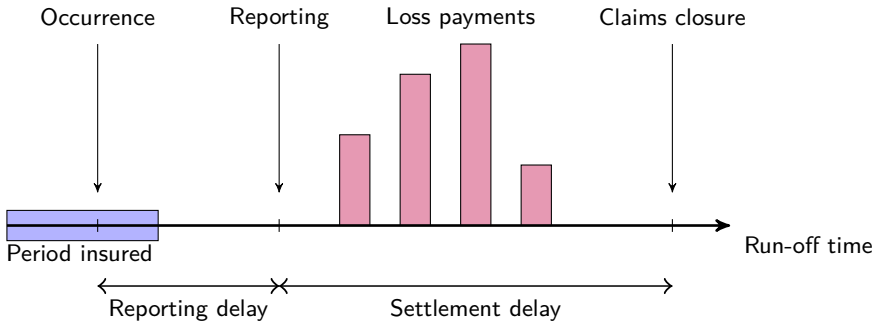

Figure 1.1: Time line representing the development of a non–life claim.

date is called the reporting delay, and the time gap between reporting date and settlement date is called the settlement delay. As the possible claims cash flows are only paid later, the insurance company should build claims reserves to be able to fulfill these outstanding loss liabilities. Insurers often distinguish between reserves for RBNS, claims that are reported to the insurer but not settled, and IBNR, claims that incurred but are not reported to the company. An IBNR claim has occurred before the present moment, but its declaration and settlement follow afterwards. This kind of distinction of claims is necessary in the prediction routine.

Accurately estimated loss reserves are extremely important for insurers for many reasons. The claims reserves should be sufficiently high to fulfill all liabilities and allow appropriate dividends to the shareholders at the same time. Moreover, the claims reserves should be best-estimate such that they can be used for pricing the future insurance contracts. In general, accurate loss reserves are essential for proper decision-making in almost every aspect of insurance practices, including underwriting, rate-making and investment. Besides the owners and management of the insurance company, many other stakeholders, such as investors, customers and regulators, are making decisions that depend on the insurer's loss reserves. See Chapter 1 of Friedland (2010) for a detailed overview for the importance of loss reserving.

There is a variety of methods for the actuary to choose amongst for reserving purposes. Many of the reserving methods mentioned below in Sections 1.2–1.3 have been applied and discussed widely in the academic literature over the past thirty years. The new solvency guidelines (Solvency II) have added a dynamic component to claims reserving entailing tremendous developments in the field. Roughly speaking, reserving models can categorized into three classes: non-stochastic macro-level models, stochastic macro-level models, and

stochastic micro-level models. A non-stochastic claims model expresses data as a deterministic algebraic combination of parameters, whereas stochastic claims model quantifies the uncertainties in deterministic claims reserving algorithms using appropriate stochastic models. Micro-level models use claims related data on an individual basis, rather than aggregating by underwriting or accident year and development year. The remainder of this chapter introduces various reserving methods and reviews their strengths and limitations.

1.2 Aggregate claims reserving

The most widely used claims reserving models are non-stochastic macro-level models which are merely deterministic algorithms using aggregate claims data. Macro-level models are often based on claims data organized in a run-off triangle. A run-off triangle, such as presented in Table 2.1, is a table showing aggregate losses for a series of loss periods at various valuation dates, reflecting the change in claims amounts and in claims payments as claims mature. The terms "claims development" and "claims run-off" are used interchangeably. For an extensive overview of techniques based on the run-off triangle see England and Verrall (2002) and Wüthrich and Merz (2008). The period considered in the run-off triangle is usually a year, but it could be also a day, a week, a month, a quarter or some other preferred amount of time. Also, the data does not have to be claim payments, even if it often is. It could also be the number of reported claims at the time.

1.2.1 Non-stochastic macro-level models

Traditional macro-level methods are dominant among practitioners in loss reserving and the chain-ladder (CL) method, being a non-stochastic macro-level model, is the most widely used loss reserving model. Besides the chain-ladder method, another commonly-used macro-level reserving method is the expected claims technique. It projects the ultimate claims based on actuaries' prior estimates rather than the claims experience. Other macro-level models, such as Bornhuetter-Ferguson method and Cape-Cod method, are constructed as a blend of the chain-ladder and the expected claims techniques (Friedland (2010)). An important assumption for the use of these techniques is that the actuaries can provide a reliable expected claims estimate. However, when the environment undergoes many rapid and complex changes, it may be questionable to assume that actuaries' expectations are reliable to reflect the

impact from the environment. There are many other non-stochastic macro-level methods, e.g., Frequency-Severity, and Berquist-Sherman techniques. See Friedland (2010) for a detailed review of the existing non-stochastic reserving methods.

The chain-ladder method relies on the simplest assumption that payments will emerge in a similar way in each accident year. The proportionate increases in the known cumulative payments from one development year to the next can then be used to calculate the expected cumulative payments for future development years. See Wüthrich and Merz (2008) for a detailed step-by-step introduction of the chain-ladder technique. In general, macro-level models are easy to understand and models can also be implemented in a single spreadsheet. Nevertheless, macro-level models have also major drawbacks. Many articles in the literature have discussed issues of the chain-ladder method and other macro-level reserving models, for instance, unstable predictions for recent accident years (Bornhuetter and Ferguson (1972)), over-parametrization due to the small sample size (Wright (1990)), difficulties in separating RBNS from IBNR claims (Schnieper (1991); Liu and Verrall (2009)), problems with the presence of zero or negative cells in run-off triangles (Verrall and Li (1993), Kunkler (2004)), and difficulties in the simultaneous use of incurred and paid claims (Quarg and Mack (2008)). Neglected detailed available information and high prediction errors were re-addressed in Charpentier and Pigeon (2016) in the context of comparing micro and micro-level models. This literature also provides adjustments to address some of the issues, but the adjustments are often suggested in a heuristic fashion and not applied simultaneously.

1.2.2 Stochastic macro-level models

The disadvantages of macro-level models discussed in Section 1.2.1 apply to both stochastic and non-stochastic macro-level models. However, as the non-stochastic methods give only a point estimate and one of the main interests in loss reserving being the likely variability of the claims reserves, stochastic macro-level models were first to assess the variability of the claims reserves. An overview of stochastic macro-level models is given by England and Verrall (2002), Wüthrich and Merz (2008) and Kaas et al. (2008), different stochastic chain-ladder methods were compared in Hess and Schmidt (2002). Stochastic claims reserving starts with constructing a model that produces the actuary's best estimate and then using this model for estimating the prediction error of the model. While working with different stochastic reserving models, the chain-ladder mean is often kept as a benchmark.

Researchers have extended the chain-ladder method with distributional assumptions regarding the underlying stochastic process generating the chain-ladder estimate. Mack (1993) was the first to model the stochastic process underlying the chain-ladder method, presenting a distribution-free formula to derive the standard errors of the chain-ladder estimates. Furthermore, the use of generalized linear models (GLM) has been a popular approach to model the stochastic process underlying the chain-ladder method. Within this group of models, Kremer (1982) introduced a standard lognormal model, Mack (1991) described a parametric model for the claims amount using the gamma distribution, Renshaw and Verrall (1998) used a quasi-likelihood approach to fit an over-dispersed Poisson model and Verrall (2000) considered a range of stochastic models including a quasi negative binomial model. The distribution-free method by Mack specifies only the first two moments whereas the last four models specify the full distribution of the incremental losses.

Another popular alternative approach to obtain inference is to use the bootstrap technique. The bootstrap technique has proved to be a very useful tool in many fields and can be particularly interesting to assess the variability of the claim reserving predictions and to construct upper limits at an adequate confidence level. When England and Verrall (1999) and England (2002) introduced bootstrapping in claims reserving, it soon became a popular method in practice as well as in the literature. Pinheiro et al. (2003) used the bootstrap technique to obtain prediction errors for different claim reserving methods, namely methods based on the CL technique and on GLMs. They extended the work of England and Verrall in many aspects, for instance, they addressed the use of proper residual definitions and discussed the consequences of the introduction of the residual corrections. Björkwall et al. (2009) relaxed the model assumption in England and Verrall (1999), England (2002) and Pinheiro et al. (2003) in order to obtain a bootstrap approach which could be used for other development factor methods than the chain-ladder, and investigated a bootstrap procedure which is based on unstandardized prediction errors and defined a parametric bootstrap approach. Björkwall et al. (2010) implemented a bootstrap procedure for the deterministic separation method, which also considers calendar year effects. Given the possible sensitivity of standard procedures to the possible presence of outliers, Peremans et al. (2016) implemented several robust bootstrap procedures in the claims reserving framework.

The chain-ladder method has been extended as well as combined with other methods. The Munich chain-ladder method proposed by Quarg and Mack (2004) is based on and generalizes the classical Mack's chain-ladder model. It is a joint model for paid and incurred triangles that takes into account the

dependence of the development factors on the paid/incurred ratios. Then there is an arithmetic chain-ladder, where the development factors are considered as ratios of arithmetic means of claims in the current development period relative to claims in the previous period, see Kuang et al. (2009) and Taylor (2011) for a recent discussion on how the arithmetic chain-ladder technique arises through maximum likelihood estimation in a Poisson model. In addition, a well-known alternative to the arithmetic chain-ladder is a model involving the same parametrization but using lognormal distributed variables. This alternative approach was developed further by Kuang et al. (2015) providing closed form expressions for estimators for individual parameters. They derived log development factors which are differences of logs of geometric means and therefore it was referred to as the geometric chain-ladder model. One of the recent developments of the classical chain-ladder method is the extension of considering both the triangle of paid losses and a triangle of incurred claim counts, leading to the double chain-ladder method (Martínez Miranda et al. (2012)).

Regardless of all the stochastic methods suggested over the years, the work of finding a better version of or adding an appealing extension to the classical chain-ladder method keeps evolving and the method itself continues to being the subject of many research papers.

1.3 Individual claims reserving

All these macro-level models discussed in the Section 1.2 are based on aggregate data formulated in a run-off triangle, which makes the models easier to handle, but can be somewhat limiting and is considered as a possible weakness. One of the main advantages of aggregate methods are the low data requirements and computational power, which on the other hand causes a loss of useful (detailed) information and therefore individual claims prediction is not available. It has been questioned by several authors about possible use of a detailed micro-level information. Moreover, these days insurance companies do have access to extensive micro-level data. Micro-level models handle claims related data on an individual basis, rather than aggregating by underwriting year and development period. Due to the aggregation of the data, useful information regarding the claim development process is lost, but micro-level models work on the individual claim level to deal with the development throughout each claim's lifetime. One single insurance claim may have a life that develops over a number of years with a multiple transactions related to the claim.

Despite many advantages of micro-level loss reserving models, practitioners and researchers have paid rather little of attention to this topic, primarily due to the modeling complexity. In most research articles on this topic, the marked Poisson point process has been the dominant modeling technique. Research papers by Arjas (1989), Jewell (1989) and Norberg (1993, 1999) are the first contributions in the area of micro-level loss reserving. It was proposed in Arjas (1989) to model the development of a claim by a marked Poisson process, in which the occurrence of transactions follows a non-homogeneous Poisson process and the payment amount in each transaction is treated as the "mark". Jewell (1989) focused on modeling the number of IBNR claims with a marked Poisson process, in which the occurrence of claims is modeled by a non-homogeneous Poisson process and the reporting delay is treated as the "mark". Also the use of parametric Bayesian approach to predict the number of IBNR claims was proposed in the latter.

Later, the work of Arjas and Jewell was extended by Norberg (1993, 1999), who built a mathematical framework for applying a marked Poisson process in loss reserving on the individual claim level. He modeled claims using a marked Poisson process, in which claims occur in accordance with a non-homogeneous Poisson process and other stochastic characteristics about the claims are treated as the time dependent "marks". The first detailed implementation of a micro-level loss reserving model was done by Haastrup and Arjas (1996). The authors used the theoretical framework set up by earlier contributions and provided a case study with a portfolio of accident insurance to illustrate the implementation of a marked Poisson process.

Only in recent years, micro-level loss reserving has attracted more research interest again. Literature in the field of micro-level models contains several proposals of reserving models that are based on individual level claims data and several studies based on simulation or empirical data have been conducted. For instance, a substantial case study of the marked Poisson process method is provided by Antonio and Plat (2014), where a detailed stochastic hierarchical model for each part of the development of a claim is specified: occurrence time, reporting delay, transaction times and their severity, and the final settlement of the claim. The model is then calibrated using historical claim-level data of material and injury claims from a personal-line general liability insurance portfolio and the model is used to simulate the development process of each open claim. Jin and Frees (2013) estimate a similar model without the use of exposure data from the insurance portfolio. Verrall and Wüthrich (2016) construct an inhomogeneous marked Poisson process to explicitly model the claims arrival process and reporting delay in continuous time based on individual claims data. Two real individual claims data sets from property

and casualty insurance are calibrated to the statistical model of Jewell and Norberg.

Slightly differing from the regular Poisson process approach, Badescu et al. (2016a,b) and Avanzi et al. (2016) propose to model the claim arrival process along with its reporting delays as a marked Cox process to allow for over-dispersion and serial dependency. A Cox process, or doubly stochastic Poisson process, extends a Poisson process by modeling the intensity as a non-negative stochastic process. Badescu et al. (2016b) use a weekly piecewise constant stochastic process generated by a hidden Markov model (HMM) with state-dependent Erlang distributions. The discrete process of the number of observed claims during each week then follows a Pascal-HMM with scale parameters depending on the exposure and the reporting delay distribution.

Recently, new innovative approaches have been presented in claims reserving. Namely, machine learning techniques have been introduced by Wüthrich (2016) in individual claims reserving. Many contributions that are based on individual claims data assume a rather fixed structural form, for example, Pigeon et al. (2013) fit a multivariate skew normal distribution to the claims payments. Such fixed structural forms are not very flexible and the consideration of detailed feature information is difficult to implement. Machine learning techniques make it feasible to calculate claims reserves on individual claims data. Paper by Wüthrich (2016) illustrates how machine learning techniques can be used by providing an explicit example in individual claims reserving using regression trees. This approach has not yet been fully developed, but is an interesting contribution in the given field.

1.4 Motivation

Despite the fact that micro-level models have recently emerged in an increasing steam of academic literature, these models are not substantially used by practitioners. It can be explained by several reasons, for instance, the complex structure of the proposed model, which make the application of the models difficult. Proposed detailed models are usually computationally not efficient, and practitioners are slow in adapting with the novel approaches. Insurers consider developing new models only when they believe the models are implementable. It is clear that the industry-wide standard to estimate the future claim counts in the lower triangle is the chain-ladder model and its related extensions.

Reserving is a practical activity and the ultimate goal of studies on reserv-

ing models is to help insurers improve their reserving practice. The aim of this research is to shift the focus from creating new methodologies to testing effectiveness of existing ones in practical situations. Like mentioned beforehand, one of the main interests in actuarial practice is to estimate the likely variability of the claims reserve. The prediction error has been used as the precision measure for the reserve estimates in most literature. The first part of this dissertation concentrates mainly on the prediction error issue. The prediction error can be decomposed into two components: parameter uncertainty and process uncertainty (see, e.g., England and Verrall (2002), Taylor (2014)). The former comes from the uncertainty in the estimation of parameters of the reserving model due to the limited sample size, whereas the latter comes from the intrinsic randomness of the claims development in the future. However, obtaining estimates for the standard error of prediction can be a difficult task. There are several analytical results for computing the prediction error (see England and Verrall (1999)), but those estimates can be difficult to calculate or are only approximate values. In addition, calculating the prediction error certainly provides great insight into the performance of reserve estimates, but other information such as the cash flow or risk measures are also of interest. For both the prediction errors and the full predictive distribution of reserve estimates, bootstrapping can be used for a solution.

The bootstrap technique has been extensively studied in the claims reserving framework by various authors, such as Ashe (1986), England and Verrall (1999) and Pinheiro et al. (2003). The definition of the proper residuals to base the bootstrap technique on is definitely an open subject when bootstrapping. We extend the work of Pinheiro et al. (2003) by using another useful type of residual with bootstrapping, and we carry out a comparative study among several stochastic models. We will use claims data from an Estonian insurance company for the empirical study, where we discuss the impact of the chosen models and the residuals on the reserve estimates and prediction errors. Moreover, as contributions involving case studies are usually based on one or two data sets, we provide an extensive case study with the Schedule P database. In addition, similar comparison studies do not straightforwardly report the best model, indicating a clear need for more proper tools to validate and assess the quality of predictions when comparing different reserving methods. In order to validate the reserving method and identify any needed modifications, we need to rank the competing predictive models. We propose to consider scoring rules to measure the accuracy of probabilistic predictions.

There are many recent contributions conducting a comparison of macro and micro-level models (see, e.g., Charpentier and Pigeon (2016)). Several studies with empirical comparisons between macro- and micro-level models (see, e.g.,

Jin and Frees (2013)) show that their calibrated micro-level model produces a more realistic reserve compared to the macro-level models. For instance, the proposed micro-level model by Antonio and Plat (2014) outperforms the macro-level models providing the actuary with detailed information regarding the development process of each claim and obtained on real data analysis lower variance on the total amount of reserves than with the macro-level model. On the other hand, Johansson (2015) shows with the empirical study that the variances of the chain-ladder reserve and the double chain-ladder reserve are larger than the variance of the micro-level method reserve, but the micro-level method does not estimate a better reserve than the two other methods. Thus, it is not straightforward that micro-level models generate reserve estimates with better quality in every situation, but it is indeed a natural and sensible direction to pursue a better reserve estimate. Given the data availability and advances of computational power, it does make sense to use that data.

It could be that micro-level methods perform better than macro-level models but it is not clear whether using sophisticated models is worth the extra effort. Perhaps it would be more sustainable to use the available granular data but hold to a simpler method, which can be extended or corrected to overcome its shortcomings. At the end of the day, the claims reserving actuary wants to keep the company solvent and profitable giving the best estimate for the reserves.

One of the most recent chain-ladder extensions is the method by Martínez Miranda et al. (2013), which extends the traditional chain-ladder framework towards the continuous use of individual claims data, where granular data is given in a run-off triangle and reserves are estimated with non-parametric estimation of the underlying density. The continuous chain-ladder (CCL) is the first chain-ladder extension up to our knowledge that does not assume the data aggregation and the model can be considered as a micro-level model due to the use of continuous time. In theory, the CCL model is designed for data recorded in continuous time. However, insurance companies do not always keep track of insurance data in continuous time and the lowest level of data aggregation that could be considered, is daily data. The CCL method has not been extensively developed and investigated from the first moment it was published and we have not seen contributions involving simulation or comparison studies using the CCL model. In this thesis, we look into the CCL method, implement the method and discuss several associated problems using kernel functions. As this dissertation serves a practical input to the existing literature, we describe the method in a discrete framework for practical situations.

The classical chain-ladder method has mostly been used with annual data and we have not seen contributions where the reserve estimation is based on more refined level of data, such as quarterly, monthly or even daily data. We investigate whether and how much the use of different aggregation levels of claim data can improve the reserving process with both the CCL and the CL method. The data is simulated on daily basis and evaluation of the impact of each level of data aggregation (monthly, quarterly and annual) on the predictive distribution of the IBNR claim counts will follow. We provide a simulation study and also an empirical study to examine the possible gain of a micro-level approach and to judge whether the different levels of data aggregation influence the reserve estimates.

While the existing literature has contributed a solid mathematical framework for both macro and micro-level reserving models, this dissertation supplements this literature by providing a more practical approach to demonstrate how to implement some of the discussed reserving models and the benefits that one receives from them. The reserving methods suggested in the research literature seldom fit practical applications perfectly and, hence, the actuaries often have to make ad hoc adjustments. Moreover, in the literature the suggested methods tend to be illustrated using neat and uncomplicated data sets and are therefore not so often questioned. In insurance practices, each data set poses specific challenges to modeling and model selection, imposing an another flaw of the current literature: there is a wide range of methods giving slightly different results, but the current literature lacks any systematic guide to choosing between them. In this thesis, the emphasis is on analyzing real-life insurance data which is complemented by simulation studies, as the main drive of this research has been the practical relevance and applicability. An important objective of the proposed work is the guidance in practical situations making our findings accessible to insurance practitioners. Only a correct use of the available data, proposed models and prediction methods, with an appropriate study of the statistical uncertainties involved, will lead to the expected benefits. In addition, the use of proper statistical techniques should help to get reliable uncertainty estimates. The overall objective is to improve actuarial practices for reserving by using sound and flexible statistical methods shaped for the actuarial data. These objectives clearly require an interdisciplinary research approach, combining actuarial science and mathematical statistics. The highly advanced information that is currently gathered by insurance companies justifies the challenges formulated in this thesis as well as the statistical methodology.

Chapter 2

Stochastic reserving methods and bootstrapping

In this chapter, we consider the well-known stochastic reserve estimation methods on the basis of generalized linear models, such as the (over-dispersed) Poisson model, the gamma model and the lognormal model. For the likely variability of the claims reserve, bootstrap method is considered. In the bootstrapping framework, we discuss the choice of residuals, namely the Pearson residuals, the deviance residuals and the Anscombe residuals. In addition, several possible residual adjustments are discussed and compared in a case study. We carry out a practical implementation and comparison of methods using real-life insurance data to estimate reserves and their prediction errors. We propose to consider proper scoring rules for model validation, and the assessments will be drawn from an extensive case study.

This chapter is an extended version of Tee, L., M. Käärik, and R. Viin (2017). On comparison of stochastic reserving methods with bootstrapping. *Risks* *5*(1), 2.
In comparison with the original publication, we have adjusted descriptive sections, developed the subsection of residuals and extended the case study section with Schedule P database to involving data from different lines of businesses.

2.1 Chain-ladder method as a generalized linear model

In this section, we introduce briefly the basic chain-ladder method, recall how the chain-ladder method is reformulated in the context of generalized linear models and give a brief review of stochastic macro-level models, which will be used in the analysis. For a general introduction to GLM, we refer to McCullagh and Nelder (1989).

Stochastic macro-level models use aggregate claims data, and some of the main advantages over non-stochastic macro-level models are the possibilities to obtain first two moments or the predictive distribution of the reserve estimate. Several often-used and traditional actuarial methods to complete a run-off triangle can be described by GLM. The actuarial literature has also shown a close connection between the chain-ladder method and the multiplicative Poisson model.

Without loss of generality, we assume that the data that have been collected for $i = 1, ..., n$ and $j = 1, ..., n$ consist of a triangle of incremental claims:

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

where the row index i refers to the year of origin and, depending on a particular situation, indicates the accident year, reporting year or underwriting year. The column index j refers to the development year, indicating the delay, more precisely loss disbursal, reporting year or accident year. Claims data are given as a run-off triangle as shown in Table 2.1.

	Development Period j				
Year of Origin i	**1**	**2**	**3**	...	n
1	C_{11}	C_{12}	C_{13}	...	C_{1n}
2	C_{21}	C_{22}	...		
3	C_{31}	...			
⋮	⋮				
n	C_{n1}				

Table 2.1: Run-off triangle with incremental claim amounts.

The cumulative claim amounts with accident year index i reported up to, and including, the delay index j are defined as:

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

Thus, D_{ij} is the total claims amount of accident year i, $i = 1, \ldots, n$, either paid or incurred up to development year j, $j = 1, \ldots, n$. The development factors of the chain-ladder technique are estimated as:

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

Generalized linear modeling is a methodology for modeling the relationships between variables. It generalizes the classical normal linear model, by relaxing some of its restrictive assumptions, and provides methods for the analysis of non-normal data. GLM is important in the analysis of insurance data, because with insurance data, the assumptions of the normal model are often not applicable. See De Jong and Heller (2008) for a detailed description of generalized linear models for insurance data.

Following Renshaw and Verrall (1998); Pinheiro et al. (2003), the structure of the stochastic models for claim reserving in the terminology of GLM can be given by:

(1) incremental claim amounts C_{ij} belong to the exponential family,

(2) $\mathbb{E}(C_{ij}) = \mu_{ij}$,

(3) $\eta_{ij} = g(\mu_{ij})$, where $g(\cdot)$ is the link function,

(4) linear predictor $\eta_{ij} = c + \tilde{\alpha}_i + \tilde{\beta}_j$ with an intercept c and factor effects $\tilde{\alpha}_i$ and $\tilde{\beta}_j$.

