May 17, 2019

MASTER THESIS

ADAPTING A HIERARCHICAL GAUSSIAN PROCESS MODEL TO PREDICT THE LOSS RESERVE OF A NON-LIFE INSURER

P.L. Ruitenberg
Industrial Engineering and Management (IEM)
Specialisation Financial Engineering and Management (FEM)

Faculty of Behavioural, Management and Social Sciences (BMS)
Department Industrial Engineering and Business Information Systems (IEBIS)

Exam committee:
dr. B. Roorda (University of Twente)
dr. P.K. Mandal (University of Twente)
M.K. Lam, PhD. (KPMG)

KPMG

UNIVERSITY OF TWENTE.

Preface

Dear reader,

This thesis concludes approximately 8 months of research that I have conducted at the Financial Risk Management department of KPMG Advisory N.V., in order to conclude the master Financial Engineering and Management.

First of all, I want to thank my supervisor and colleagues of KPMG. I have had numerous meetings with Miekee during my internship, all of which were insightful and productive. No matter how big or small an issue was, Miekee was always willing to assist me, guide me in the right direction and supply me with feedback. Furthermore, I want to thank everyone who has helped me in the process of this research, including (but not limited to) Rinze, Helen, Peter and Luuk.

Furthermore, I want to thank my supervisors Berend and Pranab from the University of Twente. Both Berend and Pranab have been lecturers during my Master's study. Berend's lectures have sparked my interest in the field of Risk Management, while Pranab's expertise in statistics and Risk Theory has been inspirational. The meetings with Berend and Pranab during this research have been productive, and their input has ensured the academic level of this thesis.

Delivering this thesis also concludes my student life. During these eight years, I have met many people and made a lot of friends, be it by working together, being in a committee (or board!) together, living in the same home or just by being in "Beneden Peil" at the same time. I want to thank everyone who has been a part of this for making it as unforgettable as it has been.

Last, but definitely not least, I want to thank my family, but my parents in particular, for their unconditional support.

I hope you enjoy reading this thesis.

Patrick Ruitenberg
Amstelveen, May 8th, 2019.

Executive Summary

In this thesis, we have researched potential improvements of a hierarchical Gaussian process model on the actuarial challenge of predicting the Loss Reserve of a non-life insurer. The Hierarchical Gaussian process model is described in a paper by Lally and Hartman (2018) and is able to give adequate Best Estimate predictions, but the uncertainty of the prediction results in a wider confidence interval than other methods currently applied in actuarial practice.

The research performed consists of multiple parts: The model performance and design is validated, by assessing the model performance on a more extensive data set. The data set contains historical losses of 1988-2006 of multiple insurers of multiple Lines of Business. Furthermore, we will validate design choices by varying the prior distributions on hyperparameters in the model to other weakly informative priors used in literature on Gaussian processes.

Furthermore, we have researched if the model can be extended with external information, outside of the run-off triangle. We research various methods, including Bornheutter-Ferguson, to incorporate premium information in the model, and assess if performance is improved.

We compare the GP model by Lally and Hartman (2018) with the Chain Ladder method - which is commonly used in actuarial practice. Furthermore, we use the model results of the model by Lally and Hartman (2018) as a benchmark for our variations in design choices and model extension. As the data set used contains all observed losses, we can compare performance of the Best Estimate of the Loss Reserve, and the Root Mean Square Error of prediction of the estimated losses. Furthermore, we analyse the density of the prediction of the Loss Reserve.

Our results indicate that the performance on volatile run-off triangles by the model of Lally and Hartman (2018) still has room for improvement. However, in some cases the Gaussian process model outperforms the Chain Ladder. The design choices of the prior distributions of the hyperparameters made by Lally and Hartman (2018) are adequate. Most notably, changing the prior distribution of the Bandwidth parameter applied to a $\text{Cauchy}(0, 2.5)$ distribution has varying results, and should not be used in combination with a Squared Exponential covariance function. For the model in general, we conclude that most weakly informative priors give adequate results.

With regards to supplying the model with premium information, the results show that a transformation to Loss Ratio's have a positive effect on the prediction of the Best Estimate of the Loss Reserve. Adding a Bornheutter-Ferguson estimation to the model, and allowing the Gaussian Process model to add noise to these estimations, give good results regarding to the uncertainty of the prediction. However, this implementation is very reliant on the quality of these estimations, as the Best Estimate of the Loss Reserve is reliant on them.

We conclude that the Gaussian process model is well designed and generally applicable on a

multitude of data sets. However, we recommend testing the model performance on incremental data, or run-off triangles of incurred claims. We also recommend tuning the kernel function of the model to triangle specific characteristics. As for implementing external data in the model, other data such as the number of claims could be researched. Moreover, validating a faster Gaussian process approximation as described by Flaxman et al. (2016) is recommended.

Contents

1 Introduction **1**
1.1 Organisation 1
1.2 Loss Reserve of an non-life insurer 1
1.3 Gaussian Process regression 3
1.4 Research 5
1.4.1 Research Questions 5
1.4.2 Methodology 5

2 Predicting Loss Reserves **7**
2.1 An example of a Claim Timeline 7
2.2 The Run-off Triangle 8
2.3 Predicting the Loss Reserve 8
2.3.1 Chain Ladder 9
2.3.2 Bornheutter-Ferguson 10
2.3.3 Recent Research for Loss Reserve prediction 13

3 A Gaussian process model to predict loss reserves **15**
3.1 Introduction on the model 15
3.2 Covariance functions 16
3.2.1 Squared Exponential 17
3.2.2 Matérn 3/2 and Matérn 5/2 17
3.2.3 Signal and Noise 18
3.3 Input Warping 19
3.4 Input data preparation 19
3.5 Hyperparameters and prior distributions 21
3.6 Model limitations 22

4 Research setup **25**
4.1 Methodology 26
4.2 Prior distribution of hyperparameters 26
4.2.1 Bandwidth parameter 26
4.2.2 Signal and noise parameters 27
4.2.3 Warping parameters 27
4.3 Extending the model with premium information 29
4.3.1 Transform observations to Loss Ratio's 29
4.3.2 Supply Premiums as input instead of Accident Years 29

4.3.3 Add Estimators of the Ultimate Loss to the model 30

5 Data set used **33**
5.1 Medical Malpractice (MM) 33
5.1.1 Physicians Reciprocal Insurers (Triangle 1.1) 34
5.1.2 Promutual Group (Triangle 1.2) 34
5.1.3 Scpie Indemnity Company (Triangle 1.3) 36
5.2 Commercial Automobile (CA) 36
5.2.1 Farmers' Automobile (Triangle 2.1) 36
5.2.2 State Farm (Triangle 2.2) 38
5.2.3 The Ins Co (Triangle 2.3) 38
5.3 Personal Automobile (PA) 40
5.3.1 Farmers' Automobile (Triangle 3.1) 40
5.3.2 Federal Insurance Company (Triangle 3.2) 40
5.3.3 State Farm (Triangle 3.3) 42
5.4 Product Liability (PL) 42
5.4.1 Allstate Insurance Company (Triangle 4.1) 42
5.4.2 Federal Insurance Company (Triangle 4.2) 44
5.4.3 Federated Mutual Group (Triangle 4.3) 44
5.5 Workers' Compensation (WC) 46
5.5.1 Allstate Insurance Company (Triangle 5.1) 46
5.5.2 Lumbermen's Underwriting Alliance (Triangle 5.2) 46
5.5.3 State Farm (Triangle 5.3) 48

6 Results **49**
6.1 Benchmark measurements 49
6.2 Prior distributions 51
6.2.1 Bandwidth parameter (ψ) 51
6.2.2 Noise parameter (σ^2) 54
6.2.3 Signal parameter (η^2) 57
6.2.4 Warping parameters (α, β) 58
6.2.5 Optimal prior configuration 59
6.3 Premium Information 63
6.3.1 Loss Ratio's 63
6.3.2 Premium as input 67
6.3.3 Bornheutter-Ferguson Estimations 71

7 Conclusion **75**
7.1 Data applicability 75
7.2 Prior distributions and model design 75
7.3 Adding premium information to the model 76

8 Discussion and Recommendations **77**
8.1 Data set 77
8.2 Priors on Loss Ratio's 78
8.3 Bornheutter-Ferguson Estimation 78
8.4 Implementation in practice 78

8.5 Technical possibilities and limitations 79

References **81**

Appendix A Glossary of abbreviations **85**

Appendix B Run-off triangle: Example **87**

Appendix C Run-off triangles of the data set **89**

Appendix D Results: 3D-plots of benchmark measurements **99**

Appendix E Results: Inference of prior parameters **115**

E.1 Bandwidth parameter (ψ) 115

E.2 Noise parameter (σ^2) 117

E.3 Signal parameter (η^2) 117

E.4 Warping parameters ($\alpha_1, \beta_1, \alpha_2, \beta_2$) 117

Appendix F Results: Bornheutter-Ferguson Estimators **121**

1 | Introduction

Contents

1.1 Organisation **1**
1.2 Loss Reserve of an non-life insurer **1**
1.3 Gaussian Process regression **3**
1.4 Research **5**
1.4.1 Research Questions 5
1.4.2 Methodology 5

This chapter will give a brief introduction to this thesis, discussing the basics of the loss reserve that we wish to estimate, an introduction on a Gaussian process and applying it for regression, and the organisation at which this research has been performed. This section will conclude with our research question and methodology.