The given structure of GLM can be used to describe several often used actuarial methods. We consider the following multiplicative model (Kaas et al. (2008)), with a parameter for each row i, each column j and each diagonal $k = i + j - 1$:

$$C_{ij} \approx \alpha_i \cdot \beta_j \cdot \gamma_k, \tag{2.1}$$

where parameter α_i describes the effect of year of origin i, parameter β_j corresponds to development year j and γ_k describes the effect of calendar year $k = i + j - 1$. The approximation sign in Equation (2.1) expresses a difference caused by a chance, i.e., there is a possible deviation of the observation on the left-hand side from its mean value on the right-hand side. The model involves three time scales, which give rise to the well-known identification problem. Parametrization using three time scales has been introduced for instance by Zehnwirth (1994). The identification problem has been revisited by several authors; see, for example, Kuang et al. (2008b,a), who have proposed

a canonical parametrization that is uniquely identified. In the framework of three time scales, we also face a problem with extrapolating the calendar estimates. Namely, we have no data on the values of γ_k for the future calendar years, e.g., if $k > n$. This can be overcome by assuming that the γ_k have a geometric pattern, with $\gamma_k \propto \gamma^k$ for some real number γ. Typically, the model (2.1) is simplified by taking $\gamma_k \equiv 1$, and the condition $\sum_{j=1}^{n} \beta_j = 1$ is imposed. If the parameters $\alpha_i > 0$ and β_j are estimated by using the maximum likelihood method, then the simplified model is a multiplicative GLM with log-link.

In the terminology of GLM, to linearize the multiplicative model (2.1), the logarithm is chosen as a link function (log-link). Hence:

$$\begin{aligned} \mathbb{E}(C_{ij}) = \mu_{ij} &= \alpha_i \cdot \beta_j \cdot \gamma_k \\ &= \exp(\ln \alpha_i + \ln \beta_j + \ln \gamma_k), \end{aligned}$$

or, equivalently,

$$\ln \mathbb{E}(C_{ij}) = \ln \alpha_i + \ln \beta_j + \ln \gamma_k. \tag{2.2}$$

Parameters of the given model are estimated by using the maximum likelihood method. After obtaining the estimates of the parameters, it is easy to complete the run-off triangle, simply by taking:

$$\widehat{C}_{ij} := \widehat{\alpha}_i \cdot \widehat{\beta}_j \cdot \widehat{\gamma}_k. \tag{2.3}$$

This simple model allows one to generate quite a few reserving techniques, depending on the assumptions set on the distribution of C_{ij}. It is common in claim reserving to consider the Poisson, gamma or lognormal distribution for the variable C_{ij}. We proceed with reviewing the following methods from model (2.1).

The (over-dispersed) Poisson model: Already in 1975, a stochastic model corresponding to the Poisson model, which leads to the chain-ladder technique, was proposed. This model works on the incremental amounts C_{ij} from a Poisson distribution, where $\mathbb{E}(C_{ij}) = \alpha_i \beta_j$ with unknown parameters α_i and β_j. Here, α_i is the expected ultimate claims amount (up to the latest development year so far observed), and β_j is the proportion of ultimate claims to emerge in each development year with the restriction $\sum_{k=1}^{n} \beta_k = 1$. The restriction immediately follows from the fact that β_j is interpreted as the proportion of claims reported in development year j. Obviously, the aggregate proportion over all periods has to be one.

Throughout this chapter, we use the notation Δ for the triangle of known data, i.e., the set of all (i, j), where C_{ij} is known, and ∇ for the triangle of unknown data, i.e., the set of all (i, j), where C_{ij} is unknown. We also distinguish $\Delta_i = \{j : (i, j) \in \Delta\}$ and $\Delta^j = \{i : (i, j) \in \Delta\}$. Using this notation, we can define the reserve R_i in origin year i by $R_i = \sum_{j\in\nabla_i} C_{ij}$, and the total reserve by $R = \sum_{(i,j)\in\nabla} C_{ij}$. The corresponding estimates are given by $\widehat{R}_i = \sum_{j\in\nabla_i} \widehat{C}_{ij}$ and $\widehat{R} = \sum_{(i,j)\in\nabla} \widehat{C}_{ij}$, respectively.

We estimate the unknown parameters α_i and β_j from the triangle of known data with the maximum likelihood method. Assume that all C_{ij} are independent with a Poisson distribution, and $\mathbb{E}(C_{ij}) = \alpha_i\beta_j$ holds. Then, the maximum likelihood estimators $\widehat{\alpha}_i$ and $\widehat{\beta}_j$ are given by

$$\widehat{\alpha}_i = \frac{\sum_{j\in\Delta_i} C_{ij}}{\sum_{j\in\Delta_i} \beta_j}, \quad i = 1, \ldots, n \tag{2.4}$$

and

$$\widehat{\beta}_j = \frac{\sum_{i\in\Delta^j} C_{ij}}{\sum_{i\in\Delta^j} \alpha_i}, \quad j = 1, \ldots, n. \tag{2.5}$$

The maximum likelihood estimates for unknown parameters α_i and β_j are derived with the likelihood function $\mathscr{L} = \mathscr{L}(\boldsymbol{\alpha}, \boldsymbol{\beta})$, where $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n)$, as follows

$$\mathscr{L} = \prod_{(i,j)\in\Delta} \frac{(\alpha_i\beta_j)^{C_{ij}}}{C_{ij}!} \exp(-\alpha_i\beta_j).$$

Therefore, the log likelihood function is:

$$\ell = \ln(\mathscr{L}) = -\sum_{(i,j)\in\Delta} \alpha_i\beta_j + \sum_{(i,j)\in\Delta} C_{ij} \ln(\alpha_i\beta_j) - \sum_{(i,j)\in\Delta} \ln(C_{ij}!),$$

where the summation is for all (i, j) where C_{ij} is known. The maximum likelihood estimator consists of values of α_i, β_j, which maximize $\mathscr{L}$ or equivalently $\ln(\mathscr{L})$. They are given by the equations:

$$0 = \frac{\partial\ell}{\partial\alpha_i} = -\sum_{j\in\Delta_i} \beta_j + \sum_{j\in\Delta_i} C_{ij}\frac{1}{\alpha_i}, \quad i = 1, \ldots, n$$

and:

$$0 = \frac{\partial\ell}{\partial\beta_j} = -\sum_{i\in\Delta^j} \alpha_i + \sum_{i\in\Delta^j} C_{ij}\frac{1}{\beta_j}, \quad j = 1, \ldots, n.$$

Thus, the likelihood estimator α_i and β_j is given, respectively, by Formulas (2.4) and (2.5). The proportion factors β_j express the ratio of the sum of

observed incremental values for certain development year j with respect to certain ultimate claims, i.e., β_i denotes the proportion of claims reported in development year j. The parameters α_i refer to the ratio of the sum of observed incremental values for a certain origin year i to corresponding proportion factors. In other words, if the incremental claim amounts and respective proportions factors are known, it is simple to derive the corresponding ultimate claim α_i for origin year i. One can note the principal similarities with the chain-ladder technique, where development factors are also the outcomes of certain ratios.

The Poisson model can be cast into the form of a GLM, and to linearize the multiplicative model, we need to choose the logarithm as a link function, $\eta_{ij} = \ln(\mu_{ij})$, so that:

$$\mathbb{E}(C_{ij}) = \mu_{ij} = \exp(\ln(\alpha_i) + \ln(\beta_j))$$

or, equivalently,

$$\ln(\mathbb{E}(C_{ij})) = \ln(\alpha_i) + \ln(\beta_j) \tag{2.6}$$

where the structure of linear predictor (2.6) is still a chain-ladder type, because parameters for each row i and each column j are given. Hence, the structure (2.6) is defined as a GLM in which the incremental values C_{ij} are modeled as Poisson random variables with a log-link. Reparametrizing (2.6) gives us a structure of property (4) defined in a GLM setting, i.e., we obtain a linear predictor:

$$\eta_{ij} = c + \tilde{\alpha}_i + \tilde{\beta}_j, \tag{2.7}$$

where parameter c can be considered as an intercept, which corresponds to the incremental amount in the cell (1, 1). This means, we take

$$\begin{aligned} c &= \ln \alpha_1 + \ln \beta_1, \\ \tilde{\alpha}_i &= \ln \alpha_i - \ln \alpha_1, \quad i = 1, \ldots, n, \\ \tilde{\beta}_j &= \ln \beta_j - \ln \beta_1, \quad j = 1, \ldots, n. \end{aligned}$$

The Poisson model was studied in further detail by Kuang et al. (2009), where also a new canonical parametrization was proposed.

We recall that the only distributional assumptions used in GLMs are the functional mean-variance relationship and the fact that the distribution belongs to the exponential family. When defining a GLM, we can omit the distribution of C_{ij}'s and use only the most elementary information about the response variable, namely the relationship between variance and mean. This

introduces a quasi-likelihood as an alternative, and using this elementary information alone can be often sufficient to stay close to the full efficiency of maximum likelihood estimators. Therefore, we can estimate the parameters by the maximum quasi-likelihood (McCullagh and Nelder (1989)) instead of the maximum likelihood, and the estimators remain consistent. However, it is necessary to impose the constraint that the sum of the incremental claims in every row and column has to be non-negative. This means that quasi-likelihood could not be used, for instance, when modeling incurred data with a large number of negative incremental claims in the later development periods.

In the case of the Poisson distribution, the mentioned relationship is $\mathrm{Var}(C_{ij}) = \mathbb{E}(C_{ij})$, and allowing for more or less dispersion in the data can be generalized to $\mathrm{Var}(C_{ij}) = \phi\mathbb{E}(C_{ij})$ without any change in form and solution of the likelihood equations. This kind of generalization allows for more dispersion in the data, and one speaks of an over-dispersed Poisson (ODP) model. The ODP model is a generalization of the Poisson model and overcomes many of the limitations while retaining the same basic structure and the desirable feature that the reserve estimates are identical to the estimates obtained with the CL method. The restriction to non-negative values takes the much weaker form that only the column totals must be positive in the triangle of incremental claim amounts, but note that this implies that the expected value of the incremental claims amounts cannot be negative. It is shown in Schmidt (2002) that every ODP model can be transformed into the Poisson model by dividing all incremental claims by a certain parameter. The general form for the ODP model can be given as follows:

$$\mathbb{E}(C_{ij}) = \mu_{ij} = \alpha_i\beta_j, \tag{2.8}$$

$$\mathrm{Var}(C_{ij}) = \phi\alpha_i\beta_j, \tag{2.9}$$

where:

$$\sum_{k=1}^{n} \beta_k = 1.$$

The over-dispersion is introduced through the parameter ϕ, which is unknown and estimated from the data. The scale parameter ϕ may either be constant, or allowed to vary by development period. In practice it is usually more appropriate to allow the scale parameter to vary by development period. Considering a single incremental payment C_{ij} with the origin year i and claim payments in development year j (yet to be observed), we obtain the estimates of future payments from the parameter estimates by inserting them into Equation (2.6) and exponentiating, resulting as:

$$\widehat{C}_{ij} = \widehat{\alpha}_i\widehat{\beta}_j = \exp(\widehat{\eta}_{ij}). \tag{2.10}$$

Given the Equation (2.10), the reserve estimates for any origin year can be derived by:

$$\widehat{R}_i = \widehat{\alpha}_i\widehat{\beta}_{n+2-i} + \ldots + \widehat{\alpha}_i\widehat{\beta}_n, \quad i = 2, \ldots, n, \tag{2.11}$$

and the reserve estimate for the total amount can be easily derived by summation:

$$\widehat{R} = \sum_{i=2}^{n} \widehat{R}_i = \sum_{i=2}^{n} (\widehat{\alpha}_i\widehat{\beta}_{n+2-i} + \ldots + \widehat{\alpha}_i\widehat{\beta}_n), \quad i = 2, \ldots, n. \tag{2.12}$$

The negative binomial model can be derived from the Poisson model, and thus, these models are very closely related, but with a different parameterization. The negative binomial model was first derived by Verrall (2000), by integrating out the row parameters from the Poisson model. The predictive distributions of both models are basically the same and give identical predicted values.

Lognormal model: When considering the lognormal distribution to describe claim amounts (see for a reference Kremer (1982)), we can still continue to use GLM for the logs of the incremental claim amounts. The lognormal class of models are given as:

$$\ln(C_{ij}) \sim \mathcal{N}(\mu_{ij}, \sigma^2),$$

i.e.,

$$\mathbb{E}(\ln(C_{ij})) = \mu_{ij} = \ln(\alpha_i) + \ln(\beta_j) \text{ and } \quad \text{Var}(\ln(C_{ij})) = \sigma^2.$$

Now, the identity link function is used, and the normal responses $\ln(C_{ij})$ are assumed to decompose (additively) into a deterministic non-random component with mean $\mu_{ij} = \eta_{ij}$ and normally-distributed random error components with zero mean.

The fitted values on a log scale, given the estimates for the parameters in the linear predictor η_{ij} and the process variance σ^2, are obtained by forming the appropriate sum of estimates. Obtaining the estimates for the mean on the untransformed scale is not that simple. We cannot just exponentiate the linear predictor, since that would give an estimate of the median. Therefore, the fitted values on the untransformed scale are given by:

$$\widehat{C}_{ij} = \exp(\widehat{\eta}_{ij} + \frac{1}{2}\widehat{\sigma}^2), \tag{2.13}$$

which is in the standard form of the expected value of a lognormal distribution. The reserve estimate in origin year i is given by summing the predicted values

in row i of ∇, i.e., $\widehat{R}_i = \sum_{j \in \nabla_i} \widehat{C}_{ij}$, and the total reserve estimate, summing the predicted values in row i and in column j of ∇, is given by $\widehat{R} = \sum_{(i,j) \in \nabla} \widehat{C}_{ij}$. The lognormal model is also referred to as the geometric chain-ladder model; see this additional analysis in Kuang et al. (2015).

Gamma model: Mack (1991) proposed a model with a multiplicative parametric structure for the mean incremental claims amounts which are modeled as gamma response variables. As noted in Renshaw and Verrall (1998), the same model can be fitted as the GLM described in the ODP model, except that the incremental claim amounts are modeled as independent gamma response variables with a logarithmic link function and the same linear predictor; and also requiring a slight change in (2.9). The gamma model implemented as a generalized linear model gives exactly the same reserve estimates as the gamma model implemented by Mack (1991). The gamma model is given with the mean:

$$\mathbb{E}(C_{ij}) = \mu_{ij} = \alpha_i \beta_j,$$

subject to:

$$\sum_{k=1}^{n} \beta_k = 1,$$

and with the variance:

$$\mathrm{Var}(C_{ij}) = \phi(\mathbb{E}(C_{ij}))^2 = \phi \mu_{ij}^2.$$

Estimates for the single incremental payments C_{ij} are then given by:

$$\widehat{C}_{ij} = \widehat{\alpha}_i \widehat{\beta}_j = \exp(\widehat{\eta}_{ij}). \tag{2.14}$$

To obtain reserve estimates with the gamma model for any origin year or for the overall amount, the same formulas as defined in the ODP model, (2.11) and (2.12), respectively, can be used. The limitation of both the gamma and the ODP model is that each incremental value should be non-negative. As with the lognormal model, the predicted values provided by the gamma model are usually close to the chain-ladder estimates, but it cannot be guaranteed.

2.2 The bootstrap technique

Bootstrapping is a popular technique in stochastic claims reserving because of the simplicity and flexibility of the approach. We are using bootstrapping to

estimate the prediction error and to approximate the predictive distribution. An analytical derivation of the prediction error of the total reserve estimate may be preferable from a theoretical perspective, but it is often impracticable due to complex reserve estimators. For the classical chain-ladder method, Mack (1993) derived an analytical expression of the Mean Squared Error of Prediction (MSEP) within an autoregressive formulation of the claims development using a second-moment assumption. The first order Taylor approximation of the corresponding MSEP within the GLM framework was derived by England and Verrall (1999). As said, known theoretical estimators are difficult to calculate and are still merely approximate values.

For both the classical and generalized linear model, it is common to adopt either a paired bootstrap where resampling is done directly from the observations or the residuals bootstrap where resampling is applied to the residuals of the model (Pinheiro et al. (2003)). The paired bootstrap is more robust than the residual bootstrap, but only the residual bootstrap can be implemented in the context of the claim reserving, given the dependence between some observations and the parameter estimates. If the type of residuals adopted is the same, then mixing GLMs with bootstrapping is similar to combining the chain-ladder method with bootstrapping. The residuals obtained from applying a GLM to the past claims data are used in the resampling process of bootstrapping. With each re-sampled set of residuals, an upper triangle can be constructed, and the stochastic chain-ladder can be applied again. The lower triangle is then simulated from the assumed distribution with the first two moments determined by the stochastic chain-ladder. Thereafter, an empirical distribution is formed, from which the required inferences can be drawn.

2.2.1 Residuals

The process of creating a distribution for the reserve can be done by either parametric bootstrapping or non-parametric bootstrapping. Given the context of this research, we use a non-parametric bootstrapping method, i.e. use the residuals to bootstrap a claims reserves distribution.

The most often used residuals in model diagnostics are the Pearson residuals and the deviance residuals. Furthermore, the Anscombe residual is often mentioned as a possible residual to consider, but is rarely applied in further work due to being known as a less commonly-used residual. However, following De Jong and Heller (2008), the Anscombe and the deviance residuals are mathematically different, but numerically, they give similar results. The

Anscombe residual tries to make the residuals "as close to normal as possible", and given that the response distribution has been correctly specified, the deviance residuals are also approximately normally distributed. Thus, contrary to the usual practice, we explore in the following the use of the Anscombe residuals. A version of the deletion residual is also available under the GLM setting, which is related to the Pearson residual, but their forms are rather complicated and therefore omitted.

The Pearson residuals are just rescaled versions of the raw or response residuals and are defined as:

$$r_{ij}^{P} = \frac{C_{ij} - \widehat{\mu}_{ij}}{\sqrt{V(\widehat{\mu}_{ij})}}, \tag{2.15}$$

where $V(\cdot)$ is a variance function. The Pearson residuals need to be adjusted in order to obtain (approximately) equal variance, and there are different adjustments suggested by several authors. It was proposed by England and Verrall (1999, 2002) to adjust the residuals by multiplying them by a correction factor $\sqrt{\frac{n_\Delta}{n_\Delta - p}}$, where n_Δ is the sample size and p is the number of estimated parameters. In correspondence with the classical linear model, often the "hat" matrix of the model is used to standardize the Pearson residuals, which are given as:

$$r_{ij}^{P*} = \frac{r_{ij}^{P}}{\sqrt{\widehat{\phi}(1 - h_{ij})}}, \tag{2.16}$$

where $\widehat{\phi}$ is a scale parameter estimated from the data, and the factor h_{ij} is the corresponding element of the diagonal of the "hat" matrix. The scale parameter can be estimated, for example, as the Pearson chi-squared statistic divided by the degrees of freedom:

$$\widehat{\phi} = \frac{\sum (r_{ij}^{P})^2}{n_\Delta - p}, \tag{2.17}$$

where n_Δ is the number of data points in the sample and p is the number of parameters estimated, and the summation is over the number of residuals. The "hat" matrix is given for classical linear models by $\mathbf{H} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$, and it can be generalized for GLM as follows:

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W},$$

where $\mathbf{X}$ is a design matrix and $\mathbf{W}$ is a diagonal matrix with elements:

$$w_{ii} = \left(V(\mu_{ij}) \left(\frac{\partial \eta_{ij}}{\partial \mu_{ij}} \right)^2 \right)^{-1}$$

on the diagonal (see McCullagh and Nelder (1989) for details). In this thesis, residual adjustment procedure (2.16) is used in Chapters 2.3 and 2.4.

The distribution of Pearson residuals for non-normal distributions is often markedly skewed and, thus, may fail to have properties similar to those of a normal-theory residual. Then, the Anscombe residual can be a good alternative to the Pearson residual. The Anscombe residuals do not use the variable C_{ij} directly, but instead a transformation $A(C_{ij})$. The function $A(\cdot)$ is chosen to make the distribution of $A(C_{ij})$ as normal as possible and is given by

$$A(x) = \int_{-\infty}^{x} V^{-\frac{1}{3}}(t)dt,$$

where $V(t)$ is the variance function. In the context of GLM the Anscombe residual is defined as

$$r_{ij}^{A} = \frac{A(C_{ij}) - A(\widehat{\mu}_{ij})}{A'(\widehat{\mu}_{ij}) \cdot \sqrt{V(\widehat{\mu}_{ij})}}, \tag{2.18}$$

where $A'(\mu)$ is the derivative of $A(\mu)$. For the Poisson model, the Anscombe residuals are defined by:

$$r_{ij}^{A} = \frac{\frac{3}{2}\left(C_{ij}^{\frac{2}{3}} - \widehat{\mu}_{ij}^{\frac{2}{3}}\right)}{\widehat{\mu}_{ij}^{\frac{1}{6}}}$$

and for the gamma model the residuals are defined as:

$$r_{ij}^{A} = 3\left(\left(\frac{C_{ij}}{\widehat{\mu}_{ij}}\right)^{\frac{1}{3}} - 1\right).$$

It is easy to see that in case of the normal model, the Anscombe residuals (2.18) are equivalent to the classical residuals, and thus, for the lognormal model the residuals are defined as:

$$r_{ij}^{A} = \ln(C_{ij}) - \widehat{\mu}_{ij},$$

since $V(\mu_{ij}) = 1$. Note that similar standardization procedure (2.16) could be defined for the use of another kind of residuals, such as the Anscombe residuals. For a detailed overview of residuals, see De Jong and Heller (2008).

Prediction errors obtained with the bootstrapping method are compared based on the type of residuals used and if or how we have adjusted the residuals. It is important to notice that the residuals of the calculated values of the first column in the last row and of the first row in the last column are always

equal to zero, i.e. $\widehat{\mu}_{1n} - C_{1n} = 0$ and $\widehat{\mu}_{n1} - C_{n1} = 0$. These are zeros due to the defined linear structure adopted in the models implying the estimates for some of the parameters depend on one observation only. The reason for the correction of zeros is that the bootstrap method assumes the random variables (in this case, residuals) to be i.i.d. random variables, but in this case, there are two non-random residuals, which are always fixed as zeros. Thus, we remove zero-residuals and replace them with residuals resampled from the remaining ones. In the following of this chapter, the correction of zeros is referred to as the zero-correction.

2.2.2 Prediction error and confidence limits

Once the model has been chosen the variability of the claims reserve can be obtained either analytically or by simulation. The bootstrap is often adopted when a standard error is difficult or impossible to estimate analytically. A commonly-used measure of variability is the prediction error, also known as the root mean square error of prediction. In this context, we use the expected value as the prediction. The prediction error consists of two parts: the process variance and the estimation variance. The root mean squared error of the prediction (RMSEP) $\widehat{R}$ is given by:

$$\begin{aligned} \text{RMSEP}(\widehat{R}) &= \sqrt{\mathbb{E}((R-\widehat{R})^2)} \\ &= \sqrt{(\mathbb{E}(R)-\mathbb{E}(\widehat{R}))^2 + \text{Var}(R-\widehat{R})} \\ &\approx \sqrt{\text{Var}(R)+\text{Var}(\widehat{R})}, \end{aligned} \tag{2.19}$$

where $\text{Var}(R)$ denotes the process variance and $\text{Var}(\widehat{R})$ denotes the estimation variance. The explicit form of the two components of the variance in 2.19 depends on which prediction model is used; see for instance Renshaw (1994); England and Verrall (1999).