1.1 | Organisation

This research has been performed from the 1st of October 2018 to the 17th of May 2019 at KPMG Advisory N.V. in Amstelveen, at the department of Financial Risk Management (FRM). This department advises financial institutions on various risk-related topics. Clients of this department include, for instance, large banks, insurers and pension funds. Topics that FRM advises on can range from regulatory questions (i.e. Basel IV, Solvency II), advising on Mergers & Acquisitions and assisting the colleagues of KPMG Accountants N.V.. As such, the FRM department is interested in gaining knowledge in more sophisticated methods to adequately model risks.

1.2 | Loss Reserve of an non-life insurer

Insurers are inherently, due to the nature of their business, exposed to various sources of risk. Insurance companies take on a specific risk that an individual (or business) wants to hedge, at the cost of a fixed premium (Kaas, Goovaerts, Dhaene, & Denuit, 2008). Besides the risks that are inherent to this business model (*Insurance Risk*), insurers have exposure to the financial market (*Market Risk*), there are several *Business Risks*, such as unforeseen large expenses, and they also have an *Operational Risk*, for instance when safeguards or internal models fail.

In insurance, there is a distinction in Life and Non-Life insurance companies. A life insurer writes life policies (e.g.: in case of death of the policyholder, pay out a specified amount to the beneficiaries), and a non-life (also known as Property & Casualty) insurance company writes all kind other insurance policies except life insurance, such as automobile and healthcare insurance.

This separation is made for both legal reasons and due to the difference between the two products: contract terms and claim types are different. As such, life and non-life insurance is usually modelled differently. (Wüthrich & Merz, 2006)

A non-life insurer usually has two main actuarial reserves: a premium reserve and a loss reserve. The premium reserve of a non-life insurer consists of premiums that are expected to still be received. In this thesis, the loss reserve of an non-life insurance company will be the main focus. A loss reserve contains provisions for payment obligations from losses that have occurred, but have not yet been settled (Radtke, Schmidt, & Schnaus, 2016). As this reserve should cover losses/payments that will occur in the future, there is a considerable amount of uncertainty while attempting to adequately assess the size of this reserve. This poses a risk to the insurance company: when the loss reserve is too low, it can run into problems when making payments on claims. On the other hand, when too much money is pooled in the loss reserve, this might be detrimental to business - as that money cannot be used for other purposes.

Distinctions can be made in determining this loss reserve, usually along several Lines of Business (LoB's) that a non-life insurer has, such as commercial or personal automobile. Every branch can have a unique pattern in claims development, and claims usually have a delay in settlement unique to the type of claim. For instance, liability products can have a substantial delay due to litigation/lawsuits (Kaas et al., 2008). As such, a distinction can be made in short-tail LoB's (of which losses are known relatively quickly) and long-tail LoB's (which take longer to develop).

The Loss Reserve is built up by of two types of claims: claims that are IBNR or RBNS. Claims that are known to the insurer, but have not yet been paid in full yet, are referred to as *RBNS: Reported, But Not Settled.* Furthermore, claims that have been incurred, but are not yet known to the insurance company are referred to as *IBNR: Incurred, but not reported.* Both types of claims form a future payment obligation, and as such are included in the loss reserve.

This research focuses on improving the prediction of the loss reserve. For this, historical payment data is considered in order to model the size of this reserve. Depending on the Line of Business, and/or business size, the level of detail of this data may be yearly, monthly or even daily. The lag between when the claim is incurred and when a payment is made is known as the development lag.

Incurral Year	Development lag				
	1	2	3	4	5
1997	1,188,675	3,446,584	4,141,821	4,308,633	4,400,762
1998	1,235,402	4,485,415	5,135,343	5,346,687	·
1999	2,209,850	5,928,544	6,746,912	·	·
2000	2,662,546	6,149,580	·	·	·
2001	2,457,265	·	·	·	·

Table 1.1: Example of a cumulative run-off triangle. Source: Frees (2009)

The aggregated data is usually illustrated in a so-called run-off triangle, of which an example is given in Table 1.1[1]. In any run-off triangle, each payment is categorised by its' Incurral Year and a Development Lag. This triangle has a total time span of 5 years. The Incurral Year is defined as the origin of the claim: The claim was incurred on a policy written in that specific year. However, there is a delay in payments. In 1997, a total of €1,188,675 was paid. In the

[1]A more elaborate (but simplified) example is given in Appendix B

subsequent year (1998), thus in development year 2, this insurer has paid out a grand total of €3,446,584 for claims incurred in 1997, and so forth.