The bootstrap process involves resampling, with replacement, from the residuals. A bootstrap data sample is then created by inverting the formula for the residuals using the resampled residuals, together with the fitted values. Having obtained the bootstrap sample, the model is refitted and the reserve estimate is calculated. It is necessary to repeat the process a large number of times, each time creating a new bootstrap sample, and obtaining the reserve estimates. Then, the bootstrap standard error (SE_{bs}) is the standard deviation of the bootstrap reserve estimate, thus being an estimate of the square root of the estimation variance. However, the bootstrap standard error cannot be compared directly with the analytic equivalent since the bootstrap standard

error does not take into account the number of parameters used in fitting the model, i.e., the bootstrap process simply uses the residuals with no regard as to how they are obtained. To obtain the bootstrap prediction error (PE_{bs}), it is necessary to add an estimate of the process variance. Therefore, the bootstrap prediction error is the square root of the sum of the process variance and the squares of estimation variance:

$$\mathrm{PE}_{bs}(\widehat{R}) = \sqrt{\mathrm{Var}(R) + (\mathrm{SE}_{bs}(\widehat{R}))^2}, \tag{2.20}$$

where R stands for the total reserve and $\widehat{R}$ for its corresponding estimate. Note that the formula (2.20) can be applied analogously in case of origin year reserves R_i. $\mathrm{PE}_{bs}(\widehat{R})$ is the bootstrap standard error of the total reserve estimate, and process variance $\mathrm{Var}(R)$ has an explicit form depending on the considered model. In the case of ODP model, the estimate of the process variance would be

$$\widehat{\mathrm{Var}(R)} = \sum_{(i,j)\in\nabla} \widehat{\phi}\widehat{\mu}_{ij} = \widehat{\phi} \sum_{(i,j)\in\nabla} \widehat{\mu}_{ij} = \widehat{\phi}\widehat{R},$$

and in the case of gamma model

$$\widehat{\mathrm{Var}(R)} = \sum_{(i,j)\in\nabla} \widehat{\phi}\widehat{\mu}_{ij}^2 = \widehat{\phi} \sum_{(i,j)\in\nabla} \widehat{\mu}_{ij}^2.$$

In the case of lognormal model, the process variance is

$$\widehat{\mathrm{Var}(R)} = \sum_{(i,j)\in\nabla} \Big(\exp(\widehat{\sigma}^2) - 1\Big) \exp(2\widehat{\mu}_{ij} + \widehat{\sigma}^2).$$

Following Davison and Hinkley (1997); Pinheiro et al. (2003), we consider an alternative bootstrapping procedure to obtain an upper confidence limit for the forecasts of the aggregate values. This approach (in the following named as the PPE-method) includes two resampling procedures in the same bootstrap iteration, but the results should be more robust against deviations from the hypothesis of the model. The idea is to define an adequate prediction error as a function of the bootstrap estimate and a bootstrap simulation of the future reality and to record the value of this prediction error for each bootstrap iteration. Then, use the desired percentile of this prediction error, and combine it with the initial prediction to obtain the upper limit of the prediction interval. See the appendix for a step-by-step explanation of this alternative approach as well as the first bootstrap procedure which we refer to as the regular method.

2.3 Empirical study

To enable a comparison with previously-discussed methods in the framework of bootstrapping with defined residuals, we use the real-life data set from an Estonian insurance company. We want to avoid using the well-known data sets that have been used by many others in the literature over the years and to contribute an analysis with a new data set. The data considered describe the paid out claims and are shown here in incremental form. We are interested in the impact of the choice of the model and, mainly, in the effect of the choice of residual and its adjustments.

We use both the Pearson and the Anscombe residuals (see definitions (2.15) and (2.18), respectively) first without any corrections, then with the zeros corrected and lastly standardized residuals together with the zero-correction (see (2.16) for the definition of standardization). It is clear that using just standardized residuals will lead to the same results as obtained with the zero-corrected residuals; thus, we do not consider standardized residuals independently in the comparative study. In addition, we compare the obtained prediction errors and upper limits using both bootstrap approaches, i.e., the regular bootstrap standard error of prediction (2.20) and the alternative (using pseudo-reality) PPE-method. We present the PPE prediction errors only for the total reserve. When comparing the bootstrap prediction errors PE_{bs} and the PPE prediction errors, we have to take into account that different units are used: bootstrap prediction error equals one standard deviation, and PPE prediction error equals (approximately) 1.645 standard deviations (95%-quantile of normal distribution). This means that we have to multiply the obtained bootstrap prediction error (2.20) by 1.645 and add it to the reserve estimate to obtain an upper confidence limit for the total reserve. In the case of PPE-method, we simply sum the prediction error and the mean to obtain the upper limit.

Reserve estimates obtained using the over-dispersed Poisson model, the gamma model and the lognormal model in the framework of bootstrapping with residuals outlined in this chapter are shown in Tables 2.3 – 2.8 below. As one can see, the data considered are rather inconvenient (see Table 2.2), i.e., the large fluctuation of the values in the triangle is obvious: the smallest incremental value is 1022, and the largest one is 10,660,074, which is a 10,430-fold difference. The second column in Tables 2.3 – 2.8 shows a point estimate for the reserve. These estimates are obtained directly from the defined model (not depending on the bootstrap procedure), and the point estimates do not depend on the choice of residual or on its correction.

	1	2	3	4	5	6	7	8	9	10
2000	4,734,994	1,885,305	281,240	504,341	524,449	365,049	100,761	32,449	3,697	56,901
2001	4,344,093	1,783,774	243,849	339,985	49,482	178,961	508,272	78,125	1,022	
2002	5,288,867	1,795,855	303,246	351,320	316,038	33,501	88,774	31,102		
2003	5,357,617	2,548,383	336,749	403,501	348,378	236,017	12,982			
2004	5,737,732	2,574,724	971,320	280,140	226,212	152,127				
2005	5,635,064	2,758,392	241,734	268,113	429,503					
2006	6,629,504	3,045,252	356,119	200,420						
2007	6,824,829	2,669,579	166,400							
2008	8,116,439	3,428,535								
2009	10,660,074									

Table 2.2: Full run-off triangle for paid out claims.

The most problematic stage in the bootstrap method is the formation of the pseudo-data. If the magnitudes of the incremental values differ significantly, it is quite likely that the values of simulated residuals (simulated from the initial set of residuals) are sufficiently high compared to the predicted incremental values to cause the negative values to appear in the (pseudo-)data due to the use of the inverse function. Most of the probability distributions used in loss reserving are non-negative (or positive) valued; thus, the problem with negative values in the (pseudo-)data can often appear. For example, in the case of the Poisson distribution, the negative incremental values are often replaced by zeros and in case of the gamma model they are replaced by ones in practice. This kind of replacement can of course cause the non-convergence of the parameters, but well chosen initial values help to avoid it. See Table 2.9 for an overview of the experienced negative values in the pseudo-data for each considered model with the given data set in Table 2.2.

Tables 2.3 – 2.8 present the point estimates ("Est. Reserve") along with the bootstrap prediction errors ("PE") as well as the upper limits ("Upper 95%") for a confidence level of 95%. To compare the behavior of the two discussed bootstrapping approaches, we have the last two lines of each table presenting the prediction errors and the upper confidence limits of the total reserve obtained by the PPE-method and the ratio of the results by the PPE-method and the bootstrap prediction error. We first have a look at the results obtained by the ODP model with using the Pearson residuals (see Table 2.3) and the Anscombe residuals (Table 2.4). The total reserve estimate with the given model is 13.4 millions and the prediction error using the uncorrected Pearson residuals is 1.95 millions. Table 2.3 shows that the prediction errors are not reflected significantly when the zero-correction or standardization is applied to the residuals, and consequently, also the upper limits are rather similar. The bootstrap prediction error decreases slightly if we use standardization with the zero-correction.

Year	Est. Reserve	Without Corrections		Zero-Correction		Zero-Correction & Stand.	
		PE	Upper 95%	PE	Upper 95%	PE	Upper 95%
2	50,795	93,020	203,813	92,998	203,777	92,216	202,490
3	57,836	100,596	223,316	100,731	223,538	100,100	222,501
4	120,028	138,277	347,494	138,874	348,476	137,976	346,999
5	348,993	223,848	717,223	225,245	719,521	223,718	717,009
6	552,215	275,089	1,004,736	275,921	1,006,105	273,860	1,002,715
7	1,024,516	379,154	1,648,224	379,513	1,648,815	377,637	1,645,729
8	1,406,289	443,167	2,135,299	444,117	2,136,861	442,360	2,133,971
9	2,283,616	582,104	3,241,177	585,913	3,247,443	581,658	3,240,443
10	7,560,816	1,254,499	9,624,467	1,256,388	9,627,574	1,244,243	9,607,596
Total	**13,405,108**	**1,959,079**	**16,627,793**	**1,962,403**	**16,633,261**	**1,939,728**	**16,595,961**
PPE		2,991,495	16,396,603	3,028,286	16,433,395	2,998,899	16,404,008
PPE/PE		1.527	0.986	1.543	0.988	1.546	0.988

Table 2.3: Over-dispersed Poisson model with Pearson residuals.

The prediction errors (PE) in the case of the Poisson model with Pearson residuals are varying from 1.94 million–1.96 million, depending on the residual adjustment, whereas in the case of the Anscombe residuals (see Table 2.4), the prediction errors vary from 1.74 million–1.94 million. This means that the 95% confidence limits for the total reserve prediction are between 16.60 million and 16.63 million in the case of the Pearson residuals and 16.27 million and 16.60 million in the case of the Anscombe residuals, given the Poisson model and residual adjustments. Thus, the Anscombe residuals are slightly more sensitive to the use of the residuals adjustments as the results vary more than with the Pearson residuals.

Year	Est. Reserve	Without Corrections		Zero-Correction		Zero-Correction & Stand.	
		PE	Upper 95%	PE	Upper 95%	PE	Upper 95%
2	50,795	88,545	196,452	90,094	199,000	91,999	202,133
3	57,836	96,181	216,054	98,013	219,067	98,917	220,554
4	120,028	133,040	338,879	135,019	342,134	137,301	345,888
5	348,993	214,176	701,313	216,583	705,272	225,633	720,159
6	552,215	262,405	983,871	264,390	987,137	273,456	1,002,050
7	1,024,516	358,928	1,614,953	362,319	1,620,531	378,666	1,647,422
8	1,406,289	418,820	2,095,248	422,350	2,101,055	445,735	2,139,523
9	2,283,616	545,624	3,181,167	550,420	3,189,057	580,020	3,237,749
10	7,560,816	1,118,001	9,399,928	1,124,329	9,410,337	1,239,100	9,599,136
Total	**13,405,108**	**1,743,656**	**16,273,422**	**1,772,161**	**16,320,313**	**1,941,261**	**16,598,482**
PPE		2,376,659	15,781,768	2,400,911	15,806,020	2,979,436	16,384,544
PPE/PE		1.363	0.970	1.355	0.968	1.535	0.987

Table 2.4: Over-dispersed Poisson model with Anscombe residuals.

The zero-corrected and standardized Anscombe residuals result in higher prediction errors. The uncorrected Anscombe residuals give the lowest pre-

diction errors. We see that the zero-correction do not effect the prediction errors significantly also in this case. The prediction error of the total reserve estimate obtained without any residual corrections with the Anscombe residuals are 11% smaller than with the Pearson residuals. The corresponding number with the zero-correction is 10%, but with the standardization and the zero-correction, the prediction errors of the total reserv estimate are about the same. We can see that the upper confidence limits for the total reserve are lower with the PPE-method (all of the corresponding ratios $\frac{PPE}{PE}$ are smaller than one). On the other hand, the prediction errors (depending on the residual adjustments) obtained by the PPE-method are higher than the estimates obtained by (2.20). The ratios seem to increase if we correct the residuals. In the case of Pearson residuals using the zero-correction and standardization, the corresponding ratio is the highest (1.546) as well as in the case of the Anscombe residuals (1.535).

Fitting the gamma model gives considerably higher reserve estimates (see Tables 2.5 – 2.6) compared to the results obtained by the ODP. The point estimate for the total reserve with the gamma model is 12.1 million, whereas with the Poisson model it was 13.4, indicating that the point estimate with the gamma model is 10.7% lower. In case of the gamma model and the Pearson residuals, the prediction errors for the total reserve vary from 5.77 million–6.06 million. Note that the zero-corrected and standardized residuals give the highest prediction errors, which was vice versa in the ODP model with the Pearson residuals. Also the prediction errors by the PPE-method are higher with the gamma model.

Year	Est. Reserve	Without Corrections		Zero-Correction		Zero-Correction & Stand.	
		PE	Upper 95%	PE	Upper 95%	PE	Upper 95%
2	50,011	44,643	123,449	45,671	125,140	46,180	125,977
3	37,118	31,996	89,751	32,803	91,079	33,134	91,623
4	93,432	55,698	185,055	56,962	187,134	57,583	188,156
5	332,152	177,861	624,733	180,928	629,779	182,393	632,188
6	454,013	214,667	807,140	220,470	816,686	222,762	820,456
7	782,168	368,402	1,388,189	377,559	1,403,253	381,636	1,409,959
8	1,031,663	494,109	1,844,472	504,639	1,861,794	510,817	1,871,957
9	2,090,954	1,127,346	3,945,438	1,153,719	3,988,822	1,172,064	4,018,999
10	7,270,704	5,522,714	16,355,569	5,694,928	16,638,861	5,771,588	16,764,966
Total	**12,142,220**	**5,767,157**	**21,629,193**	**5,970,660**	**21,963,956**	**6,056,343**	**22,104,904**
PPE		7,552,630	19,694,850	7,876,209	20,018,430	8,041,991	20,184,211
PPE/PE		1.310	0.911	1.319	0.911	1.328	0.913

Table 2.5: Gamma model with Pearson residuals.

In case of the gamma model and the Anscombe residuals, the prediction errors for the total reserve vary from 4.80 million–5.03 million. It seems that the

lower prediction errors are obtained with the Anscombe residuals. The upper limit for the total reserve in case of the gamma model reaches 22 million. We conclude, that based on the comparison of the Poisson and the gamma model in this particular data set, the latter gives us a smaller total reserve estimate, but higher prediction errors and, thus, higher upper limits for the reserve.

Year	Est. Reserve	Without Corrections		Zero-Correction		Zero-Correction & Stand.	
		PE	Upper 95%	PE	Upper 95%	PE	Upper 95%
2	50,011	39,307	114,671	39,667	115,263	40,032	115,864
3	37,118	28,228	83,553	28,454	83,925	28,706	84,339
4	93,432	49,537	174,920	50,035	175,740	50,478	176,468
5	332,152	161,253	597,413	163,026	600,330	164,015	601,957
6	454,013	192,684	770,978	196,031	776,484	197,565	779,007
7	782,168	324,763	1,316,403	330,583	1,325,977	333,394	1,330,601
8	1,031,663	424,982	1,730,758	433,923	1,745,466	438,507	1,753,007
9	2,090,954	943,237	3,642,579	962,403	3,674,107	976,113	3,696,660
10	7,270,704	4,614,429	14,861,440	4,752,683	15,088,868	4,810,173	15,183,439
Total	**12,142,220**	**4,808,814**	**20,052,719**	**4,970,282**	**20,318,334**	**5,033,233**	**20,421,888**
PPE		6,337,605	18,479,826	6,607,874	18,750,094	6,727,019	18,869,239
PPE/PE		1.318	0.9216	1.329	0.923	1.337	0.924

Table 2.6: Gamma model with Anscombe residuals.

In the case of both models, the PPE-method tends to give higher prediction errors than the bootstrap prediction errors. Likewise with the ODP model, also the gamma model with the uncorrected residuals result in lower prediction errors compared to adjusted residuals. We also note, that the $\frac{PPE}{PE}$ ratio of the corresponding upper limits are around one, indicating that similar confidence limits are obtained with both bootstrap methods.

From Tables 2.7 and 2.8, we can see the results of the lognormal model. The point estimate among all of the considered models is the highest with the lognormal model. The total reserve estimate using the lognormal model is 13.6 million and is 1.5% and 11% higher of the total reserve estimates obtained by the ODP and the gamma model, respectively. Note the high increase in the prediction errors, especially in the case of zero-corrected and standardized residuals. The prediction errors for the total reserve with the lognormal model with the Pearson residuals vary from 12 million–14 million, depending on the type of residual adjustment and the upper limits for the total reserve vary from 33.5 million–36.6 million; this shows a great difference of the estimates obtained by different models.

Year	Est. Reserve	Without Corrections		Zero-Correction		Zero-Correction & Stand.	
		PE	Upper 95%	PE	Upper 95%	PE	Upper 95%
2	54,060	85,331	194,429	90,471	202,885	96,159	212,242
3	46,399	71,200	163,523	75,678	170,889	80,440	178,723
4	101,016	98,142	262,460	102,990	270,435	108,580	279,630
5	271,424	198,419	597,823	207,885	613,395	216,101	626,910
6	442,472	285,962	912,879	298,890	934,146	307,842	948,872
7	756,516	486,210	1,556,331	498,810	1,577,058	513,128	1,600,612
8	1,031,985	696,039	2,176,969	703,776	2,189,697	726,645	2,227,316
9	2,255,719	1,813,279	5,238,563	1,889,497	5,363,942	1,979,344	5,511,740
10	8,658,523	12,584,162	29,359,469	13,392,437	30,689,082	14,340,838	32,249,202
Total	**13,618,118**	**12,073,216**	**33,478,558**	**12,957,863**	**34,933,803**	**13,976,414**	**36,609,319**
PPE		7,912,197	21,530,315	8,053,281	21,671,399	8,191,420	21,809,538
PPE/PE		0.655	0.643	0.621	0.620	0.586	0.596

Table 2.7: Lognormal model with Pearson residuals.

The Anscombe residuals lower the prediction errors roughly 40%. The prediction errors with the Anscombe residuals are between 7.2 million and 8.1 million; thus, the 95% confidence limit for the total reserve is between 25.5 million and 26.9 million, depending on the type of residual adjustment. However, higher values of the prediction errors should not be surprising, as the lognormal model is a more conservative model than, for example, the Poisson model or the gamma model. We see that the same pattern follows as before; if we use the zero-correction and standardization, then the prediction errors (and consequently, the upper limits, as well) are the highest. The lowest prediction errors are obtained with the uncorrected residuals. Note the change in the $\frac{PPE}{PE}$ ratios: the PPE-method gives lower prediction estimates than the regular bootstrap prediction error, which is not the case with the Poisson and the gamma model. Also the difference in obtained upper limits with the PPE-method is increased with the lognormal model.

Year	Est. Reserve	Without Corrections		Zero-Correction		Zero-Correction & Stand.	
		SEP	Upper 95%	SEP	Upper 95%	SEP	Upper 95%
2	54,060	60,881	154,209	62,416	156,734	64,501	160,164
3	46,399	51,105	130,467	52,185	132,243	53,947	135,142
4	101,016	73,835	222,475	75,545	225,288	77,599	228,666
5	271,424	161,902	537,753	164,606	542,201	167,711	547,309
6	442,472	240,304	837,772	245,296	845,984	249,108	852,255
7	756,516	403,916	1,420,958	409,625	1,430,349	416,062	1,440,938
8	1,031,985	549,315	1,935,608	558,675	1,951,005	569,359	1,968,581
9	2,255,719	1,329,510	4,442,763	1,368,829	4,507,443	1,410,281	4,575,631
10	8,658,523	8,317,733	22,341,194	8,632,691	22,859,300	9,013,141	23,485,140
Total	**13,618,118**	**7,227,323**	**25,507,064**	**7,611,566**	**26,139,144**	**8,062,255**	**26,880,527**
PPE		6,425,181	20,043,298	6,316,504	19,934,622	6,464,284	20,082,402
PPE/PE		0.889	0.786	0.830	0.763	0.802	0.747

Table 2.8: Lognormal model with Anscombe residuals.

We see from Tables 2.3 – 2.8 that on the 10th year, the estimated reserve is the highest and is approximately three times higher than the estimated reserve on the previous year. The reserve estimate on the 10th year makes nearly 56.7% of the total reserve estimate in the case of Poisson model, 60.3% in the case of gamma model and 64% in the case of lognormal model, which is the highest percentage. This high proportion of the total reserve estimate on one particular year can be explained by having a look at the initial data set, Table 2.2, where we see that on the last year, 2009, we have the largest value in the whole data set.

We can draw four main conclusions from analyzing this particular data set:

1. The gamma model produces the lowest estimated claim reserve, and the lognormal model produces the highest estimated claim reserves. The corresponding figures of the gamma model are not that different from the ODP model.

2. The standard errors of prediction are quite different and consequently the estimated upper limits. These differences tend to be greater especially on the first years, since estimations are based on few predictions. The highest prediction errors are produced by the lognormal model, and the lowest prediction errors were obtained by the over-dispersed Poisson model.

3. With this particular data set, the prediction errors are the lowest with the Anscombe residuals. In general, no matter which residual of the two is used, the highest prediction errors are obtained by using the zero-correction with standardization.

4. When comparing the two bootstrap procedures, we can conclude that using the (alternative) PPE-method, the upper confidence limits for the total reserve are lower with each considered model.

As we mentioned beforehand the possible problem associated with the negative values in the pseudo-data, we present Table 2.9, which gives an overview of the amount of the negative values appearing in the procedure of creating a pseudo-data in the case of 10,000 iterations. Roughly speaking, we observe that with the Poisson models 1–2, negative incremental pseudo values appeared with every iteration step. This is rather expected since the incremental values in the data differ largely. Note that using the Pearson residual caused more negative values than using the Anscombe residuals. There were no negative values in the pseudo-data in the case of the gamma model and the Anscombe

residuals. We do not present the appeared negative values of lognormal model, as its non-negative incremental pseudo values are obvious.

Model	Residual	Type of Adjustment		
		1	2	3
Poisson	Pearson	22,557	23,184	22,306
	Anscombe	11,632	12,172	12,110
Gamma	Pearson	20,112	21,315	21,315
	Anscombe	0	0	0

Table 2.9: Amount of negative values appearing in the pseudo-data during 10,000 iterations. 1 - without corrections; 2 - zero-correction; 3 - zero-correction & standardization.

2.4 Comparative analysis with the Schedule P database

In the previous section, we implemented different stochastic reserving methods on a real-life insurance data and we assessed the impact of the considered predictive models and residuals. Apart from the analytical perspective of the methods, it is also essential to compare and rank competing forecasting methods. Academic contributions in reserving context often lack a thorough comparative study to evaluate and compare the performance of the models. In most of the comparative studies, up to our knowledge, the chain-ladder mean is often kept as a benchmark, but this should not be the only criteria when deciding which model is the best (or the most precise). There are enough statistical tools and methods to measure the prediction accuracy, we conduct a model validation with a specific scoring rule.

2.4.1 Schedule P database

We apply the defined models, residuals and their adjustments to the run-off triangles from practice. We use data sets of paid net loss triangles from the Schedule P – Analysis of Losses and Loss Expenses in the National Association of Insurance Commissioners (NAIC) database, which is available on the website of the Casualty Actuarial Society [1]. The data include major

[1] `http://www.casact.org/research/index.cfm?fa=loss_reserves_data`

personal and commercial lines of business from U.S. property and casualty insurers. The database contains data on six lines of business, and we chose to use (1) *workers' compensation*; (2) *commercial auto/truck liability/medical* and (3) *private passenger auto liability/medical*. The triangle data correspond to the claims of the accident years 1988–1997 with a 10-year development lag. Not all of the data sets there were applicable; some of them contained too many negative values in the upper triangle, which lead to a problem in the parameter estimation procedure with the given models, and many triangles contained a high number of zeros both in the upper and the lower triangle. Thus, we had to carefully extract the data sets, which would fulfill the requirements of the models' assumptions. Both upper and lower triangles are included, so that we can use the data to test the models' performance retrospectively, i.e., the validation process is based on the back-testing idea, and all of the methods provide reserve estimates by predicting in the same lower triangle. We extracted the full triangles (from each line of business) with the following identifiers: (1) 337; 1767; 2135; 2712; 7080; 8672; 34576; 21172; 18767; 14176; (2) 353; 388; 1767; 2003; 2135; 2712; 26077; 26433; (3) 620; 1767; 2003; 7080; 25275; 33499; 34592. Anyone interested could easily find these chosen data sets from the corresponding website.

2.4.2 Model validation

This subsection describes the validation process for the three methods discussed in Section 2.1 in combination with the possible residual definitions in the bootstrap procedure discussed in Section 2.2. We consider the scoring rule to measure the accuracy of probabilistic predictions. There are many scoring rules available to apply, including entire parametrized families of proper scoring rules. In accordance with the Dawid (1984) prequential principle, the evaluation of probabilistic forecasts is required to be based only on the predictive distributions and the observations.

Scoring rules provide summary measures of predictive performances, by assigning numerical scores to the probabilistic forecasts and on the value that materializes. Sharpness and calibration are combined here in one measure. Sharpness refers to the concentration of the predictive distributions and is a property of the forecasts only. The less variability in the predictions, the more concentrated the predictive distributions are. Consequently, forecasts will be more sharper, and subject to the calibration, the sharper the forecasts the better. Following Gneiting et al. (2007), we denote $s(P, x)$ as the assigned score for the issued predictive distribution P and materialized observation x

drawn from Q. We take scores to be penalties that the forecaster wishes to minimize. A scoring rule is proper if the expected value of the penalty $s(P, x)$ for an observation x is minimized if $P = Q$. We talk about a strictly proper scoring rule if the minimum is unique. For an introduction to scoring rules, we refer to Gneiting and Raftery (2007); Riebler et al. (2012).