For both the insured and the insurer, it is important that all (previously) incurred claims can be paid. These payments will have to be made from our loss reserve. Referring back to our example in Table 1.1, we want to estimate the loss reserve by inferring the blanks in the triangle. For this, several methods are available. We will elaborate on some of the methods in Chapter 2.

When making an estimation of risks and/or reserves, a distinction should be made between a Best Estimate (BE) and the uncertainty of this estimate. When using a strictly discrete method, the outcome of such a model would strictly result in a point estimate. On the other hand, when modelling a risk, a realistic assumption of the spread of the possible outcomes is of interest, and thus the volatility that underlies the estimation. Solvency II requires insurers to determine their provisions (such as the loss reserve) at Best Estimate.

1.3 Gaussian Process regression

A Gaussian process (GP) is stochastic process, by generalisation of the Gaussian distribution (Rasmussen & Williams, 2006). Stochastic processes are commonly known in Mathematical/Financial literature. An example of a financial application of a stochastic process is the Black-Scholes formula used to price the value of option contracts, which applies Geometric Brownian Motion to model a random process. (Black & Scholes, 1973). We define a Gaussian process in Definition 1.1.

Definition 1.1: Gaussian process

A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution.

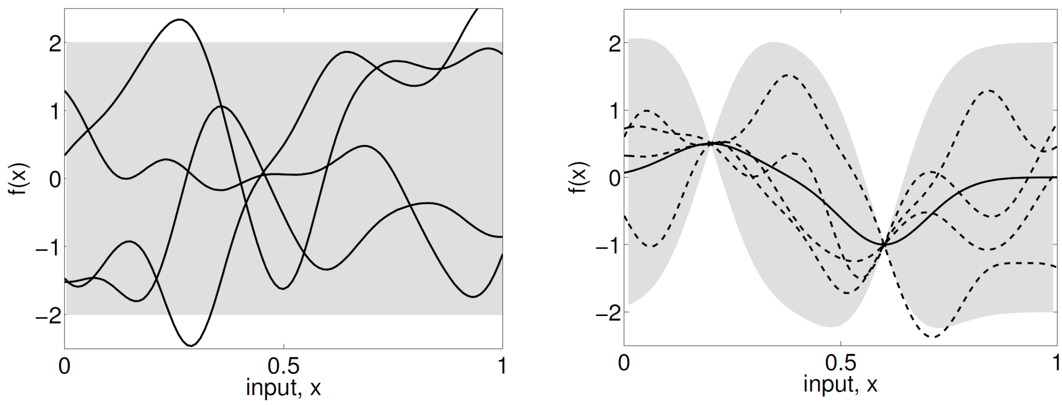
Source: Rasmussen and Williams (2006)

In a GP, we assume that the underlying process $f(x)$ follows a Gaussian distribution. Where a Gaussian Distribution is defined by its' mean and covariance *parameters*, a Gaussian Process is defined by it's mean and covariance *function*. We define $m(x)$ as the mean function and $k(x, x')$ as the covariance function. An example of a GP is given mathematically in (1.1), where notation of Rasmussen and Williams (2006) is followed.

$$\begin{aligned} m(x) &= \mathbb{E}[f(x)], \\ k(x, x') &= \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))], \\ \\ f(x) &\sim GP(m(x), k(x, x')) \end{aligned} \tag{1.1}$$

For simplification purposes, the mean of the GP is often set equal to zero (Rasmussen & Williams, 2006). Therefore, a GP is most commonly defined by its' covariance function: also known as the kernel function (Rasmussen & Williams, 2006).

A Gaussian process can be applied for regression purposes. The goal of any regression is to find a relationship between various variables: given input parameters, we want to predict a certain output that is related to these inputs. A GP regression intends to estimate a function that can relate a series of known inputs to a series of outputs. Here, the input vector will be



(a) Prior

(b) Posterior

Figure 1.1: (a): 4 draws from the prior distribution, with mean 0 and standard deviation 1. There are no known observations. (b) 4 draws from the posterior distribution, after adding two noiseless observations. The solid line indicates the mean, the shaded region indicates a distance of two standard deviations. Source: Rasmussen and Williams (2006)

defined as $x = \{x_1, x_2, ..., x_n\}$, and the corresponding output as $f(x)$. In case a value of $f(x_i)$ is known for a certain value of x_i, that data point will be referred to as an observation. The covariance function is determined on the input vector x, which is known for all data points, both the data with a known observation $f(x_i)$ and the data we want to infer. An example of a covariance function is the squared exponential function. Taking fictive inputs x_p and x_q, this function is given by:

$$\text{cov}(f(x_p), f(x_q)) = k(x_p, x_q) = \exp(-\frac{1}{2}|x_p - x_q|^2) \tag{1.2}$$

The covariance can thus be defined for any input pair (x_p, x_q), resulting in a covariance matrix K of size $[n, n]$ which describes the covariance between any two inputs.