Czado et al. (2009) study the tools for the evaluation of probabilistic forecasts for count data. We consider the scoring rule that depends on the predictive distribution P only through the first and the second moment. This type of proper scoring rules were studied by Dawid and Sebastiani (1999). In this thesis, we use the Dawid–Sebastiani scoring rule (DSS), which is defined as:

$$\text{DSS} = \left(\frac{x - \mu_P}{\sigma_P}\right)^2 + 2\ln(\sigma_P),$$

where x is the observation that realizes, μ_P is the mean and σ_P is the standard deviation of the predictive distribution. To assess the predictive performance of each model with different residual adjustments discussed in this thesis, we obtain an overall performance measure by averaging the DSS scores over all of the cells in the lower triangle and over each considered data set. Let $k = 1, \ldots, d_s$ denote the number of data sets used. Then, the DDS scoring rule specifies to:

$$\text{DSS} = \frac{1}{d_s}\sum_{k=1}^{d_s}\sum_{(i,j)\in\nabla}\left(\left(\frac{(C_{ij})_k - (\widehat{C}_{ij})_k}{(\widehat{\sigma}_{ij})_k}\right)^2 + 2\ln((\widehat{\sigma}_{ij})_k)\right), \tag{2.21}$$

where $C_{ij},\ (i,j) \in \nabla$ denote the cells in the lower triangle, i.e., the observation that realizes (true observation in the lower triangle), $\widehat{C}_{ij}$ is the estimate of the corresponding mean obtained by the predictive model and $\widehat{\sigma}_{ij}$ is the estimate of the corresponding standard deviation. Under the over-dispersed Poisson model the mean estimate $\widehat{C}_{ij}$ is given by (2.10) and the standard deviation is estimated by the bootstrap prediction error for cell C_{ij}:

$$\widehat{\sigma}_{ij} = \text{PE}_{bs}(\widehat{C}_{ij}) = \sqrt{\widehat{\phi}\widehat{\mu}_{ij} + (\text{SE}_{bs}(\widehat{C}_{ij}))^2}.$$

Likewise, under the gamma model the mean estimate $\widehat{C}_{ij}$ is given by (2.14) and the standard deviation is estimated by:

$$\widehat{\sigma}_{ij} = \text{PE}_{bs}(\widehat{C}_{ij}) = \sqrt{\widehat{\phi}\widehat{\mu}_{ij}^2 + (\text{SE}_{bs}(\widehat{C}_{ij}))^2},$$

and under the log-normal model the mean estimate $\widehat{C}_{ij}$ is given by (2.13) and the standard deviation is estimated by:

$$\widehat{\sigma}_{ij} = \text{PE}_{bs}(\widehat{C}_{ij}) = \sqrt{(\exp(\widehat{\sigma}^2) - 1)\exp(2\widehat{\mu}_{ij} + \widehat{\sigma}^2) + (\text{SE}_{bs}(\widehat{C}_{ij}))^2},$$

The first term in (2.21) focuses on calibration and the second term on sharpness. The DSS score therefore measures the quality of the predictive distributions as a trade-off between calibration and sharpness.

2.4.3 Results

In this section, we present the model assessment results obtained by the scoring rule (2.21). In Tables 2.10 – 2.12 below, we present the overall performance measure for the over-dispersed Poisson model, the gamma model and the lognormal model with the considered residuals adjustments, i.e., using residuals without corrections, the zero-corrected residuals and then the zero-corrected residuals with taking into account the influence of the observation (i.e., using the standardization). Thus, we have 18 different setups and combinations.

Table 2.10 presents the model scores obtained using the workers' compensation data. The ODP model fits clearly data the best, but the scores of the lognormal model are quite the same. The gamma model is fitting the data rather poorly compared to the ODP and the lognormal model. The use of zero-corrected and standardized residuals give the best results with all three models. The differences are small, but the highest values of the scores are obtained with the uncorrected residuals. The smallest score is obtained using the zero-corrected and standardized Pearson residuals. In case of the ODP and the log normal model, results by the Anscombe and the Pearson residuals are nearly the same, the biggest difference in the given residuals is seen with the gamma model.

Model	**Residual**	**Type of Adjustment**		
		1	**2**	**3**
ODP	Pearson	**13.76**	**13.74**	**13.75**
	Anscombe	13.98	13.97	13.96
Gamma	Pearson	22.30	21.98	21.59
	Anscombe	25.11	24.58	24.47
Lognormal	Pearson	14.55	14.53	14.53
	Anscombe	14.60	14.58	14.58

Table 2.10: Model validation on workers' compensation data using the Dawid–Sebastiani scoring rule (DSS). The three lowest scores are indicated in bold. 1 - without corrections; 2 - zero-correction; 3 - zero-correction & standardization.

Table 2.11 describes the model assessment results when the data from the business line (2) was used. Also in the case of this data, the ODP model remains the best model, no matter which residual or its adjustment is used. Note that the use of Pearson residuals gives still more precise predictions than the use of Anscombe residuals, only in case of the ODP model the score measures obtained by the Pearson and the Anscombe residuals are very close. Likewise with the workers' compensation data, the lognormal model fits the data better than the gamma model and all three models predict with the zero-corrected and standardized residuals the most precisely.

Model	**Residual**	**Type of Adjustment**		
		1	**2**	**3**
ODP	Pearson	**12.59**	**12.58**	**12.58**
	Anscombe	12.62	12.61	12.62
Gamma	Pearson	59.75	58.49	57.84
	Anscombe	73.86	72.16	71.96
Lognormal	Pearson	26.27	25.74	25.32
	Anscombe	30.91	30.38	30.02

Table 2.11: Model validation on commercial auto/truck liability/medical data using the Dawid–Sebastiani scoring rule (DSS). The three lowest scores are indicated in bold. 1 - without corrections; 2 - zero-correction; 3 - zero-correction & standardization.

The business line (3) does not bring any surprises: the ODP model continues to being the best model in the sense of the score. We have used the data from three different lines of business and the considered models have performed in a similar manner with each given line of business.

Model	Residual	Type of Adjustment		
		1	**2**	**3**
ODP	Pearson	**14.71**	**14.70**	**14.72**
	Anscombe	14.77	14.76	14.77
Gamma	Pearson	25.52	25.10	24.99
	Anscombe	28.44	28.02	27.88
Lognormal	Pearson	17.43	17.38	17.32
	Anscombe	17.92	17.87	17.83

Table 2.12: Model validation on private passenger auto liability/medical data using the Dawid–Sebastiani scoring rule (DSS). The three lowest scores are indicated in bold. 1 - without corrections; 2 - zero-correction; 3 - zero-correction & standardization.

As we are interested in which model minimizes the score the most, we pointed out in bold in the tables three setups with the smallest numerical value; see Tables 2.10 – 2.12. In general, we can draw the following conclusions from validating the considered models:

1. The over-dispersed Poisson model fits the data best. This confirms the results obtained in the previous section, where we obtained the smallest prediction errors precisely with the ODP model.

2. Surprisingly, the ODP model and the lognormal model are behaving rather similarly. The gamma model is fitting the data slightly worse.

3. With the gamma and the lognormal model, the lowest values of the scores are obtained with the zero-corrected and standardized residuals, however the differences are small. In Section 2.3, the highest prediction errors were given by the models with the zero-corrected and standardized residuals.

4. Overall, the smallest score was obtained using the zero-corrected Pearson residuals. The ODP model fits different insurance data the best.

5. If comparing just the choice of residuals, we see that the Anscombe residuals perform slightly worse than the Pearson residuals, but in case of the ODP and the lognormal model, the score measures of both residuals are quite close.

In Section 2.3, the results showed that the lowest prediction errors were obtained with the Anscombe residual and its score measures were rather comparable with the Pearson residuals in this section. The highest values of the score measures are obtained strictly with the uncorrected residuals. Recall that in Section 2.3, we saw that the lowest prediction errors were obtained specifically by the uncorrected residuals, with the given data set of an Estonian insurer. This is a good example to show that in the comparative study, the focus should not be only on which model gives the lowest errors, but which method actually fits the data the best. An actuary has to be always ready to use his/her own expertise and experience in addition to well-known or most-used models in the estimation problems, as every data set is different. We considered 18 different setups in the model validation, and as we wanted to rank the models, then we rank the first three models based on the used scoring rule: the ODP model with the Pearson residual with the zero-correction and standardization; the ODP model with the Pearson residuals with the zero-correction and the ODP model with the Pearson residual without any corrections. In general, this model validation section shows clearly that the Anscombe residuals could be considered as an alternative to the Pearson residuals.

According to this case study and comparative analysis, we can say that given the obtained results in Sections 2.3 and 2.4, the method that gives the lowest prediction errors should not be confused with being the model that fits the best. Like in our case study, methods that result in the lowest variability may be, for instance, suffering from the underestimation and may not fit the data after all. We considered only one scoring rule, but more investigation is required in the model validation part.

2.5 Conclusions

In this chapter, we studied the impact of the methods and the residuals on the reserve estimates and their predictive distributions. Caution is necessary when dealing with the latest development periods of the earlier accident years. The residuals of the tail are often volatile, and adjustment is hence required if they are used in the bootstrapping process. Therefore, we implemented and compared the (over-dispersed) Poisson, the gamma and the lognormal distributions in combination with the residual adjustments in the bootstrapping framework. We saw that the (over-dispersed) Poisson model and the gamma model tend to give similar point estimates, as expected, but there are bigger differences in the estimates of the prediction errors. In our

case study (Section 2.3), we obtained the highest prediction errors using the standardized residuals with zero-correction, but in the model validation (Section 2.4), we saw that the use of zero-corrected and standardized residuals lead to the lowest score with the gamma and the lognormal model. The ODP model works the best with the zero-corrected residuals. In general, based on the case study and the model validation part, we could conclude that the over-dispersed Poisson model with the Pearson residual fits the data the best, and it seems to not matter too much, which adjustments is applied in the residuals. The Poisson model with the Pearson or the Anscombe residuals appears to be a good choice in the sense that it yields the most reliable results based on the scoring rule. The gamma model shows the poorest fit and the lognormal model should be used by a more conservative scientist.

We can conclude that there are many different possibilities that we have to take into account before applying the bootstrap method as the prediction errors obtained by using different combinations of possible options are quite different. It is up to an actuary which result should be taken into account when making decisions in setting up the fund for reserves. The choice of a particular model remains the main struggle, but based on our research, we can draw the following conclusions:

- The large fluctuation of the values in the data substantiates the use of the over-dispersed Poisson model. The gamma model and the ODP model tend to give similar point estimates, whereas the lognormal model produces the highest estimated claim reserves. Here, the expertise of an actuary would help to finalize a decision in model selection, depending on the company's balance of hazard and conservatism.

- When the emphasis is on prediction errors, then the ODP model should be used for the lowest prediction errors. The lognormal model tends to give much higher errors.

- The choice of residuals matters in bootstrapping. The Pearson residual could be preferred, but in some cases (see Tables 2.10 – 2.12), the Anscombe residual could be considered as well.

- The adjustment of residuals is as important as the choice of the residual; the most precise predictions are obtained with zero-corrected and standardized residuals if the gamma or the lognormal model is applied, but they also bring an increase in the prediction errors. The ODP predicts the best with the zero-corrected residuals.

- Different types of claims data require different methods. Understanding the data and using professional expertise are crucial in claims reserving.
- It is important to realize that each method is appropriate only under certain conditions, and that a thorough understanding of a variety of models and their disparities leads to more flexibility in handling different situations.
- The proposed model validation and assessment ideas are generic and do not depend on a particular data set, thus constituting a useful tool in reserve estimation.

The analysis between the estimates and the actual future payments has to be carried out by the expert in the long run, in order to validate the functionality of a reserving method and identify any needed modifications. It is contended in England and Verrall (2002) that the effectiveness of a particular reserving method and modeling can be completely tested only with an extensive case study with data from various lines of business and companies. Then, the estimated results are compared with the claims development over time, and only then, we can get closer to the best choice of the reserving models. Comparative studies could have a higher value when the model validation is included in the analysis. The model assessment should become a default procedure when deciding on (reserving) models. In this thesis, we considered only one scoring rule, but other statistical approaches for the model assessment could be considered in the future.

Chapter 3

Loss reserving on different levels of data aggregation

Recently there have been many proposals of reserving models that are based on individual level claims data. Organizing claims data in discrete time is just an approximation to the real situation, where data are recorded in continuous time. In recent developments, an extension towards the continuous use of individual claims data in the chain-ladder framework has been proposed. The classical chain-ladder method (CL) is regularly applied to annual data, but the question arises whether the reserve estimates based on granular (daily or monthly) data outperform results obtained by annual and quarterly data. We investigate whether and how much different aggregation levels of claim data can improve the reserving process and compare the performances of the CL method and the continuous chain-ladder method (CCL) on each level of data aggregation (daily, monthly, quarterly and annual) in simulation study. To further investigate the impact of aggregation, we present an empirical study with the application of discussed models using an insurance data. Finally, our analysis shows that the point estimation precision increases if the basic chain-ladder method is applied to more granular data and results are comparable with the continuous model approach.

This chapter is a considerably extended version of Tee, L. and M. Käärik (2017). Loss reserving on different levels of data aggregation: chain-ladder vs. its continuous extension. *Submitted*.
In comparison with the original publication, we have adjusted descriptive sections, added a number of graphs and figures, added many corrections and conducted an additional simulation study.

3.1 Loss reserving: aggregate vs. individual data model

The access to micro-level information in insurance companies allows to go more granular with the data, i.e., daily, weekly or monthly data could be used instead of annual data. It could be that the micro-level methods perform better, but it remains an open question whether it is worth of all the extra effort. However, using a solid statistical framework, but keeping the model straightforward could be appealing to practitioners. The paper of Martínez Miranda et al. (2013) reformulates the well-known chain-ladder method as a histogram type approach and improves the well-known classical technique by replacing the histogram by a kernel smoother, leading to a continuous chain-ladder method. The model assumes that the data have not been aggregated and it uses data recorded in continuous time, see Table 3.1 and Figure 3.1 for explanatory examples. It is clear from Figure 3.1 that the data are arranged in a two dimensional space, but the data still forms a triangle. When the data are not grouped, it can be treated as having a density on the triangle.

Origin	Development year				
year	1	2	3	4	5
1	147	79	56	44	21
2	166	120	69	10	
3	202	142	82		
4	251	170			
5	310				

Table 3.1: Aggregated data.

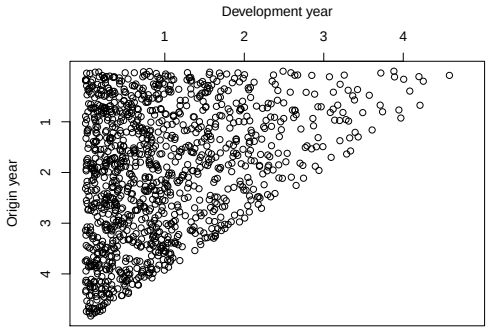

Figure 3.1: Individual claim data.

3.1.1 Estimating the density in the observation triangle

We consider a random sample of m i.i.d. observations $\{(X_i, Y_i), i = 1, \ldots, m\}$ from a population (X, Y) having a density f with support on a subset $\mathscr{S} = \{(x, y)|x, y \geq 0, x + y \leq T\}$ of the whole square $\mathcal{S} = \{(x, y) \mid 0 \leq x, y \leq T\}$, where x is the accident or the underwriting time and y is the claims development time within the time frame $[0, T], T > 0$. The CCL method is

developed for the individual (or continuous) claims data context, but can still be applied to aggregated data as well. In this chapter, we consider only claim counts in order to keep the density approach straightforward and we assume, without loss of generality, that the data are available as a triangle. Then the aggregated reported counts triangle can be written by $\{C_{ij} : (i,j) \in \Delta\}$, where the business has been observed during n time periods, i.e. $\Delta = \{(i,j)|\ i = 1,\dots,n, j = 1,\dots,n-i+1\}$ and C_{ij} is the total number of claims of insurance incurred in period i, which are reported in period $i+j-1$, i.e. with $j-1$ periods delay from year i. Clearly, i denotes the origin period and j the delay period. If we assume that the claims are settled within the n observed time periods, then the aim is to predict the sum of the delayed claim counts in the lower, unobserved future triangle $\{C_{ij} : (i,j) \in \nabla\}$, where $\nabla = \{(i,j)|\ i = 2,\dots,n; j = n-i+2,\dots,n\}$. See Tee et al. (2017) for more detailed explanation concerning the notation.

In the context of the CCL method, the reserving problem is reformulated in terms of a multivariate density estimation problem, where the target function is a continuous two-dimensional density function. One of the simplest density estimators is a histogram, which can handle the underlying distribution of continuous data. The classical CL model can be described as a histogram of the granular data projected on a multiplicative structure for forecasting the future. The densities in the underwriting and development directions are piece-wise constant. Considering a histogram estimator of the density in the triangle gives an insight for a regression view of the density estimation problem. There are several contributions that have described the connection between the density and the regression problems, moreover, the classical CL method clearly approaches the density through regression. The classical CL method estimates a two-dimensional density f supported in the triangle Δ using a multiplicative structure. Technically speaking, an equally-spaced grid points are set up which define the bins (contiguous intervals) such that the support of f is contained in defined bins, and the bin centers z_{ij} and the bin counts C_{ij} (the number of data falling in the corresponding interval) are constructed. Figure 3.2 depicts the idea of describing the CL method as a histogram of the granular data. Thus, when considering aggregated claim count data C_{ij} in the form of a run-off triangle, the classical chain-ladder method approaches the density problem through the regression model

$$C_{ij} = r(z_{ij}) + \varepsilon_{ij},$$

based on the data $\{(z_{ij}, C_{ij}), i,j \in \Delta\}$, where $z_{ij} = (x_i, y_j)$ are the points in the (run-off) grid. The points z_{ij} are the coordinates of the data point C_{ij} and we assume that C_{ij} is a middle point of the corresponding area

(of a data cell in a triangle), see Figure 3.3 for a clarifying example. The regression function $r(z_{ij}) = m\Lambda^2 f(z_{ij})$ is defined through the bin length Λ in the grid and $m = \sum_{(i,j)\in\Delta} C_{ij}$. We refer for the detailed explanation to Martínez Miranda et al. (2013) and Fan and Gijbels (1996).

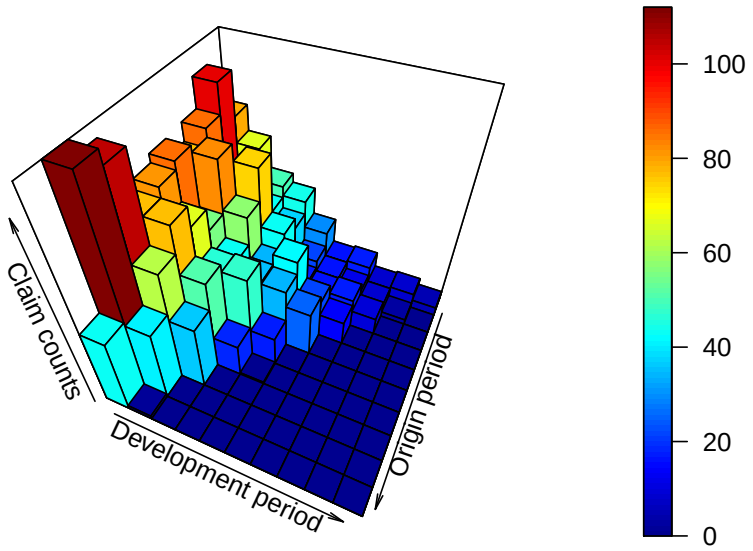

Figure 3.2: The CL method described as a histogram of the granular data projected on a multiplicative structure.

The assumption of the CCL method for forecasting the target density in the future is a multiplicative structure, i.e. $f(x, y) = f_1(x) \cdot f_2(y)$ for origin period x and development period y. Since histogram leads to discrete time effects, the continuous approach improves on histogram with a kernel estimator and assumes continuous densities. The general estimator at any point $(x_0, y_0) \in \mathscr{S}$ using the kernel density estimator is defined as follows

$$\widehat{f}(x_0, y_0; h) = m^{-1} \sum_{i=1}^{m} K_{h_1}(X_i - x_0) K_{h_2}(Y_i - y_0),$$

where $K_{h_1}(x) = h_1^{-1} K\left(\frac{x}{h_1}\right)$ and $K_{h_2}(y) = h_2^{-1} K\left(\frac{y}{h_2}\right)$ are one-dimensional kernels. Here K is a function satisfying $\int K(x)dx = 1$ and bandwidth parameters $h = (h_1, h_2) \in \mathbf{R}_+^2$, where h_1 controls the degree of smoothing in the origin direction and h_2 controls the development direction. Usually K is chosen to be a unimodal probability density function that is symmetric about zero, ensuring that $\widehat{f}(x_0, y_0; h)$ is also a density itself. For a general

introduction to kernel smoothing, we refer to Wand and Jones (1994). Kernel smoothing is a natural way to improve on histograms, however, it brings some issues to tackle which will be discussed in the following.

The boundary problem. The simple kernel methods suffer from the well-known boundary problems. The kernel density estimator does not take into account the potential finite support of the variables. Given that we have non-negative data and the model aims to estimate two-dimensional density f based on the observed triangle, the boundary bias problem becomes severe as the boundary region increases with the dimension of the support. This leads to a bias in the boundary region, since the support of variables is bounded and the standard kernel estimator continues to give weight outside the supports. As clarified in Martínez Miranda et al. (2013), reserving could be seen as a density estimation problem, where the claims triangle defines the boundary region. There have been proposed several solutions to this problem. In the nonparametric regression context, the problem of boundary bias for multivariate data is developed by Staniswalis and Messer (1996) and Bouezmarni and Rombouts (2010), among others. Also Müller and Stadtmüller (1999) proposes boundary kernels for multivariate data defined on arbitrary support. The local linear estimator introduced in Nielsen (1999) proposes a non-parametric multivariate density estimation with arbitrary boundary regions. Nielsen's local linear estimator at each point $(x_0, y_0) \in \mathscr{S}$ is obtained by solving the following minimization problem:

$$\begin{aligned} \widehat{\Theta} &= \arg\min_{\Theta} \lim_{b\to 0} \int_{\mathscr{S}} \left\{ \widetilde{f}_b(z) - \Theta_0 - \Theta_{1,1}(x - x_0) - \Theta_{1,2}(y - y_0) \right\}^2 \\ &\times K_{h_1}(x - x_0) K_{h_2}(y - y_0) dz \end{aligned} \tag{3.1}$$

where $\Theta = (\Theta_0, \Theta_{1,1}, \Theta_{1,2})$, $\widehat{\Theta} = (\widehat{\Theta}_0, \widehat{\Theta}_{1,1}, \widehat{\Theta}_{1,2})$ and $\widetilde{f}_b(z) = m^{-1} \sum_{i=1}^{m} K_{b_1}(X_i - x) K_{b_2}(Y_i - y)$ is the standard kernel estimator at the point $z = (x, y) \in \mathscr{S}$ with bandwidth parameters $b = (b_1, b_2) \in \mathbf{R}_+^2$. The local linear density of $f(x, y)$ is then given by

$$\widehat{f}(x, y; h) = \widehat{\Theta}_0, \tag{3.2}$$

where $h = (h_1, h_2) \in \mathbf{R}_+^2$.

The data in an insurance company is usually aggregated at some level (such as on daily, weekly, monthly, quarterly or yearly basis) and the aim of this research is to compare the performances of the classical chain-ladder method and the continuous chain-ladder method on different levels of data aggregation. Therefore, for practical situations the rewritten version of the local linear estimator (3.1) using the regression formulation for aggregated data is given

below. The regression formulation for aggregated data is closely related to the marginal regression method of Nielsen and Linton (1998). Recall that $z_{ij} = (x_i, y_j)$ are the points in the run-off grid, i.e. z_{ij} are the coordinates of the data point C_{ij}, and we assume that C_{ij} is a middle point of the corresponding data cell in a triangle. Since we want to estimate density in any random point in the upper triangle, we need to define the coordinates of the corresponding points. For density estimation in a discrete framework we have to extend the run-off grid to estimation grid Δ_{grid} with the points where the density will be estimated at. The estimation grid Δ_{grid} with its points are defined as follows:

$$\Delta_{grid} = \left\{ (d_k, d_l) \mid d_k = \Lambda\left(k - \frac{1}{2}\right), d_l = \Lambda\left(l - \frac{1}{2}\right), k = 1, \ldots, \frac{n}{\Lambda}, \right.$$
$$\left. l = 1, \ldots, \frac{1}{\Lambda}(n - \lfloor \Lambda(k-1) \rfloor) \right\}.$$

The points d_{kl} in the estimation grid Δ_{grid} can be set as tightly as preferred, but one has to take into account the highly increasing computational time, for that reason we assume $\Lambda = \frac{1}{2}$ in the following. It is clear that for the points $z_{ij} = (x_i, y_j)$ in the run-off grid, we have $x_i = i - \frac{1}{2}, y_j = j - \frac{1}{2}$, $i, j = 1, \ldots, n$. We are using monthly data with the CCL method in Sections 3.3 – 3.5 and for illustrative explanation see Figure 3.3 below, where the length of period (in our case, months) is chosen $n = 4$.

The local linear estimator for the density f (with a support in a triangle) for any given point $d_{kl} = (d_k, d_l) \in \Delta_{grid}$ can be derived by solving the following minimization problem:

$$\begin{aligned} \widehat{\Theta} &= \arg\min_{\Theta} \sum_{(i,j)\in\nabla} \{C_{ij} - \Theta_0 - \Theta_{1,1}(x_i - d_k) - \Theta_{1,2}(y_j - d_l)\}^2 \quad (3.3) \\ &\times K_{h_1}(x_i - d_k) K_{h_2}(y_j - d_l). \end{aligned}$$

Here the solution $\widehat{\Theta}_0$ gives an estimator for $r(d_{kl})$. Then the density $f(d_{kl})$ can be estimated by

$$\widetilde{f}(d_{kl}) = \frac{\widehat{r}(d_{kl})}{m\Lambda^2}, \quad d_{kl} \in \Delta_{grid}. \tag{3.4}$$

Note the square of the bin length Λ in the formula – since we are estimating two-dimensional density, bin lengths are considered in both directions (origin and development direction).

The bandwidth selection problem. The kernel density estimator requires the specification of the bandwidth h as the choice of smoothing parameter

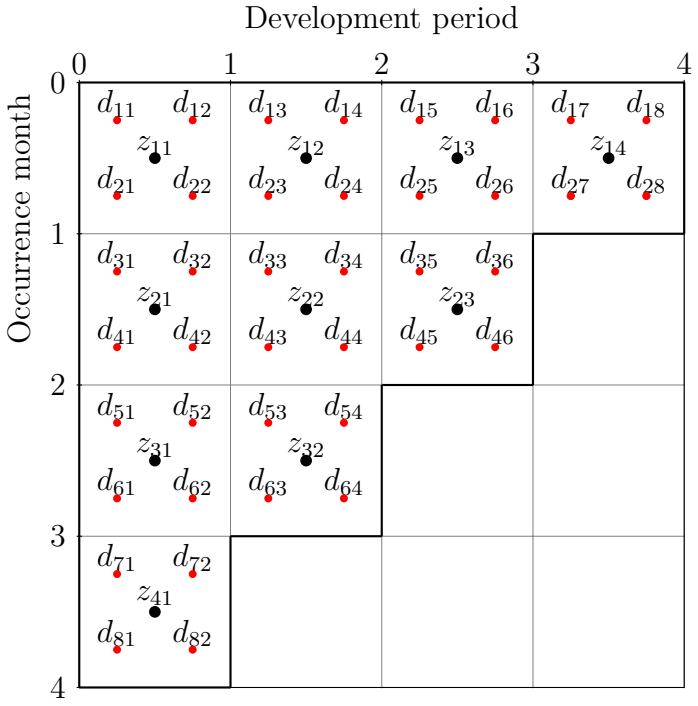

Figure 3.3: The grid layout. The black dots denoted by z_{ij} represent the data points in the centers of each cell in the run-off triangle. The red dots denoted by d_{kl} represent the grid points, at which the density is being estimated.

determines a trade-off between bias and variance. Bandwidth selection controls the smoothness or roughness of a density estimate and bears danger of under- or oversmoothing. There are many occasions where the bandwidth could be chosen by eye, but that approach requires prior knowledge about the structure of the data and a subjective choice can be very time consuming. Often the bandwidth selection problem is solved by automatic selection. Among the earliest automatic and consistent bandwidth selectors were those based on cross-validation ideas. The most commonly used cross-validation method (see for example Wand and Jones (1994)) is also considered in Martínez Miranda et al. (2013), where the data-driven *least square cross-validation* score is defined by

$$\mathrm{LSCV}(h) = \int_{\mathscr{S}} \widehat{f}(x,y;h)^2 dz - 2\sum_{i=1}^{m} \int_{\mathscr{S}} \widehat{f}^{[-i]}(x,y;h) d\widetilde{F}_n(x,y) \tag{3.5}$$

with $\widehat{f}^{[-i]}(x,y;h)$ being the leave-one-out version of the estimator $\widehat{f}(x,y;h)$ and $\widetilde{F}_n$ being the empirical distribution function from the sample. Therefore it

is reasonable to choose h which minimizes LSCV(h) and we denote the chosen bandwidth by $\widehat{h}_{\text{LSCV}}$. It was recommended in Mammen et al. (2015) to also consider *integrated squared error* (ISE) for finding the optimal bandwidths in this framework, but we have limited ourselves to using only cross-validation score (3.5) in this study. Clearly the choice of the smoothing parameter is a central issue in kernel smoothing and it is still a burgeoning area of research. There is a great deal of literature on choosing bandwidths under various conditions (for a comprehensive description, see Scott (2015) or Wand and Jones (1994)), but the development of a bandwidth selector is beyond the scope of this thesis.

The choice of kernel. The choice of the shape of the kernel function is not a particularly important compared to the choice of value for the bandwidth. Kernel density estimators are sensitive to the choice of bandwidth, but the choice of a kernel function does not usually affect the results considerably. Most unimodal kernels perform about the same as every other kernel and the choice between kernels can be made on other grounds such as computational efficiency. The CCL model assumes a multiplicative structure, thus multiplicative kernels are considered. The Epanechnikov kernel is often the default kernel function in the statistical software if no other kernel is specified. The perfomance of kernel is measured by MISE (mean integrated squared error) or AMISE (asymptotic MISE) and the Epanechnikov kernel is the most efficient kernel in minimizing MISE and AMISE. We bring out various popular kernels in multiplicative form and we present kernels' efficiency relative to the Epanechnikov kernel, see Table 3.2. All the kernel functions are derived for the multiplicative structure. The efficiency is a value that represents the ratio of sample sizes necessary to obtain the same minimum AMISE (for a given density) with a chosen kernel as when using the Epanechnikov kernel.

Kernel	Formula	Efficiency
Epanechnikov	$K(x,y) = \left(\frac{3}{4}\right)^2 (1-x^2)(1-y^2)\mathbb{1}_{\{\lvert x\rvert<1 \wedge \lvert y\rvert<1\}}$	1.000
Biweight	$K(x,y) = \left(\frac{15}{16}\right)^2 (1-x^2)^2(1-y^2)^2\mathbb{1}_{\{\lvert x\rvert<1 \wedge \lvert y\rvert<1\}}$	0.994
Triweight	$K(x,y) = \left(\frac{35}{32}\right)^2 (1-x^2)^3(1-y^2)^3\mathbb{1}_{\{\lvert x\rvert<1 \wedge \lvert y\rvert<1\}}$	0.987
Gaussian	$K(x,y) = \left(\frac{1}{\sqrt{2\pi}}\right)^2 e^{-\frac{1}{2}(x^2+y^2)}$	0.951
Uniform	$K(x,y) = \frac{1}{4}\mathbb{1}_{\{\lvert x\rvert<1 \wedge \lvert y\rvert<1\}}$	0.930

Table 3.2: Efficiencies of several kernels compared to the Epanechnikov kernel.

3.1.2 Estimating the density in the whole square

In the previous subsection the methodology of the CCL method was explained. The first part of the method included estimating the two-dimensional density in the observed run-off triangle. In this section, a technique to obtain the estimates of the IBNR claim counts in the lower triangle is given.

The IBNR claim counts in the framework of the CCL method are obtained by integrating a two-dimensional density, i.e. extrapolating the estimated density (3.2) (or (3.4)) to the full support (or to the full triangle) allows us to forecast the future. More precisely, the marginal integration method introduced by Linton and Nielsen (1995) is extended to the density estimation problem. If the data is assumed to be continuous, the two-dimensional density in the observation set $\mathscr{S}$ should be estimated by an estimator (3.2). Then, assume the target density to be multiplicative, $f(x,y) = f_1(x)f_2(y)$, and estimate f_1 and f_2 through the following iterative algorithm:

1. Choose an initial estimator of the component f_1 denoted by $\widehat{f}_1^{(0)}$. For simplicity, the estimator (3.2) derived above is denoted by $\widehat{f}^{(0)}$.

2. Using $\widehat{f}_1^{(0)}$, $f(x,y) \approx \widehat{f}_1^{(0)}(x)f_2(y)$ so that $\int_{\mathscr{S}_y} f(x,y)dx \approx f_2(y) \int_{\mathscr{S}_y} \widehat{f}_1^{(0)}(x)dx$ with $\mathscr{S}_y = \{x|(x,y) \in \mathscr{S}\}$. Then the density f_2 is estimated by

$$\widehat{f}_2^{(1)}(y) = \frac{\int_{\mathscr{S}_y} \widehat{f}^{(0)}(x,y)dx}{\int_{\mathscr{S}_y} \widehat{f}_1^{(0)}(x)dx}.$$

3. Using $\widehat{f}_2^{(1)}$, calculate the updated estimator for f_1 by

$$\widehat{f}_1^{(1)}(x) = \frac{\int_{\mathscr{S}_x} \widehat{f}^{(0)}(x,y)dy}{\int_{\mathscr{S}_x} \widehat{f}_2^{(1)}(y)dy}$$

   with $\mathscr{S}_x = \{y|(x,y) \in \mathscr{S}\}$.

4. Repeat steps 2 – 3 until the desired convergence criterion is achieved.

This iterative method provides estimates for any point in the square $\mathcal{S} = \{(x,y) \mid 0 \leq x,y \leq T\}$. Like mentioned before, the data in an insurance company is usually aggregated, thus we derive the algorithm above also for practical situations. If the data is assumed to be aggregated, the two-dimensional density in the estimation grid Δ_{grid} is estimated by an estimator $\widetilde{f}(d_{kl})$ following (3.4). Assuming the target density to be multiplicative, $f(d_k,d_l) = f_1(d_k)f_2(d_l)$, the components $f_1(d_k)$ and $f_2(d_l)$ are estimated by rewriting the steps 2 and 3 above respectively in the following way:

2. Using $\widehat{f}_1^{(0)}(d_k) = 1 \ \forall\, d_k$, we obtain

$$\widehat{f}_2^{(1)}(d_l) = \frac{\sum_{k=1}^{\frac{1}{\Lambda}(n-\lfloor\Lambda(l-1)\rfloor)} \widetilde{f}(d_{kl})}{\sum_{k=1}^{\frac{1}{\Lambda}(n-\lfloor\Lambda(l-1)\rfloor)} \widehat{f}_1^{(0)}(d_k)}.$$

3. Using $\widehat{f}_2^{(1)}(d_l)$, calculate the updated estimator for $f_1(d_k)$ by

$$\widehat{f}_1^{(1)}(d_k) = \frac{\sum_{l=1}^{\frac{1}{\Lambda}(n-\lfloor\Lambda(k-1)\rfloor)} \widetilde{f}(d_{kl})}{\sum_{l=1}^{\frac{1}{\Lambda}(n-\lfloor\Lambda(k-1)\rfloor)} \widehat{f}_2^{(1)}(d_l)}.$$

The steps 2 – 3 should be repeated until the desired convergence criterion is achieved. This provides estimates for any point in the square in case of aggregated data. In general, a convenient choice for the initial value of $\widehat{f}_1^{(0)}(d_k)$ is a constant function.

3.2 The prediction accuracy and data aggregation

A fundamental concern in predicting is the measure of prediction error for a given data set and a given prediction method. Accuracy can be defined as "goodness of fit" or how well the predictive model is able to reproduce data that is already known. In Sections 3.3 and 3.4 two different simulation approaches for simulating claim count data will be used. The advantage of using simulated data is that we can simulate data "to ultimate", and set aside the (otherwise unknown) losses at ultimate as a standard against which we can compare our model's predictions. The estimation and prediction routines are based on a number of simulated samples each containing the full development processes, drawn from the population distribution, which is explicitly specified with distributional assumptions in Sections 3.3 and 3.4. With respect to the valuation date, the actual IBNR claim counts for a concrete sample can be computed with the future development. After the generation, estimation, and prediction steps, a series of total IBNR claim count estimates, $\widehat{R}^{(1)}, \widehat{R}^{(2)}, \ldots, \widehat{R}^{(p)}$, where p denotes the number of simulated samples, is obtained for each considered reserving method. The last step is to compare the performances of these methods. The performance is evaluated by comparing the prediction errors using the mean absolute percentage error $\mathrm{MAPE}(\widehat{R}) = E\left(\frac{|R-\widehat{R}|}{R}\right)$, where R denotes the total number of IBNR claim

counts estimated by an estimator $\widehat{R}$ at a given valuation date. The final MAPE estimate for each method is expressed as the average estimation error of generated samples,

$$\text{MAPE}(\widehat{R}) = \frac{1}{p}\sum_{i=1}^{p}\left(\frac{|R^{(i)} - \widehat{R}^{(i)}|}{R^{(i)}}\right). \tag{3.6}$$

The MAPE is a relative measure which expresses errors as a percentage of the actual data and we find the MAPE an easy and intuitive way of judging the extent, or importance of errors. We cannot use here, for example, the mean squared error as it is not scale-dependent.

There is one question remaining: how to properly compare the results from different aggregation levels. We suggest aggregating data by calendar periods. This way the whole set of generated data is used on each aggregation level and aggregating by calendar period also avoids over or under estimation (especially in case of the annual data). In Figure 3.4 we have explained the

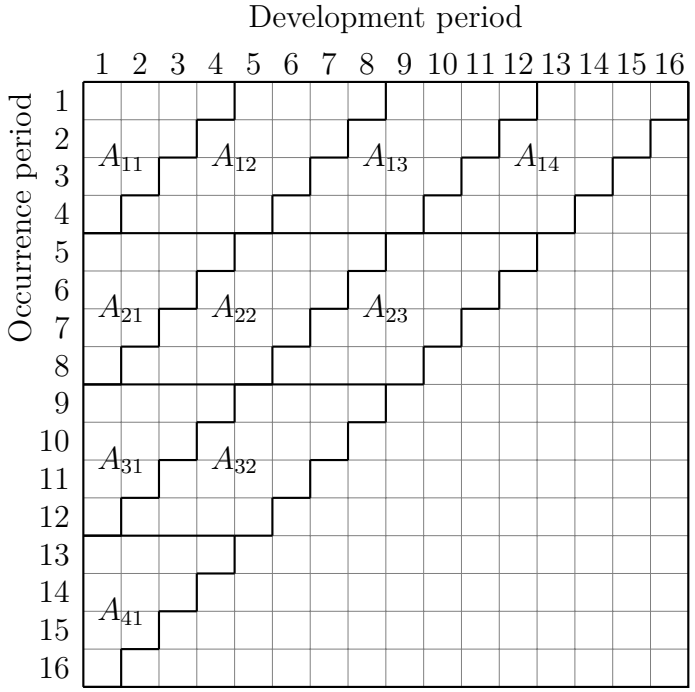

Figure 3.4: Example of data aggregation by calendar period, where the considered time period is $n = 16$ and the initial (granular) data is aggregated to the next aggregation level by 4 periods.

technique for aggregating the claims data by calendar period. Aggregated areas are denoted by $\{A_{ij} : (i, j) \in \Delta\}$ in the shape of parallelograms, and in

the given figure the time period considered $n = 16$ serves as a simple example for giving the idea. The daily claims data are aggregated into monthly data based on its date and making sure that each claim falls into a respective calendar month; monthly claims data are aggregated into quarters by every 3 months and quarterly claims data is aggregated to yearly claims data by every 4 quarters.

3.3 Simulation study I

This chapter presents a simulation study, which is the first approach among the two simulation techniques considered in this research on loss reserving. The purpose of the simulation study is to highlight and discuss the data aggregation levels in which micro-level models outperform traditional macro-level model by evaluating the performance of both the macro and micro-level model with the simulated data. The modeling procedure and the results of the following simulation studies provide guidance for the empirical study in Chapter 3.5.

3.3.1 Methodology

In this simulation study, the IBNR claims are modeled. The first simulation technique used is rather simple by nature. We simulate the claim count data over a period of $n_{years} = 5$ years in the framework of homogeneous Poisson process. The data generation scheme is described in the following:

1. We simulate monthly claim count data for each considered occurrence month, $X_i, i = 1, \ldots, n_{months}$, where $n_{months} = 12 \cdot n_{years}$ and make an ad hoc assumption about the monthly claim frequency, $\lambda_{claims} = 3000$. We generate the number of claims (for each month) from Poisson distribution, $X_i \sim Po(\lambda_{claims})$.

2. We make the simplifying assumption that there are 30 days in a month and use uniform distribution to distribute claims uniformly among each month into daily level, obtaining daily claims data as a result.

3. Given the obtained daily data, we generate the delay (in days) for each claim in the data set. We make the assumption that 95% of the claims are reported in the given time frame (the maximum observed time frame is the number of occurrence days, $n_{days} = 30 \cdot n_{months}$). The delay X_{delay}

is assumed to have an exponential distribution, $X_{delay} \sim Exp(\lambda_D)$, and the delay parameter λ_D is determined from

$$0.95 = \mathbb{P}\{X_{delay} \leq n_{days}\} = 1 - \exp(-\lambda_D \cdot n_{days})$$

as follows

$$\lambda_D = -\frac{\ln(0.05)}{n_{days}}.$$

We use generated daily claims data to aggregate it into monthly, quarterly and annual data to compare the performance on outstanding liabilities of the classical chain-ladder method on yearly, quarterly, monthly and on daily level; the continuous chain-ladder model is considered only on a monthly level due to computational reasons.

3.3.2 Simulation results

In this section, we present and discuss the estimation results of the first simulation study. We simulate the full distribution of IBNR claim counts over 5 calendar years. Then we aggregate the simulated loss counts like specified in Section 3.2 to obtain the total number of IBNR claim counts for the portfolio. To estimate the uncertainty, the simulation routine is repeated 2000 times to generate a predictive distribution of the IBNR claims. The low number of simulated reserves is due to the extensive computational time. The predictive distributions for the IBNR claims are investigated with each model and the prediction accuracy of the forecasting methods are compared.

For the bandwidth selection, the cross-validation (*leave-one-out*) score defined in Section 3.1.1 was used. Since the data is simulated 2000 times, it would be too time consuming to evaluate the optimal bandwidths for each simulated data set. Therefore, to obtain optimal bandwidths for the simulation study, 50 data sets were simulated and optimal bandwidths were obtained with each simulated data set. Among those 50 pairs of bandwidths, the most frequently appeared bandwidth pair were chosen for the simulation study. We obtained $h_1 = 3.5$ (months) and $h_2 = 8.0$ (months) as optimal bandwidths. The Epanechnikov kernel was the chosen kernel function in the study as we applied several kernel functions (see Table 3.2) but did not see a significant difference on results when using another kernel function.

Figure 3.5 shows the distribution of the IBNR claims as obtained with the different methods (from left to right: the actual total IBNR claim counts and predictions by the CCL with the monthly data, the CL with the daily

data, the CL with the monthly data, the CL with the quarterly data and the CL with the annual data). The histograms are based on 2000 simulations of the total IBNR claim counts. The total IBNR claim count predicts the complete lower triangle. The solid red line in each plot indicates what has really been observed, i.e. the average actual total number of IBNR claims. These unrealistically low number of claim counts are a disadvantage of the CL method with the annual data; the actually observed amount is in the right tail of the corresponding histogram. The predictive distribution obtained with the CCL model and the CL model with the monthly data are the most realistic. The CL model with the quarterly data has some jumps in the histogram and the same holds for the CL with the daily data. The more aggregate models (the CL with quarterly and annual data) tend to understate the total number of IBNR claim counts.

In Figure 3.6 the prediction errors of the total IBNR claim counts are given. The performance of the models is evaluated by comparing the prediction error MAPE. The box plots in Figure 3.6 clearly state that the CL method with the annual data has the highest prediction error, on average 7%. The CCL method with the monthly data and the CL method with the monthly tend to estimate the total IBNR claim counts the most accurately, the average MAPE is around 1% for both methods. The CL method with the daily data and the quarterly data are also predicting the the total IBNR claims rather precisely (compared to using the annual data). We conclude from Figures 3.5 and 3.6 that the estimate of the total IBNR claim counts with the CCL method using the monthly data as well as the CL method using the monthly data is close to the true realization. It is also very clear from Figure 3.6 that the CL method with the highly aggregated data (annual data) gives a poor prediction.

3.4 Simulation study II

This section presents the second simulation study to evaluate the performances of the considered models on simulated data. We conduct a similar simulation study as in the previous section, but change the simulation methodology. Using distinct simulation techniques allows to efficiently analyze the output of different methods used in the study. In this simulation study we use a more complex method for generating the claims data in contrast to the previously used simplistic simulation technique.

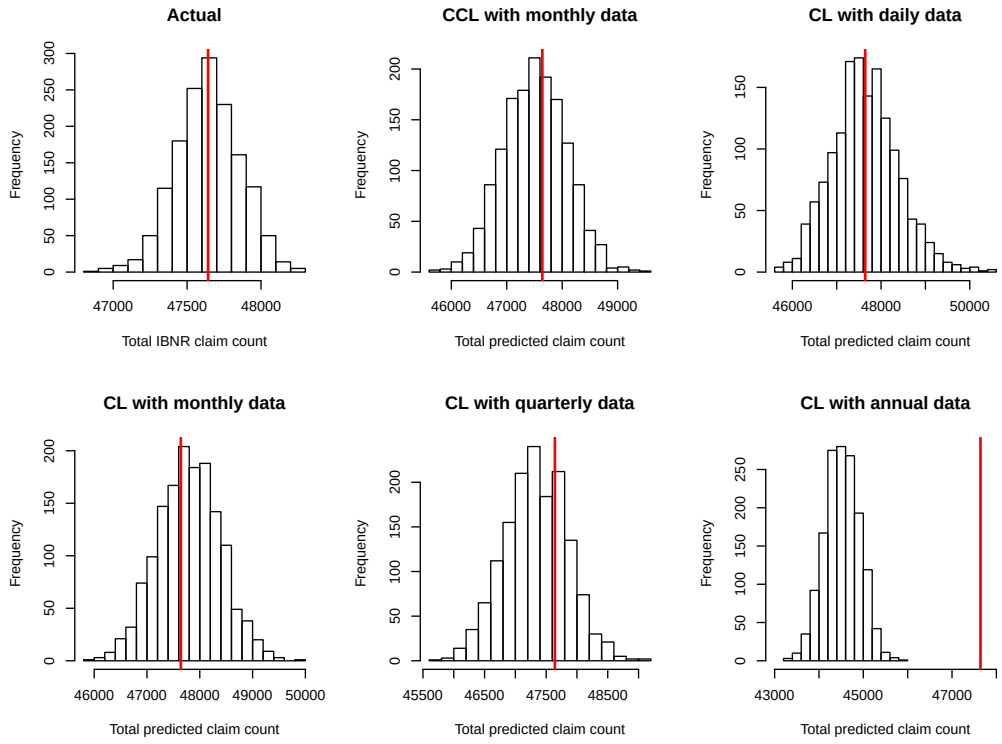

Figure 3.5: Forecasts of the total number of IBNR claims. Actual denotes actual IBNR claim counts based on the data; CL denotes the classical chain-ladder method with daily, monthly, quarterly and yearly data; CCL denotes the continuous chain-ladder method using monthly data and the selected bandwidths of the Epanechnikov kernel are $h_1 = 3.5$ and $h_2 = 8.0$. The red solid line on each histogram indicates the average actual reported claim counts based on the simulated data.

3.4.1 Methodology

It is a common practice to use the Poisson process for modeling a claim number process. For historical reasons, but also since it has very attractive mathematical properties, the Poisson process plays a central role in insurance mathematics. There are many contributions using the marked Poisson process for both simulation and empirical studies. In this simulation study, we use the approach proposed by Badescu et al. (2016b), where the claim arrival process together with its reporting delay is modeled as a marked Cox process (or a doubly stochastic Poisson process) in which the intensity function is stochastic. The stochastic intensity function $\Psi(t)$ is a piecewise stochastic process: $\Psi(t) = \Psi_l$, for $a_{l-1} \leq t < a_l, l = 1, 2, \ldots$ and $a_0 = 0$. Here t is time

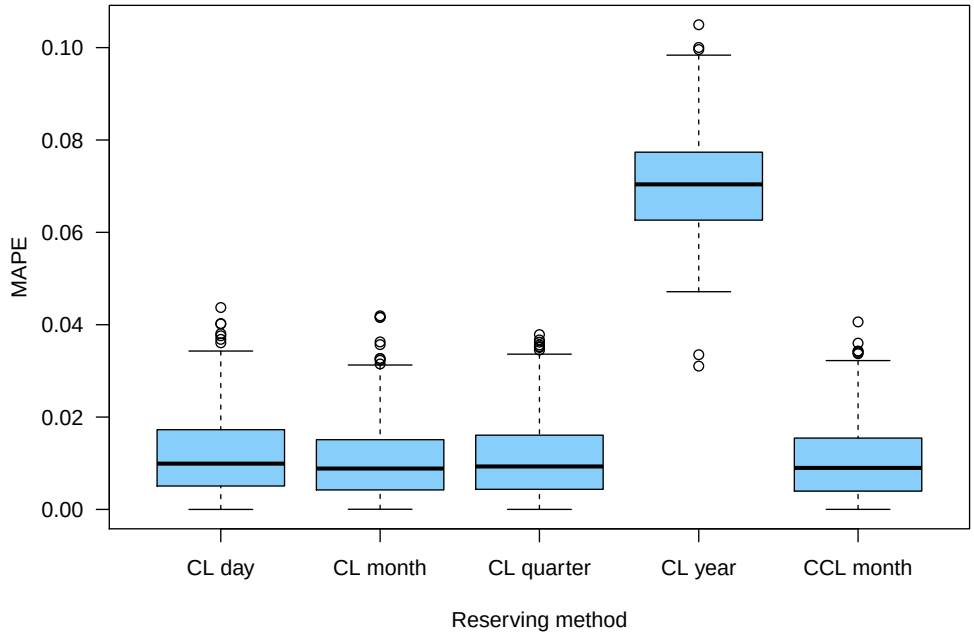

Figure 3.6: Mean absolute percentage error of total number of IBNR claims. CL denotes the classical chain-ladder method using daily, monthly, quarterly and yearly data. CCL denotes the continuous chain-ladder method using monthly data and the selected bandwidths of the Epanechnikov kernel are $h_1 = 3.5$ and $h_2 = 8.0$.

and $a_l, l = 0, 1, \ldots$, are interpreted as data collecting times and the length of each period is taken to be one week. The intensity is generated by a hidden Markov model (HMM) with Erlang state-dependent distributions with shape parameter m_i and scale parameter $\omega_l\theta$, where ω_l represents the risk exposure for the lth period. This implies that conditional on the state of the HMM in the lth period, Ψ_l follows an Erlang distribution of which the parameters depend on the state; unconditional on the state of the HMM in the lth period, Ψ_l follows a mixed Erlang distribution. Badescu et al. (2016b) show that under these assumptions the number of claims arriving during each period follows a Pascal distribution where the parameters depend on the state of the HMM. Conditional on state $i = 1, 2, \ldots, g$, the probability function of the number of claims in period l is given by the following Pascal distribution

$$P(N_l = k | H_l = i) = p(k; m_i, (a_l - a_{l-1})\omega_l\theta),$$

where

$$p(k; m, \theta) = \binom{k+m-1}{m-1} \left(\frac{1}{1+\theta}\right)^m \left(\frac{\theta}{1+\theta}\right)^k.$$

Note that N_l is the number of claims that arrived during $[a_{l-1}, a_l)$, no matter reported or not. The hidden parameter process $\{H_1, H_2, \ldots\}$ is a time-homogeneous Markov chain with a finite state space $\{1, 2, \ldots, g\}$, initial distribution $\boldsymbol{\pi}$ and transition probability matrix $\boldsymbol{\Gamma}$.

Badescu et al. (2016a) calibrate the marked Cox process model on a real insurance data set. Using weekly time intervals, the fitted model contains the following parameters:

$$\begin{aligned} \text{The common scale parameter } \widehat{\theta} &= 3.204 \times 10^{-5}, \\ \text{Shape parameters } (\widehat{m}_1, \widehat{m}_2) &= (49, 64), \\ \text{Transition probabilities } \widehat{\boldsymbol{\Gamma}} &= \begin{pmatrix} 0.986 & 0.014 \\ 0.007 & 0.993 \end{pmatrix}, \\ \text{Initial state probabilities } \widehat{\boldsymbol{\pi}} &= (1, 0). \end{aligned}$$

We refer to Badescu et al. (2016a) for all the details. The out-of-sample test for the total number of IBNR claims supports their claim that the proposed marked Cox model is more realistic compared to an over-dispersed Poisson model. We use this fitted model to generate realistic claims arrival data for our study. Additionally, we set the exposure ω_l in each week equal to 500,000. As said, the considered model allows us to generate the number of weekly claim arrivals, but we are interested in daily data. In order to transform the weekly data to daily data, we use the daily insurance data from the general liability insurance portfolio used in the case study of Section 3.5. Based on the claims with occurrence dates in between 01/01/1998 and 31/12/2002 which are reported before 31/08/2009 we report the empirical probabilities of claims arrival on each weekday in Table 3.3. By sampling from a multinomial distribution with these probabilities, we transform the weekly claim counts to daily claim counts.

	Weekday						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Probability	0.14044	0.13265	0.13995	0.13354	0.14571	0.16969	0.13803

Table 3.3: Probabilities of claims arrival on each weekday.

Badescu et al. (2016a) model the reporting delay distribution using mixture of Erlang distributions. The fitted mixture contains seven Erlang components,

but the corresponding estimates are not disclosed in the given paper. Therefore, for the reporting delay distribution in our simulation setup, we again resort to the daily insurance data from the general liability insurance portfolio used in the case study of Section 3.5. We use the empirical reporting delay distribution based on the claims with occurrence dates in between 01/01/1998 and 31/12/2002 which are reported before 31/08/2009. We generate reporting delays for each claim in our simulation study based on this empirical distribution.

3.4.2 Simulation results

In this section, we discuss and present the estimation results of the second simulation study. We use the calibrated model described in Section 3.4.1 to simulate the full distribution of IBNR claim counts over 5 calendar years. Data aggregation is done as explained in Section 3.2. The simulation routine is repeated 2000 times to generate a predictive distribution of the IBNR claim counts. The predictive distributions for the IBNR claims are investigated with each model and the prediction errors of the forecasting methods are compared.

For the bandwidth selection cross-validation (*leave-one-out*) score defined in Section 3.1.1 was used. We selected two sets of bandwidths: $h_1 = 1.2$ (months), $h_2 = 0.7$ (months) and $h_1 = 0.9$ (months), $h_2 = 0.7$ (months). The Epanechnikov kernel was the chosen kernel function in the study. Figures 3.7 and 3.8 show the distribution of the total number of IBNR claims obtained with the different methods (from left to right: the actual total IBNR claim counts and predictions by the CCL with the monthly data, the CL with the daily, the CL with the monthly data, the CL with the quarterly data and the CL with the annual data). The histograms are based on 2000 simulations of the total number of IBNR claims. The histogram figures clearly show that the data generating methodology used in this simulation study is more complex than the methodology used in the previous section. We observe that the predictive distributions are bimodal, which hints that the corresponding univariate Pascal mixture has at least two components (see Theorem 6.1 in Badescu et al. (2016b)). Since the hidden Markov chain in the fitted model has only two states, each of its simulated path fluctuates between values of 1 and 2, which can be interpreted as the "favorable" state and the "unfavorable" state. If the number of "favorable" states dominates, the number of IBNR claims will be relatively small and its simulated value will be centered around the first mode. In contrary, if there are more "unfavorable"

states than "favorable" states, the number of IBNR claims tends to be large and its simulated value will be closer to the second mode. In this case, it

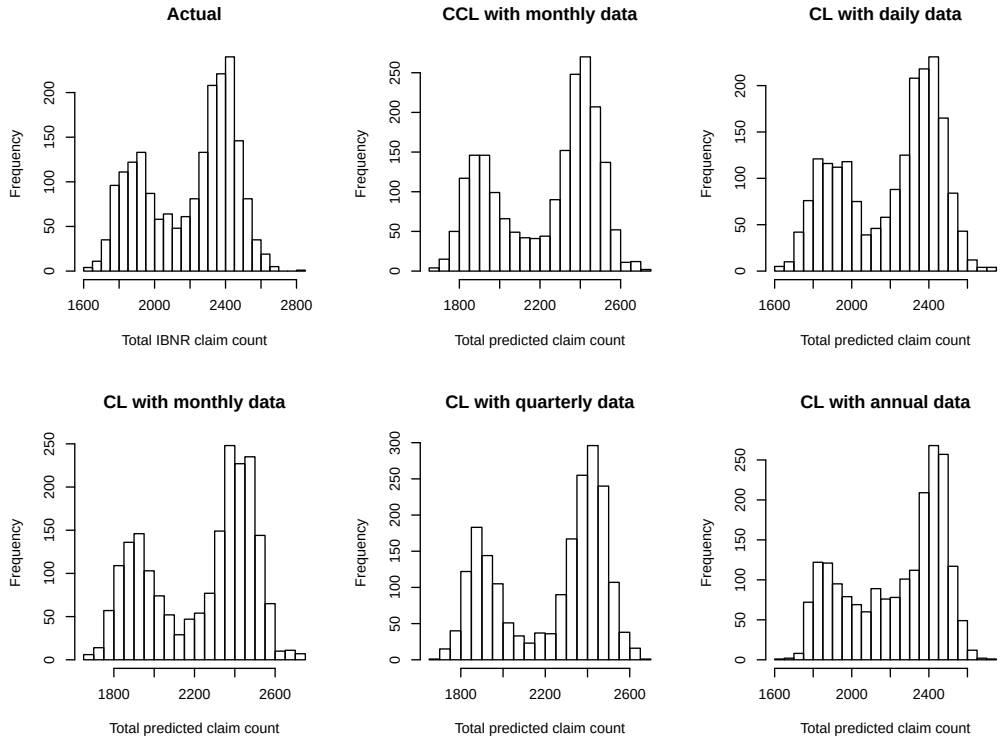

Figure 3.7: Forecasts of the total number of IBNR claims. Actual denotes actual IBNR claim counts based on the data; CL denotes the classical chain-ladder method with daily, monthly, quarterly and yearly data; CCL denotes the continuous chain-ladder method using monthly data and the selected bandwidths of the Epanechnikov kernel are $h_1 = 1.2$ and $h_2 = 0.7$.

is irrelevant to plot the average actual total number of the IBNR claims. We concentrate on comparing the full predictive distributions of the total IBNR claim counts. In contrast to the results obtained in Section 3.3, the CL method with the annual data tends to give more accurate predictions with this simulated data. If we look at the shape of the actual total IBNR claims distribution and compare it to the predictive distributions of different methods, we see that all the methods are performing surprisingly well, but the predictive distribution obtained with the CCL model and the CL model with the daily data as well as with the monthly data look the closest to the actual distribution. The more aggregate models (the CL with the quarterly and the

annual data) tend to slightly underestimate the total number of IBNR claim counts.

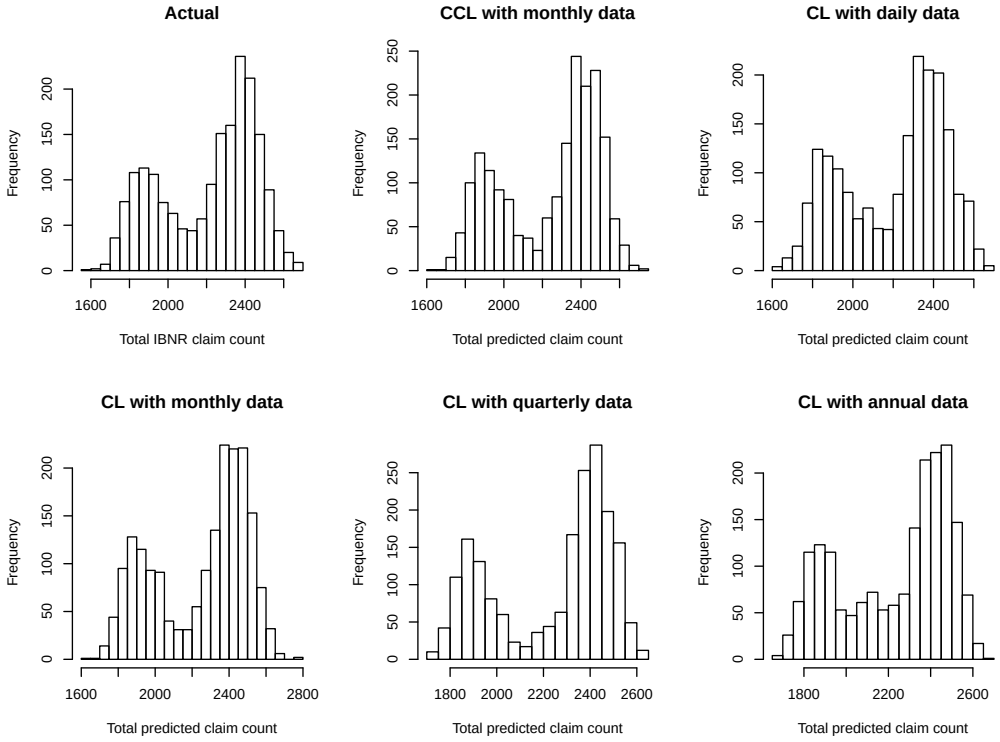

Figure 3.8: Forecasts of the total number of IBNR claims. Actual denotes actual IBNR claim counts based on the data; CL denotes the classical chain-ladder method with daily, monthly, quarterly and yearly data; CCL denotes the continuous chain-ladder method using monthly data and the selected bandwidths of the Epanechnikov kernel are $h_1 = 0.9$ and $h_2 = 0.7$.

Figures 3.9 and 3.10 present the prediction accuracy of methods in 2000 simulations. Based on the computed MAPE as the prediction accuracy, both figures show that the classical CL method predicts the future claim counts more precisely with the granular data, i.e. if using the data on daily or monthly basis, and even with the quarterly data. The given outliers of the graph show that the prediction errors obtained using annual data amounts to even 30% error, whereas the prediction errors of the other methods do not exceed 17%. In case of the CCL method, the prediction errors are as low as with the classical CL method using granular data. The given prediction errors can be large for the annual data due to small number of observations used in run-off triangles. The chosen bandwidths have an effect on the IBNR

claim counts estimates obtained by the CCL method, since on Figure 3.9 we notice less outliers, which gives an insight that the chosen bandwidths $h_1 = 1.2$ and $h_2 = 0.7$ must have been more suitable overall. If only the prediction precision is compared, then the classical CL method with the daily data seems to be the most precise.

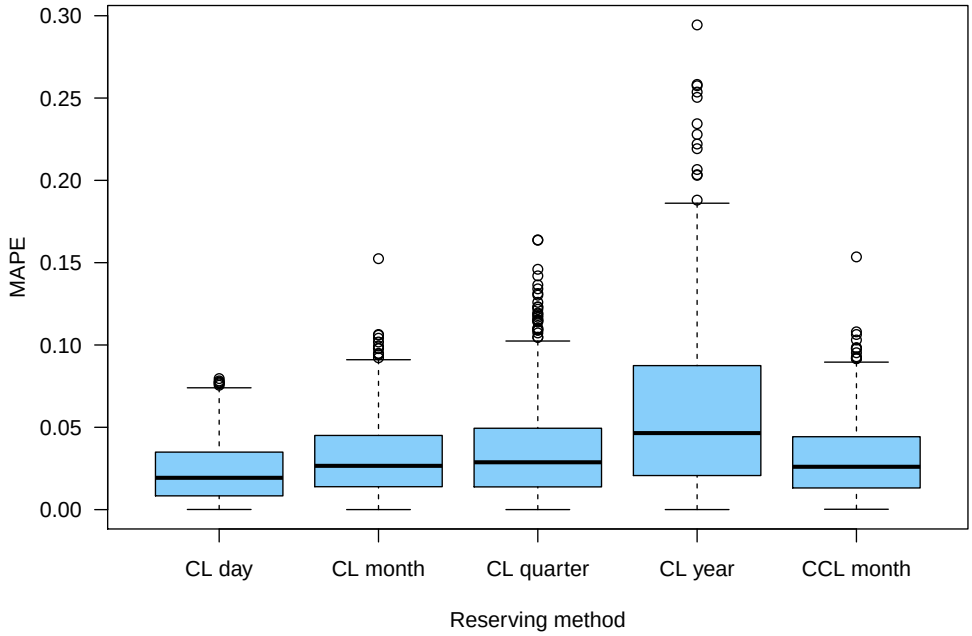

Figure 3.9: Mean absolute percentage error of total number of IBNR claims. CL denotes the classical chain-ladder method using daily, monthly, quarterly and yearly data. CCL denotes the continuous chain-ladder method using monthly data and the selected bandwidths of the Epanechnikov kernel are $h_1 = 1.2$ and $h_2 = 0.7$.

It is remarkable that the classical CL forecasts are comparable with the CCL forecasts in this simulation study. Clearly the CCL method has great advantages over the classical CL method, such as ability to predict cash-flow (in combination with the double chain-ladder method, for instance) and the likely variation in the point estimate among many other advantages. However, if an actuary is interested merely in a point estimate, then using the classical CL method with more granular data than the usual annual data can give reasonable predictions. We note that given the two different simulation studies, the CCL method remained precise in its predictions in both studies.

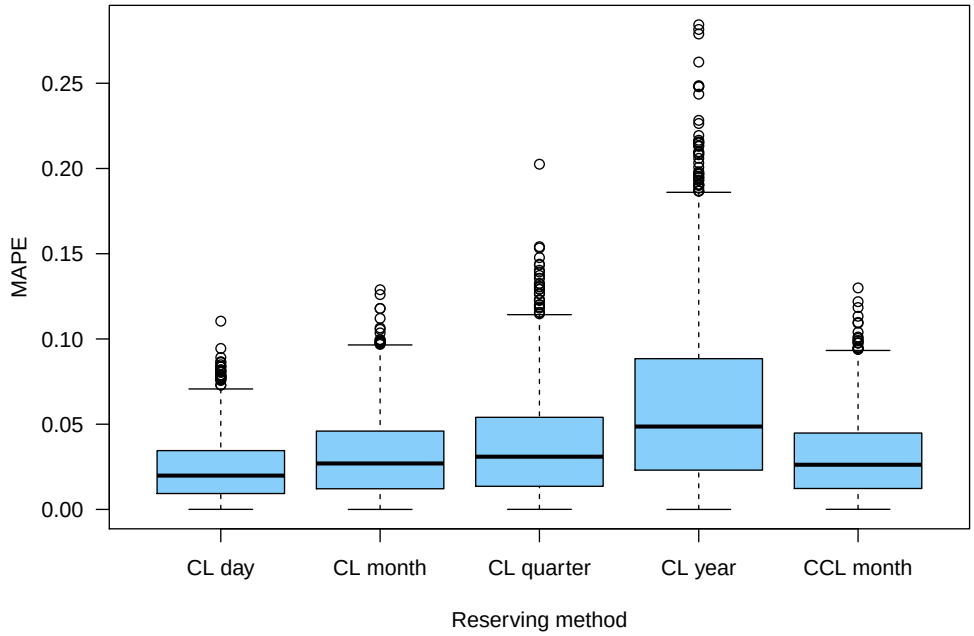

Figure 3.10: Mean absolute prediction errors of total number of IBNR claims. CL denotes the classical chain-ladder method using daily, monthly, quarterly and yearly data. CCL denotes the continuous chain-ladder method using monthly data and the selected bandwidths of the Epanechnikov kernel are $h_1 = 0.9$ and $h_2 = 0.7$.

The CL method predicted best with the monthly data in the first study and with the daily data on the second study. This shows the stability of the CCL method. We conclude that based on these two simulation studies with different simulation methodologies the well-known CL method gives more precise estimates with granular data. Simulations studies have shown that the CL method with the usual annual data understates the total number of IBNR claim counts and is far from the true realization.

3.5 Empirical study

To further investigate the research problem, we demonstrate the comparison of the macro- and micro-level reserving framework using the real-world insurance

data from a general liability insurance portfolio (for private individuals) of a European insurance company. This data set has also been studied in Pigeon et al. (2013) and Antonio and Plat (2014). The models are estimated with historic data and validated with a hold-out sample.

3.5.1 Claims data

The data consist of two types of claims, material damage and bodily injury claims, filed with the insurer between January 1997 and August 2009. Both types of claims are modeled jointly in this study as the empirical distribution of observed reporting delays is similar for material and injury claims. For each claim, a detailed record that tracks the development of the claim up to August 2009, including claim occurrence (accident) date, reporting date, payment amount of each transaction and claim's closure date (settlement date), is provided. The claim file consists of 1,525,374 records corresponding with 491,911 claims. We model the IBNR claims and consider the claim occurrence time and the reporting delay as the key development information for each claim. In the following we give a brief description of the data.

Claim occurrence time. Figure 3.11 shows the number of observed claims that occurred in each month from January 1997 to August 2009. We have not distinguished the claims types as we model all the claims jointly. Similar seasonal fluctuations are observed over each year, i.e., the highest occurrence in summer months and lowest occurrence in winter months. We can further observe an increasing trend in claim occurrences over time, which can be explained by the expanding business volume of the insurer. Note that the downward spikes correspond to the month February.

Reporting delay. The reporting delay is an important driver of the IBNR prediction. The reporting delay is calculated as the number of days from occurrence to reporting. Obviously, the delay is only available for claims that have been reported to the insurance company before the end of the observation period. Figure 3.12 shows the reporting delays in months since occurrence of the claim. For the visualization matters the reporting delay is censored at 6 months. Majority of the claims are reported within a week after the accidents occur. Table 3.4 shows the distribution of claims with a reporting delay of zero to seven days. About twelve percent of the claims have been reported at the occurrence day and the highest percentage of claims reporting takes place a day after the claim occurrence, 19.2% of the claims

have a reporting delay 1 day. Almost 70% of the claims have been reported within a week. It is clear that the IBNR claims have developed rather fast.

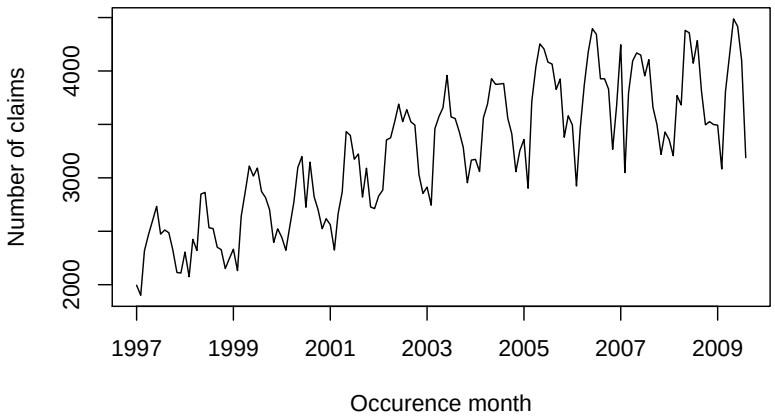

Figure 3.11: Number of claims occurred in each month from January 1997 to August 2009.

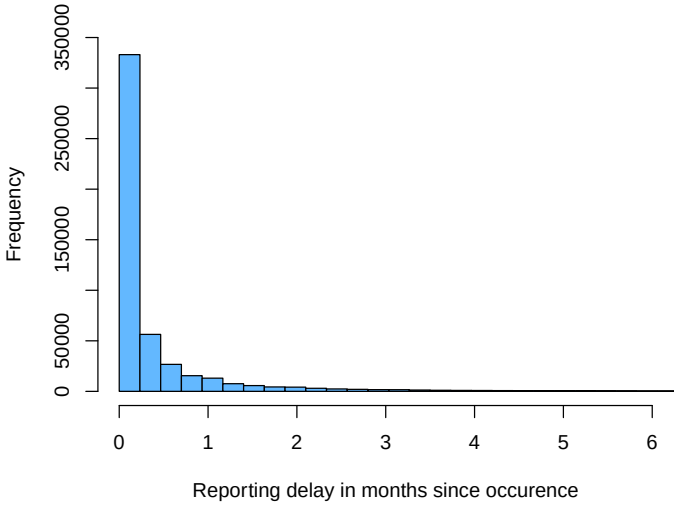

Figure 3.12: The reporting delay censored at 6 months.

Reporting delay in days	# of Claims	% of Claims	Cum % of Claims
0	61,142	12.4	12.4
1	94,599	19.2	31.6
2	57,824	11.8	43.4
3	39,355	8.0	51.4
4	27,585	5.6	57.0
5	20,672	4.2	61.2
6	16,379	3.3	64.5
7	15,504	3.1	67.6

Table 3.4: The distribution of the reporting delay for the first seven days since occurrence of claim.

3.5.2 Estimation results

We use claims from the accident years 1998 – 2002 in this study, assume a valuation date of 31/12/2002 and divide the data into a training set and a validation set according to this valuation date. The training set includes claims development up to the valuation date and the validation set includes claims development after the valuation date and up to 12/31/2006. We are not using the year 1997 for this study as there were not too many claims recorded in that year and few months had a low quality of data. Table 3.5 below shows the training set and validation set (shaded cells) when they are compiled to a run-off triangle. The same models as used in the simulation study Sections 3.3 – 3.4 are fitted to the training set. For the CCL model the Epanechnikov kernel function was chosen again and the bandwidths were obtained with the cross-validation, leading to the optimal bandwidths $h_1 = 1.9$ (months) and $h_2 = 1.5$ (months).

Origin	Development year				
year	1	2	3	4	5
1998	27,805	1,105	25	12	10
1999	31,115	1,315	34	22	3
2000	31,396	1,438	59	4	2
2001	33,122	1,792	53	17	5
2002	37,964	1,672	56	16	9

Table 3.5: Run-off triangle of incremental reported losses for claims from accident years 1998–2002 over development years 1–5.

Table 3.6 shows that the closest estimate to the actual total number of the IBNR claims is obtained with the CCL method. The classical CL method

Future	Reserving method					
	Actual	CL day	CL month	CL quarter	CL year	CCL month
2003	**1,732**	1,451	1,463	1,631	1,810	1,722
2004	**75**	81	81	81	80	72
2005	**21**	32	32	33	35	29
2006	**9**	12	13	14	14	11
Total	**1,837**	1,576	1,590	1,759	1,939	1,834

Table 3.6: Predicted number of reported claims at each future calendar year from 2003 – 2006. Actual denotes actual reported claim counts based on the data; CL denotes IBNR estimate by the classical chain-ladder method with daily, monthly, quarterly and yearly data; CCL denotes IBNR estimate by the continuous chain-ladder method with monthly data. Bandwidths for the Epanechnikov kernel: $h_1 = 1.9$ and $h_2 = 1.5$.

seems to strongly underestimate the total number of the IBNR claims when using daily or monthly data. Since the classical CL method is based only one simple assumption – the proportionate relationships between values in consecutive development years will repeat in the future – then this assumption might be working against the method when too granular data is used. Namely, if the daily (or weekly and monthly) data are used, then the run-off cells have small values and if the model is using the proportions (or so called development factors) for the predictions, then it clearly causes a lot of instability in the predictions. The CL method assumes a lot of information (lots of claims in the run-off triangle) to give reasonable predictions. For the future years 2004 – 2006, the predictions obtained by different methods are nearly equal. The biggest difference in predictions is in the first year, 2003. The estimate of the IBNR claims for the first year makes the biggest difference for the total estimate. We see that the CL method with the annual data is slightly overestimating the total number of IBNR claim counts, but in absolute values the prediction is more precise than using daily or monthly data. However, the CL method with quarterly data is predicting the best. This gives an insight to using the classical CL method with more refined data than the usual annual data.

Figure 3.13 shows the comparison of the IBNR claims on quarterly basis from January 2003 till December 2006 by the CCL method with the monthly data (the monthly estimates are aggregated to quarters) and the CL method with the quarterly data. The black dot indicates the actual IBNR claim counts on

quarterly basis, the red line with the triangular dots indicates the estimates by the CL method and the blue line with square dots indicates the estimate obtained with the CCL method. Since the biggest difference in estimates is in the first future year, we plot first the 4 first quarters and then the remaining 12 quarters. From quarter 5 till the quarter 16, the CL method and the CCL method are predicting rather similarly, the CL method showing slightly higher estimates. Also the quarters 2 – 4 are predicted in the same manner and really close to the actual number of the IBNR claims, but the biggest difference is seen on the first quarter. The CCL method predicts the first quarter very precisely, giving the estimate close to the actual number of IBNR claims.

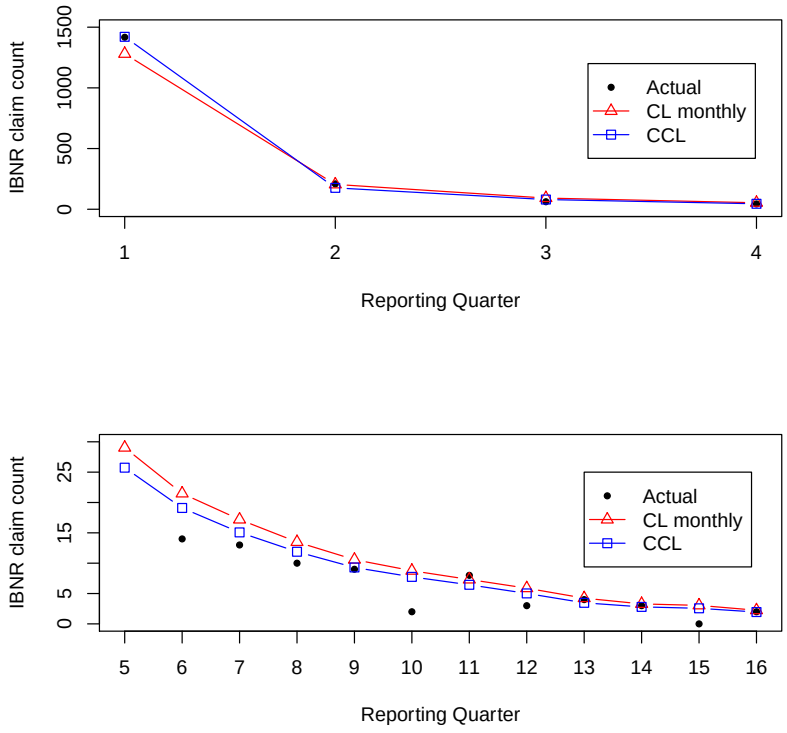

Figure 3.13: Predicted number of reported claims at each future calendar quarter from January 2003 till December 2006. The black squares indicate the actual quarterly reported claim counts, the red triangles indicate the quarterly reported claim counts estimated by the CL method using the quarterly data and the blue rectangles indicate the quarterly reported claim counts estimated by the CCL method using the monthly data.

Figure 3.14 describes the IBNR claims on monthly basis from January 2003 till December 2006 estimated by the CCL method with the monthly data and the CL method with the monthly data. The black dot indicates the actual monthly reported claim counts, the red line with the triangular dots indicates the estimates by the CL method and the blue line with round dots indicates the estimate obtained with the CCL method. The first plot in the Figure 3.14 shows the comparison on the first 12 months and the second plot depicts the estimates for the last 36 months. The biggest difference in estimates comes in only in the very first month. Clearly, the CL method is less accurate in predicting the first calendar periods.

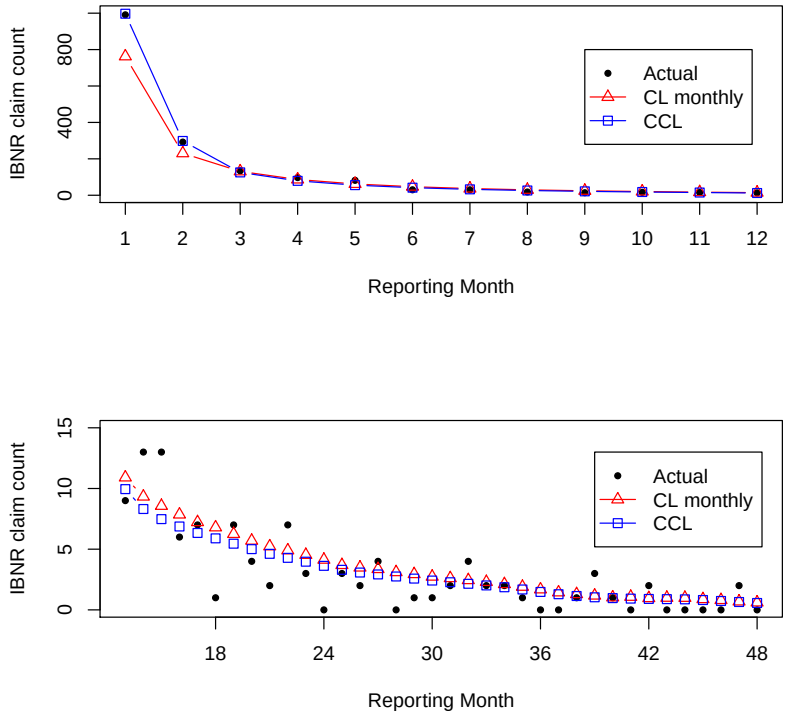

Figure 3.14: Predicted number of reported claims at each future calendar month January 2003 till December 2006. The black squares indicate the actual monthly reported claim counts, the red triangles indicate the monthly reported claim counts estimated by the CL method using the monthly data and the blue rectangles indicate the monthly reported claim counts estimated by the CCL method using the monthly data.

3.6 Conclusions

The simulation results suggest that a macro-level model with more granular data than the annual data is able to estimate the total number of IBNR claim counts with better quality (or more precisely). It is not surprising that the use of micro-level model reduces the prediction uncertainty, leading to estimates with higher precision, this was seen in simulation studies as well as in the empirical study with the insurance data. The CCL method looks like an interplay between the good old chain-ladder method and micro-level framework with the use of sophisticated yet powerful statistical estimation tools. There is no doubt that the CCL method is able to provide us a more precise cash flow with all the desired statistical indicators, but it remains a question whether the advantage is great enough to cope with the increasing computational time and difficulty level of the used statistical methods. Like mentioned beforehand, the critical point of the CCL method is the choice of the bandwidth. We saw that using the simple (yet time consuming) cross-validation score can lead us to more precise estimation but the difference in point estimates with the simple chain-ladder is not that drastic. To increase the predictive power of the CCL method, the investigation of the bandwidth selection is required. Given the intensive computational time, some rule-of-thumb bandwidth selection ideas could become useful.

There was only small steam of contributions about micro-level models not too long ago. Today, developing micro-level models is one of the key topics in reserving literature. Many advanced reserving methods have been developed, but nothing compares to having a good knowledge and understanding for the underlying data set. The insurance companies have the access to monthly or daily data, if not to development of each individual policy (or claim). The recent contributions encourage the use of individual data model in loss reserving over the classical ones based on aggregate data models, provided that individual data are available. One should not forget that at the end of the day, an actuary is interested in the best reserve estimate. Micro-level models can be more precise, but they should be easily applicable in industry as well. We have seen in this study, that using (more) granular data with simple macro-level models can also lead to accurate estimates. The problem is the instability of the CL method, thus using the daily data do not give more precise predictions with every data set. However, a fair compromise between the use of annual data and the use of daily data could be, for instance, the quarterly data.

Outlook

This chapter summarizes the main findings of this dissertation and concludes our work by presenting several suggestions for future research related to this topic. The main purpose of this dissertation is to introduce broadly the stochastic chain-ladder reserving models and its extension that can work on the individual claim level. Throughout the whole dissertation we work with real-life actuarial data. While each data set poses specific challenges to modeling and model selection, this thesis aims at providing practical solutions to actuarial problems that appear in practice.

In the first part of this thesis, we concentrate on modeling the claim payments and the focus is on testing the effectiveness of the existing methodologies. We have extended the previous reserving related contributions in bootstrapping context by broadening the choice of the residuals. A bundle of empirical studies with real actuarial data are presented to illustrate the impact of the considered predictive models. It is not possible to make a comparison with all the existing macro-level models in a single study but we conduct an extensive case study involving insurance data from three different lines of businesses along with introducing model validation in the reserving literature. Our findings suggest the use of corrected residuals and we have shown that the Anscombe residuals could be considered as an alternative to the well-known Pearson residuals. The proposed model validation and assessment ideas are generic and do not depend on a particular data set, thus constitute a useful tool in reserve estimation.

In the second part of this thesis, we shift to modeling the IBNR claim counts and concentrate on the continuous approach to the classical chain-ladder method. We consider the continuous chain-ladder method as the first chain-ladder extension in the micro-level framework and provide a thorough and in-depth demonstration of this micro-level loss reserving method. We derive the corresponding algorithms of this continuous model in order to apply it in the discrete framework and make it feasible for practitioners. We conduct two different simulation studies and an empirical study to describe

the effectiveness of the continuous approach as well as to discover and compare the performance of the chain-ladder method applied to different levels of data granularity. Our findings in this part of the thesis suggest that the use of more granular data than regularly used annual data can increase the estimation precision of the classical chain-ladder method and both the simulation and empirical study show the estimation stability as a strength of the continuous chain-ladder method. The quality of the estimation results obtained with the CL method using daily or monthly data depends on the underlying data. The CL method with the daily data outperformed all the other considered models in the simulation study, but strongly underestimated the reserve in the empirical study. Using quarterly data instead results in more similar prediction accuracy across different data sets. Our study results show that the CCL method performs well and leads to a good prediction accuracy in all the considered data setups.

Compared to the traditional models, the continuous approach of the chain-ladder method requires by nature more refined data and more groundwork concerning the used statistical methodology and the CCL also requires highly intensive computation. Fitting the model with the chain-ladder method is instant, but it can take days to estimate the continuous model and generate the reserve estimate. As one may have noticed, I have to make some compromise in both the simulation and the empirical study due to issues of computational power.

There are still many questions about both macro and micro-level reserving that are not addressed by this dissertation. Below we discuss some limitations of the conducted studies and point out several directions for future research. The CCL method was first implemented for count data. Recently there have been some contributions regarding the payments as well, but they lack a strong link with practical implementation. It would be interesting to implement the CCL method also for the claim payments. As stressed already beforehand, the key piece of the model structure is the kernel function, which implies the bandwidth selection. The choice of bandwidths have the most impact on the obtained estimates and thus also requires investigation. Right now we have used cross-validation but the heavily increased computational time is a great downside of the method. It would be useful to derive a rule-of-thumb approach for the bandwidth selection. In addition, it is not straightforward how to quantify the variability of the point estimate with the CCL method. Describing any other statistical characteristics besides the point estimates requires additional work. Future research concerning these questions could support considering the CCL method as a valid alternative micro-level method in the CL framework. We conclude that it is also important to keep track

of the development of macro-level models and to understand under what circumstances these models may fail.

Appendix

Formulation of the different steps and stages of both the regular and alternative bootstrap procedure corresponding to Section 2.2 following Pinheiro et al. (2003):

Step 1. Preliminary steps before the bootstrap cycle:

1) Choose a reserving method for estimating the outstanding liabilities.

2) Choose a type of residuals and decide on the possible residual adjustments. The overview of different residuals, the zero correction and the standardization is given in Section 2.2.1. Denote the residual function by h.

3) Estimate the model parameters c, α_i, β_j for $i, j = 1, 2, \ldots, n$, and ϕ on the initial data.

4) Calculate the fitted values $\widehat{\mu}_{ij}$ in the upper triangle for $(i, j) \in \Delta$, and the predicted values $\widehat{\mu}_{ij}$, $\widehat{C}_{ij}$, $\widehat{R}_i$ and $\widehat{R}$ in the lower triangle for $(i, j) \in \nabla$. Note that $\widehat{\mu}_{ij}$ and $\widehat{C}_{ij}$ only differ in the lognormal case.

5) Calculate the residuals $r_{ij} = h(C_{ij}, \widehat{\mu}_{ij})$ for $(i, j) \in \Delta$.

6) Fix the number of bootstrap repetitions B.

Step 3. The following steps are made in the bootstrap cycle, where $b = 1, \ldots, B$.

Step 3.1 Bootstrap iteration for the regular bootstrap procedure:

7) Resample $\frac{n(n+1)}{2}$ pseudo-residuals $r_{ij}^{(b)}$ for $(i, j) \in \Delta$ from the residuals obtained in stage 5) using replacement.

8) Apply the inverse function h^{-1} to the obtained pseudo-residuals $r_{ij}^{(b)}$ to create the pseudo-data in the upper triangle $C_{ij}^{(b)} = h^{-1}(r_{ij}^{(b)}, \widehat{\mu}_{ij})$ for $(i, j) \in \Delta$.

9) Use the obtained pseudo-data $C_{ij}^{(b)}$ to re-estimate the model and obtain the predicted values $\widehat{\mu}_{ij}^{(b)}$, $\widehat{C}_{ij}^{(b)}$, $\widehat{R}_i^{(b)}$ and $\widehat{R}^{(b)}$ in the lower triangle for $(i,j) \in \nabla$.

Step 3.2 Additional stages 10) - 12) for the alternative bootstrap method (*pseudo-reality*):

10) Resample $\frac{n(n-1)}{2}$ pseudo-residuals $r_{ij}^{*(b)}$ for $(i,j) \in \nabla$ from the residuals obtained in stage 5) using replacement.

11) Apply the inverse function h^{-1} to the obtained pseudo-residuals $r_{ij}^{*(b)}$ to create the pseudo-reality in the lower triangle $C_{ij}^{*(b)} = h^{-1}(r_{ij}^{*(b)}, \widehat{\mu}_{ij})$ for $(i,j) \in \nabla$. Note that $\widehat{\mu}_{ij}$ are the predictions obtained in stage 4) and not those obtained in the bootstrap iteration.

12) Compute the prediction errors $r^{*(b)} = R^{*(b)} - \widehat{R}^{(b)}$, where $R^{*(b)}$ is the sum of the incremental pseudo-reality $C_{ij}^{*(b)}$ in the lower triangle and $\widehat{R}^{(b)}$ is the estimate of the total reserve in the bth iteration of the bootstrap cycle.

13) Return to the beginning of Step 3 until the B repetitions are completed.

Step 4. Bootstrap data analysis.

Step 4.1 For the regular bootstrap method:

14) Obtain the bootstrap standard deviations for the incremental amounts, the yearly reserves and the total reserve as the empirical standard deviations of $\widehat{C}_{ij}^{(b)}$, $\widehat{R}_i^{(b)}$ and $\widehat{R}^{(b)}$, respectively. Use these to compute the bootstrap prediction errors (2.20).

Step 4.2 For the alternative bootstrap method (PPE):

15) Compute the $(1-\alpha)$-percentile of the empirical distribution of the prediction errors obtained in stage 12). Calculate the $(1-\alpha)$ upper limit of the provisions by adding this percentile to the prediction of the total reserve $\widehat{R}$ obtained in stage 4).

List of Figures

List of Tables

Bibliography

Antonio, K. and R. Plat (2014). Micro–level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal 7*, 649–669.

Arjas, E. (1989). The claims reserving problem in non-life insurance: Some structural ideas. *ASTIN Bulletin 19*, 139–152.

Ashe, F. (1986). An essay at measuring the variance of estimates of outstanding claim payments. *ASTIN Bulletin 16*(S1), S99–S113.

Avanzi, B., B. Wong, and X. Yang (2016). A micro-level claim count model with overdispersion and reporting delays. *Insurance: Mathematics and Economics 71*, 1–14.

Badescu, A. L., X. S. Lin, and D. Tang (2016a). A marked Cox model for the number of IBNR claims: Estimation and application. *Available at SSRN 2747223*.

Badescu, A. L., X. S. Lin, and D. Tang (2016b). A marked Cox model for the number of IBNR claims: Theory. *Insurance: Mathematics and Economics 69*, 29–37.

Björkwall, S., O. Hössjer, and E. Ohlsson (2009). Non-parametric and parametric bootstrap techniques for age-to-age development factor methods in stochastic claims reserving. *Scandinavian Actuarial Journal 2009*(4), 306–331.

Björkwall, S., O. Hössjer, and E. Ohlsson (2010). Bootstrapping the separation method in claims reserving. *ASTIN Bulletin 40*(02), 845–869.

Bornhuetter, R. L. and R. E. Ferguson (1972). The actuary and IBNR. In *Proceedings of the casualty actuarial society*, Volume 59, pp. 181–195.

Bouezmarni, T. and J. V. Rombouts (2010). Nonparametric density estimation for multivariate bounded data. *Journal of Statistical Planning and Inference 140*(1), 139–152.

Charpentier, A. and M. Pigeon (2016). Macro vs. micro methods in non-life claims reserving (an econometric perspective). *Risks 4*(2), 12.

Czado, C., T. Gneiting, and L. Held (2009). Predictive model assessment for count data. *Biometrics 65*(4), 1254–1261.

Davison, A. C. and D. V. Hinkley (1997). *Bootstrap methods and their application*, Volume 1. Cambridge University Press: Cambridge, UK.

Dawid, A. P. (1984). Present position and potential developments: Some personal views: Statistical theory: The prequential approach. *Journal of the Royal Statistical Society. Series A (General)*, 278–292.

Dawid, A. P. and P. Sebastiani (1999). Coherent dispersion criteria for optimal experimental design. *Annals of Statistics*, 65–81.

De Jong, P. and G. Z. Heller (2008). *Generalized Linear Models for insurance data*. Cambridge University Press.

England, P. (2002). Addendum to "analytic and bootstrap estimates of prediction errors in claims reserving". *Insurance: Mathematics and Economics 31*(3), 461–466.

England, P. and R. Verrall (1999). Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics 25*(3), 281–293.

England, P. D. and R. J. Verrall (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal 8*(3), 443–518.

Fan, J. and I. Gijbels (1996). *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*, Volume 66. CRC Press.

Friedland, J. (2010). Estimating unpaid claims using basic techniques. In *Casualty Actuarial Society: Arlington, VA, USA*.

Gneiting, T., F. Balabdaoui, and A. E. Raftery (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69*(2), 243–268.

Gneiting, T. and A. Raftery (2007). Strictly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association 102*(477), 359–378.

Haastrup, S. and E. Arjas (1996). Claims reserving in continuous time: A nonparametric Bayesian approach. *ASTIN Bulletin 26*, 139–164.

Hess, K. T. and K. D. Schmidt (2002). A comparison of models for the chain–ladder method. *Insurance: Mathematics and Economics 31*(3), 351–364.

Jewell, W. S. (1989). Predicting IBNyr events and delays: I. continuous time. *ASTIN Bulletin 19*, 25–55.

Jin, X. and E. W. J. Frees (2013). Comparing micro-and macro-level loss reserving models. *Presentation at ARIA*.

Johansson, A. (2015). Claims reserving on macro-and micro-level. Master's Thesis.

Kaas, R., M. Goovaerts, J. Dhaene, and M. Denuit (2008). *Modern Actuarial Risk Theory: Using R*. Springer: Berlin/Heidelberg, Germany.

Kremer, E. (1982). IBNR-claims and the two-way model of ANOVA. *Scandinavian Actuarial Journal 1982(1)*(1), 47–55.

Kuang, D., B. Nielsen, and J. P. Nielsen (2008a). Forecasting with the age-period-cohort model and the extended chain-ladder model. *Biometrika 95*(4), 987–991.

Kuang, D., B. Nielsen, and J. P. Nielsen (2008b). Identification of the age-period-cohort model and the extended chain-ladder model. *Biometrika 95*(4), 979–986.

Kuang, D., B. Nielsen, and J. P. Nielsen (2009). Chain-ladder as maximum likelihood revisited. *Annals of Actuarial Science 4*(01), 105–121.

Kuang, D., B. Nielsen, and J. P. Nielsen (2015). The geometric chain-ladder. *Scandinavian Actuarial Journal 2015(3)*(3), 278–300.

Kunkler, M. (2004). Modelling zeros in stochastic reserving models. *Insurance: Mathematics and Economics 34*(1), 23–35.

Linton, O. and J. P. Nielsen (1995). A kernel method of estimating structured nonparametric regression based on marginal integration. *Biometrika 82*(1), 93–100.

Liu, H. and R. Verrall (2009). Predictive distributions for reserves which separate true ibnr and ibner claims. *ASTIN Bulletin 39*(01), 35–60.

Mack, T. (1991). A simple parametric model for rating automobile insurance or estimating IBNR claims reserves. *ASTIN Bulletin 21*(01), 93–109.

Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin 23*, 213–225.

Mammen, E., M. D. M. Miranda, and J. P. Nielsen (2015). In-sample forecasting applied to reserving and mesothelioma mortality. *Insurance: Mathematics and Economics 61*, 76–86.

Martínez Miranda, M. D., J. P. Nielsen, S. Sperlich, and R. Verrall (2013). Continuous chain ladder: Reformulating and generalizing a classical insurance problem. *Expert Systems with Applications 40*(14), 5588 – 5603.

Martínez Miranda, M. D., J. P. Nielsen, and R. J. Verrall (2012). Double Chain Ladder. *ASTIN Bulletin 42*, 59–76.

McCullagh, P. and J. Nelder (1989). *Generalized linear models* (2nd ed.). Chapman and Hall: New York, NY, USA.

Müller, H. and U. Stadtmüller (1999). Multivariate boundary kernels and a continuous least squares principle. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 61*(2), 439–458.

Nielsen, J. P. (1999). Multivariate boundary kernels from local linear estimation. *Scandinavian actuarial journal 1999*(1), 93–95.

Nielsen, J. P. and O. Linton (1998). An optimization interpretation of integration and back-fitting estimators for separable nonparametric models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 60*(1), 217–222.

Norberg, R. (1993). Prediction of outstanding liabilities in non-life insurance. *ASTIN Bulletin 23*(1), 95–115.

Norberg, R. (1999). Prediction of outstanding liabilities II. Model variations and extensions. *ASTIN Bulletin 29*(1), 5–27.

Peremans, K., P. Segaert, S. Van Aelst, and T. Verdonck (2016). Robust bootstrap procedures for the chain-ladder method. *Scandinavian Actuarial Journal*, 1–28.

Pigeon, M., K. Antonio, and M. Denuit (2013). Individual loss reserving with the multivariate skew normal framework. *ASTIN Bulletin 43*(03), 399–428.

Pinheiro, P. J., J. M. Andrade e Silva, and M. de Lourdes Centeno (2003). Bootstrap methodology in claim reserving. *Journal of Risk and Insurance 70*(4), 701–714.

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

Quarg, G. and T. Mack (2008). Munich chain ladder: A reserving method that reduces the gap between IBNR projections based on paid losses and IBNR projections based on incurred losses. *Variance 2*(2), 266–299.

Renshaw, A. E. (1994). *On the second moment properties and the implementation of certain GLIM based stochastic claims reserving models*. Department of Actuarial Science & Statistics, City University.

Renshaw, A. E. and R. J. Verrall (1998). A stochastic model underlying the chain-ladder technique. *British Actuarial Journal 4*(04), 903–923.

Riebler, A., L. Held, and H. Rue (2012). Estimation and extrapolation of time trends in registry data - borrowing strength from related populations. *The Annals of Applied Statistics*, 304–333.

Schmidt, K. D. (2002). A note on the overdispersed poisson family. *Insurance: Mathematics and Economics 30*(1), 21–25.

Schnieper, R. (1991). Separating true IBNR and IBNER claims. *ASTIN Bulletin 21*(01), 111–127.

Scott, D. W. (2015). *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons.

Staniswalis, J. G. and K. Messer (1996). Addendum to Şkernel estimators for multivariate regressionŤ. *Journal of Nonparametric Statistics 7*(1), 67–68.

Taylor, G. (2011). Maximum likelihood and estimation efficiency of the chain ladder. *ASTIN Bulletin 41*(01), 131–155.

Taylor, G. (2014). Claims triangles/loss reserves. *Predictive Modeling Applications in Actuarial Science 1*, 449.

Tee, L., M. Käärik, and R. Viin (2017). On comparison of stochastic reserving methods with bootstrapping. *Risks 5*(1), 2.

Verrall, R. J. (2000). An investigation into stochastic claims reserving models and the chain-ladder technique. *Insurance: Mathematics and Economics 26*(1), 91–99.

Verrall, R. J. and Z. Li (1993). Negative incremental claims: chain ladder and linear models. *Journal of the Institute of Actuaries 120*(01), 171–183.

Verrall, R. J. and M. V. Wüthrich (2016). Understanding reporting delay in general insurance. *Risks 4*(3), 25.

Wand, M. P. and M. C. Jones (1994). *Kernel smoothing*. Crc Press.

Wright, T. S. (1990). A stochastic method for claims reserving in general insurance. *Journal of the Institute of Actuaries 117*(03), 677–731.

Wüthrich, M. V. (2016). Machine learning in individual claims reserving (November 11, 2016). Swiss Finance Institute Research Paper No. 16-67. Available at SSRN: https://ssrn.com/abstract=2867897.

Wüthrich, M. V. and M. Merz (2008). *Stochastic claims reserving methods in insurance*, Volume 435 of *Wiley Finance*. John Wiley & Sons.

Zehnwirth, B. (1994). Probabilistic development factor models with applications to loss reserve variability, prediction intervals and risk based capital. In *Casualty Actuarial Society Forum*, Volume 2, pp. 447–606.

Stohhastilised ahel-redel meetodid kahjukindlustuses

Kokkuvõte

Reservide hindamine on hetkel väga aktuaalne teema seoses 2016. aasta algusest jõustunud Solventsus II direktiiviga, kuna uues riskipõhises Solventsus II mudelis on reserviriski hindamine üks olulisemaid ülesandeid.

Kindlustusreservide hindamisel kasutatakse kõige sagedamini mudeleid, kus agregeeritud andmed esitatakse kahjukolmnurga kujul ja rakendatakse reservide leidmiseks teatud determineeritud meetodeid (ahel-redel) süvenemata stohhastilisse struktuuri. Samas on üha enam hakatud uurima võimalusi mudelite täpsustamiseks ja üldistamiseks, kusjuures arengud on olnud nii erinevate stohhastiliste struktuuride arvestamise ja sobitamise osas kui ka andmete agregeerimise osas. Viimaste aastate suurim trend on agregeerimata "mikro-tasandi" andmetega mudelid. Põhjalik ülevaade antud valdkonna arengutest ja erinevatest mudelitest on toodud doktoritöö esimeses peatükis.

Käesoleva väitekirja üheks eesmärgiks on anda ülevaade erinevatest (stohhastilistest) reservihindamise meetoditest, võrrelda meetodite omadusi ja käitumist ning leida kriteeriume, mis hõlbustaksid otsuste langetamist praktilises reservide hindamise ülesandes. Töö käsitleb reservide stohhastilist hindamist üldistatud lineaarsete mudelite ja bootstrap-meetodi kombineerimisel. Vaatluse all on ülehajuvusega Poissoni mudel, gammajaotuse mudel ja log-normaalse jaotuse mudel. Võrreldes deterministlike meetoditega, on stohhastilisl lähenemisel võimalik lisaks punkthinnangutele hinnata ka hinnangute varieeruvust. Hinnanguvea saab jagada kaheks komponendiks: protsessiviga, mis tuleb mudelist, ja hinnanguviga, mis tuleb parameetrite hindamisest. Hinnanguvea leidmisel kasutatakse bootstrap-meetodit hinnangujääkidel, ning tuuakse välja ja selgitatakse erinevaid võimalusi, millele Bootstrap-meetodi kasutamisel võiks tähelepanu pöörata. Seejuures vaadeldakse nii Pearsoni kui Anscombe'i jääke ja nende erinevaid täpsustusi. Lisaks mudelite teoreetilisele püstitusele rakendatakse mudeleid ühe Eesti kindlustusfirma andmetele ja erinevaid meetodeid valideeritakse Schedule P andmebaasi erinevate kahjuliikide reservikolmnurkade peal. Empiirilise analüüsi põhjal on sobivaim mudel ülehajuvusega Poissoni mudel, sedasama kinnitavad ka valideerimise tulemused, mis lisaks soovitavad kasutada "null-korrigeeritud" Pearsoni jääke.

Väitekirja teine eesmärk on uurida ahel-redel meetodi rakendamist erinevatel andmete agregeerimise tasemetel, seejuures käsitleda ka ahel-redel meetodi üldistust pideva ajaga juhule. Kuna kindlustusfirmalt ei saa eeldada andmete kogumist pidevas ajas, siis käesolevas töös on pidev ahel-redel meetod interpreeritud ka diskreetse ajaga juhule. Eesmärk on uurida erinevate agregeerimistasemete mõju kogureservi hinnangule ning võrrelda klassikalise ahel-redel meetodi ja pideva ahel-redel meetodi hinnanguid. Analüüs näitab, et nii kvartaalsete, kuiste kui päevaste andmetega on tulemused täpsemad kui aastase agregeerimise korral, aga kvartaalsete andmete korral on tulemused oluliselt stabiilsemad kui suurema detailsusega lähenemiste korral. Ka pideva ahel-redel meetodi korral on tulemused stabiilselt head erinevate simulatsioonide korral. Kas pideva ahel-redel meetodi mõnevõrra suurem stabiilsus ja täpsus kaalub üles teoreetilise keerukuse ja arvutuslikud probleemid, jääb siiski lahtiseks küsimuseks ja sõltub täpsemast ülesande püstitusest.

Curriculum vitae

Name: Liivika Tee
Date and place of birth: 29.04.1989, Valga, Estonia
Citizenship: Estonian
Contact Address: Institute of Mathematics and Statistics, J. Liivi 2, 50409 Tartu, Estonia
Phone: +372 581 372 99
E-mail: liivika.tee@ut.ee

Education:

1996–2008	Valga Gymnasium
2008–2011	University of Tartu, mathematics, BSc
Spring, 2012	Johannes Kepler Universität (Linz, Austria), exchange semester
2011–2013	University of Tartu, finance and actuarial science, MSc
2016–2017	KU Leuven (Leuven, Belgium), visiting scholar
2013–	University of Tartu, mathematical statistics, doctoral studies

Languages: Estonian, English, Dutch, French, German

Employment:

02/2014–01/2016	University of Tartu, assistant of mathematical statistics (0.5)
02/2016–	KU Leuven, voluntarily researcher (0.2)

Fields of scientific interest:
Actuarial mathematics – non-life insurance, life insurance, loss modeling and claims reserving

Curriculum vitae

Nimi: Liivika Tee
Sünniaeg ja -koht: 29.04.1989, Valga, Eesti
Kodakondsus: Eesti
Kontaktaadress: TÜ matemaatika ja statistika instituut, J. Liivi 2, 50409 Tartu, Eesti
Telefon: +372 581 372 99
E-post: liivika.tee@ut.ee

Haridus:

1996–2008	Valga Gümnaasium
2008–2011	Tartu Ülikool, matemaatika, BSc
Kevad, 2012	Johannes Kepler Universität (Linz, Austria), vahetussemester
2011–2013	Tartu Ülikool, finants-ja kindlustusmatemaatika, MSc
2016–2017	KU Leuven (Leuven, Belgia), külalisdoktorant
2013–	Tartu Ülikool, matemaatiline statistika, doktoriõpe

Keelte oskus: eesti, inglise, hollandi, prantsuse, saksa

Teenistuskäik:

02/2014–01/2016	Tartu Ülikool, matemaatilise statistika assistent (0.5)
02/2016–	KU Leuven, vabatahtlik teadur (0.2)

Teaduslikud huvid:
Kindlustusmatemaatika – kahjukindlustus, elukindlustus, kahjujaotused ja reservide hindamine

List of original publications

1. Tee, L., M. Käärik, and R. Viin (2017). On comparison of stochastic reserving methods with bootstrapping. *Risks 5*(1), 2.

2. Tee, L. and M. Käärik (2017). Loss reserving on different levels of data aggregation: chain-ladder vs. its continuous extension. *Submitted*.

DISSERTATIONES MATHEMATICAE UNIVERSITATIS TARTUENSIS

1. **Mati Heinloo.** The design of nonhomogeneous spherical vessels, cylindrical tubes and circular discs. Tartu, 1991, 23 p.
2. **Boris Komrakov.** Primitive actions and the Sophus Lie problem. Tartu, 1991, 14 p.
3. **Jaak Heinloo.** Phenomenological (continuum) theory of turbulence. Tartu, 1992, 47 p.
4. **Ants Tauts.** Infinite formulae in intuitionistic logic of higher order. Tartu, 1992, 15 p.
5. **Tarmo Soomere.** Kinetic theory of Rossby waves. Tartu, 1992, 32 p.
6. **Jüri Majak.** Optimization of plastic axisymmetric plates and shells in the case of Von Mises yield condition. Tartu, 1992, 32 p.
7. **Ants Aasma.** Matrix transformations of summability and absolute summability fields of matrix methods. Tartu, 1993, 32 p.
8. **Helle Hein.** Optimization of plastic axisymmetric plates and shells with piece-wise constant thickness. Tartu, 1993, 28 p.
9. **Toomas Kiho.** Study of optimality of iterated Lavrentiev method and its generalizations. Tartu, 1994, 23 p.
10. **Arne Kokk.** Joint spectral theory and extension of non-trivial multiplicative linear functionals. Tartu, 1995, 165 p.
11. **Toomas Lepikult.** Automated calculation of dynamically loaded rigid-plastic structures. Tartu, 1995, 93 p, (in Russian).
12. **Sander Hannus.** Parametrical optimization of the plastic cylindrical shells by taking into account geometrical and physical nonlinearities. Tartu, 1995, 74 p, (in Russian).
13. **Sergei Tupailo.** Hilbert's epsilon-symbol in predicative subsystems of analysis. Tartu, 1996, 134 p.
14. **Enno Saks.** Analysis and optimization of elastic-plastic shafts in torsion. Tartu, 1996, 96 p.
15. **Valdis Laan.** Pullbacks and flatness properties of acts. Tartu, 1999, 90 p.
16. **Märt Põldvere.** Subspaces of Banach spaces having Phelps' uniqueness property. Tartu, 1999, 74 p.
17. **Jelena Ausekle.** Compactness of operators in Lorentz and Orlicz sequence spaces. Tartu, 1999, 72 p.
18. **Krista Fischer.** Structural mean models for analyzing the effect of compliance in clinical trials. Tartu, 1999, 124 p.
19. **Helger Lipmaa.** Secure and efficient time-stamping systems. Tartu, 1999, 56 p.
20. **Jüri Lember.** Consistency of empirical k-centres. Tartu, 1999, 148 p.
21. **Ella Puman.** Optimization of plastic conical shells. Tartu, 2000, 102 p.
22. **Kaili Müürisep.** Eesti keele arvutigrammatika: süntaks. Tartu, 2000, 107 lk.

23. **Varmo Vene.** Categorical programming with inductive and coinductive types. Tartu, 2000, 116 p.
24. **Olga Sokratova.** Ω-rings, their flat and projective acts with some applications. Tartu, 2000, 120 p.
25. **Maria Zeltser.** Investigation of double sequence spaces by soft and hard analitical methods. Tartu, 2001, 154 p.
26. **Ernst Tungel.** Optimization of plastic spherical shells. Tartu, 2001, 90 p.
27. **Tiina Puolakainen.** Eesti keele arvutigrammatika: morfoloogiline ühestamine. Tartu, 2001, 138 p.
28. **Rainis Haller.** *M(r,s)*-inequalities. Tartu, 2002, 78 p.
29. **Jan Villemson.** Size-efficient interval time stamps. Tartu, 2002, 82 p.
30. Töö kaitsmata.
31. **Mart Abel.** Structure of Gelfand-Mazur algebras. Tartu, 2003. 94 p.
32. **Vladimir Kuchmei.** Affine completeness of some ockham algebras. Tartu, 2003. 100 p.
33. **Olga Dunajeva.** Asymptotic matrix methods in statistical inference problems. Tartu 2003. 78 p.
34. **Mare Tarang.** Stability of the spline collocation method for volterra integro-differential equations. Tartu 2004. 90 p.
35. **Tatjana Nahtman.** Permutation invariance and reparameterizations in linear models. Tartu 2004. 91 p.
36. **Märt Möls.** Linear mixed models with equivalent predictors. Tartu 2004. 70 p.
37. **Kristiina Hakk.** Approximation methods for weakly singular integral equations with discontinuous coefficients. Tartu 2004, 137 p.
38. **Meelis Käärik.** Fitting sets to probability distributions. Tartu 2005, 90 p.
39. **Inga Parts.** Piecewise polynomial collocation methods for solving weakly singular integro-differential equations. Tartu 2005, 140 p.
40. **Natalia Saealle.** Convergence and summability with speed of functional series. Tartu 2005, 91 p.
41. **Tanel Kaart.** The reliability of linear mixed models in genetic studies. Tartu 2006, 124 p.
42. **Kadre Torn.** Shear and bending response of inelastic structures to dynamic load. Tartu 2006, 142 p.
43. **Kristel Mikkor.** Uniform factorisation for compact subsets of Banach spaces of operators. Tartu 2006, 72 p.
44. **Darja Saveljeva.** Quadratic and cubic spline collocation for Volterra integral equations. Tartu 2006, 117 p.
45. **Kristo Heero.** Path planning and learning strategies for mobile robots in dynamic partially unknown environments. Tartu 2006, 123 p.
46. **Annely Mürk.** Optimization of inelastic plates with cracks. Tartu 2006. 137 p.
47. **Annemai Raidjõe.** Sequence spaces defined by modulus functions and superposition operators. Tartu 2006, 97 p.
48. **Olga Panova.** Real Gelfand-Mazur algebras. Tartu 2006, 82 p.

49. **Härmel Nestra.** Iteratively defined transfinite trace semantics and program slicing with respect to them. Tartu 2006, 116 p.
50. **Margus Pihlak.** Approximation of multivariate distribution functions. Tartu 2007, 82 p.
51. **Ene Käärik.** Handling dropouts in repeated measurements using copulas. Tartu 2007, 99 p.
52. **Artur Sepp.** Affine models in mathematical finance: an analytical approach. Tartu 2007, 147 p.
53. **Marina Issakova.** Solving of linear equations, linear inequalities and systems of linear equations in interactive learning environment. Tartu 2007, 170 p.
54. **Kaja Sõstra.** Restriction estimator for domains. Tartu 2007, 104 p.
55. **Kaarel Kaljurand.** Attempto controlled English as a Semantic Web language. Tartu 2007, 162 p.
56. **Mart Anton.** Mechanical modeling of IPMC actuators at large deformations. Tartu 2008, 123 p.
57. **Evely Leetma.** Solution of smoothing problems with obstacles. Tartu 2009, 81 p.
58. **Ants Kaasik.** Estimating ruin probabilities in the Cramér-Lundberg model with heavy-tailed claims. Tartu 2009, 139 p.
59. **Reimo Palm.** Numerical Comparison of Regularization Algorithms for Solving Ill-Posed Problems. Tartu 2010, 105 p.
60. **Indrek Zolk.** The commuting bounded approximation property of Banach spaces. Tartu 2010, 107 p.
61. **Jüri Reimand.** Functional analysis of gene lists, networks and regulatory systems. Tartu 2010, 153 p.
62. **Ahti Peder.** Superpositional Graphs and Finding the Description of Structure by Counting Method. Tartu 2010, 87 p.
63. **Marek Kolk.** Piecewise Polynomial Collocation for Volterra Integral Equations with Singularities. Tartu 2010, 134 p.
64. **Vesal Vojdani.** Static Data Race Analysis of Heap-Manipulating C Programs. Tartu 2010, 137 p.
65. **Larissa Roots.** Free vibrations of stepped cylindrical shells containing cracks. Tartu 2010, 94 p.
66. **Mark Fišel.** Optimizing Statistical Machine Translation via Input Modification. Tartu 2011, 104 p.
67. **Margus Niitsoo**. Black-box Oracle Separation Techniques with Applications in Time-stamping. Tartu 2011, 174 p.
68. **Olga Liivapuu.** Graded q-differential algebras and algebraic models in noncommutative geometry. Tartu 2011, 112 p.
69. **Aleksei Lissitsin.** Convex approximation properties of Banach spaces. Tartu 2011, 107 p.
70. **Lauri Tart.** Morita equivalence of partially ordered semigroups. Tartu 2011, 101 p.
71. **Siim Karus.** Maintainability of XML Transformations. Tartu 2011, 142 p.

72. **Margus Treumuth.** A Framework for Asynchronous Dialogue Systems: Concepts, Issues and Design Aspects. Tartu 2011, 95 p.
73. **Dmitri Lepp.** Solving simplification problems in the domain of exponents, monomials and polynomials in interactive learning environment T-algebra. Tartu 2011, 202 p.
74. **Meelis Kull.** Statistical enrichment analysis in algorithms for studying gene regulation. Tartu 2011, 151 p.
75. **Nadežda Bazunova.** Differential calculus $d^3 = 0$ on binary and ternary associative algebras. Tartu 2011, 99 p.
76. **Natalja Lepik.** Estimation of domains under restrictions built upon generalized regression and synthetic estimators. Tartu 2011, 133 p.
77. **Bingsheng Zhang.** Efficient cryptographic protocols for secure and private remote databases. Tartu 2011, 206 p.
78. **Reina Uba.** Merging business process models. Tartu 2011, 166 p.
79. **Uuno Puus.** Structural performance as a success factor in software development projects – Estonian experience. Tartu 2012, 106 p.
80. **Marje Johanson.** $M(r, s)$-ideals of compact operators. Tartu 2012, 103 p.
81. **Georg Singer.** Web search engines and complex information needs. Tartu 2012, 218 p.
82. **Vitali Retšnoi.** Vector fields and Lie group representations. Tartu 2012, 108 p.
83. **Dan Bogdanov.** Sharemind: programmable secure computations with practical applications. Tartu 2013, 191 p.
84. **Jevgeni Kabanov.** Towards a more productive Java EE ecosystem. Tartu 2013, 151 p.
85. **Erge Ideon**. Rational spline collocation for boundary value problems. Tartu, 2013, 111 p.
86. **Esta Kägo.** Natural vibrations of elastic stepped plates with cracks. Tartu, 2013, 114 p.
87. **Margus Freudenthal.** Simpl: A toolkit for Domain-Specific Language development in enterprise information systems. Tartu, 2013, 151 p.
88. **Boriss Vlassov.** Optimization of stepped plates in the case of smooth yield surfaces. Tartu, 2013, 104 p.
89. **Elina Safiulina.** Parallel and semiparallel space-like submanifolds of low dimension in pseudo-Euclidean space. Tartu, 2013, 85 p.
90. **Raivo Kolde.** Methods for re-using public gene expression data. Tartu, 2014, 121 p.
91. **Vladimir Šor.** Statistical Approach for Memory Leak Detection in Java Applications. Tartu, 2014, 155 p.
92. **Naved Ahmed.** Deriving Security Requirements from Business Process Models. Tartu, 2014, 171 p.
93. **Kerli Orav-Puurand.** Central Part Interpolation Schemes for Weakly Singular Integral Equations. Tartu, 2014, 109 p.
94. **Liina Kamm.** Privacy-preserving statistical analysis using secure multi-party computation. Tartu, 2015, 201 p.

95. **Kaido Lätt.** Singular fractional differential equations and cordial Volterra integral operators. Tartu, 2015, 93 p.
96. **Oleg Košik.** Categorical equivalence in algebra. Tartu, 2015, 84 p.
97. **Kati Ain.** Compactness and null sequences defined by ℓ_p spaces. Tartu, 2015, 90 p.
98. **Helle Hallik.** Rational spline histopolation. Tartu, 2015, 100 p.
99. **Johann Langemets.** Geometrical structure in diameter 2 Banach spaces. Tartu, 2015, 132 p.
100. **Abel Armas Cervantes.** Diagnosing Behavioral Differences between Business Process Models. Tartu, 2015, 193 p.
101. **Fredrik Milani.** On Sub-Processes, Process Variation and their Interplay: An Integrated Divide-and-Conquer Method for Modeling Business Processes with Variation. Tartu, 2015, 164 p.
102. **Huber Raul Flores Macario.** Service-Oriented and Evidence-aware Mobile Cloud Computing. Tartu, 2015, 163 p.
103. **Tauno Metsalu.** Statistical analysis of multivariate data in bioinformatics. Tartu, 2016, 197 p.
104. **Riivo Talviste.** Applying Secure Multi-party Computation in Practice. Tartu, 2016, 144 p.
105. **Md Raknuzzaman.** Noncommutative Galois Extension Approach to Ternary Grassmann Algebra and Graded q-Differential Algebra. Tartu, 2016, 110 p.
106. **Alexander Liyvapuu.** Natural vibrations of elastic stepped arches with cracks. Tartu, 2016, 110 p.
107. **Julia Polikarpus.** Elastic plastic analysis and optimization of axisymmetric plates. Tartu, 2016, 114 p.
108. **Siim Orasmaa.** Explorations of the Problem of Broad-coverage and General Domain Event Analysis: The Estonian Experience. Tartu, 2016, 186 p.
109. **Prastudy Mungkas Fauzi.** Efficient Non-interactive Zero-knowledge Protocols in the CRS Model. Tartu, 2017, 193 p.
110. **Pelle Jakovits.** Adapting Scientific Computing Algorithms to Distributed Computing Frameworks. Tartu, 2017, 168 p.
111. **Anna Leontjeva.** Using Generative Models to Combine Static and Sequential Features for Classification. Tartu, 2017, 167 p.
112. **Mozhgan Pourmoradnasseri.** Some Problems Related to Extensions of Polytopes. Tartu, 2017, 168 p.
113. **Jaak Randmets.** Programming Languages for Secure Multi-party Computation Application Development. Tartu, 2017, 172 p.
114. **Alisa Pankova.** Efficient Multiparty Computation Secure against Covert and Active Adversaries. Tartu, 2017, 316 p.
115. **Tiina Kraav.** Stability of elastic stepped beams with cracks. Tartu, 2017, 126 p.
116. **Toomas Saarsen.** On the Structure and Use of Process Models and Their Interplay. Tartu, 2017, 123 p.
117. **Silja Veidenberg.** Lifting bounded approximation properties from Banach spaces to their dual spaces. Tartu, 2017, 112 p.