The covariance function is one of the various design choices of the model. Based on the data and application, it is assumed that the actual process could be described by this covariance function.

Without any observations $f(x_i)$ this results in our Bayesian *prior distribution* that is based on the model design choices. By adding observations of $f(x_i)$ (i.e. adding evidence), this results in the *posterior distribution* conditioned on the observations. Figure 1.1 gives an example of the difference between generating draws from the prior and posterior distribution of a one-dimensional Gaussian process, by adding two noiseless observations.

As displayed in Figure 1.1, the noiseless observations result in changes in the mean and covariance. The *posterior distribution* is determined by applying the *joint gaussian distribution* property of the Gaussian process as described in Definition 1.1. X_1 is defined as the input that belongs to *observations* $f(X_1)$ and X_2 as the input that belongs to the *unknown output* $f(X_2)$. As $f(X_1)$ and $f(X_2)$ are generated from the same underlying process, the joint distribution of $f(X_1)$ and $f(X_2)$ can be described as follows:

$$\begin{bmatrix} f(X_1) \\ f(X_2) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m(X_1) \\ m(X_2) \end{bmatrix}, \begin{pmatrix} k(X_1, X_1) & k(X_1, X_2) \\ k(X_2, X_1) & k(X_2, X_2) \end{pmatrix} \right) \tag{1.3}$$

For simplicity we assume $m(X_1) = m(X_2) = 0$. The posterior distribution of $f(X_2)$ is then given by conditioning it on the observations:

$$f(X_2)|f(X_1), X_1, X_2 \sim \mathcal{N}\big(k(X_2, X_1)k(X_1, X_1)^{-1}f(X_1), \\ k(X_2, X_2) - k(X_2, X_1)k(X_1, X_1)^{-1}k(X_1, X_2)\big) \tag{1.4}$$

Equation 1.4 gives the posterior distribution of the Gaussian process, which we can then use for inferring $f(X_2)$.

1.4 | Research

In a paper by Lally and Hartman (2018), it has been described that a hierarchical Gaussian process model can be used in order to estimate loss reserves, by using the run-off triangle as input. While the predictions of the Best Estimate were comparable (if not better) to the current industry practice and recent research, the uncertainty surrounding the predictions is larger than would be expected in practice. Having an adequate measurement of the expected uncertainty and volatility is important in order to determine the risk that an insurer has. This makes the model in its' current state unfeasible for use in actuarial practice, as an accurate estimation of the uncertainty in the predictions made is important. As such, it is not yet an appropriate model of the true risk involved. Therefore, we will research various attempts to reduce the uncertainty, and improve the predictions that are made by the model.

1.4.1 | Research Questions

Given the model limitations described, we want to both improve the best estimate of the model, reduce the confidence interval of these results and validate the model design. Our main research question is as follows:

Can we improve the Hierarchical Gaussian Process model
to predict the loss reserve of a non-life insurer?

With the following sub-questions:

- In order to validate the design choices:
  - Is the model applicable on a more extensive data set?
  - Are the prior distributions on the hyperparameters adequately chosen?
- In order to improve the Best Estimate and/or reduce the confidence interval:
  - Can relevant, out-of-triangle information be supplied to the model?
  - Can the GP model be extended with a Bornheutter-Ferguson estimation method?

1.4.2 | Methodology

The results of our variations on the Gaussian process model is able to give a representation of the loss reserve and the uncertainty therein. As such, the results of the model can be compared to both the current industry standard (The Chain Ladder method (Shi & Frees, 2011; Côté, Genest, & Abdallah, 2016)) and the current Gaussian Process model in the paper of Lally and Hartman (2018). Furthermore, a data set is publicly available which contains loss reserving data from insurers in multiple Lines of Business in the USA (Meyers & Shi, 2011). This data set contains the actual observed losses, and can be used to both train our model and quantify the error of all models by using it as a testing set. As such, we have a ground truth that can be used for analysis, determining if the predicted Loss Reserve is sufficient and what the error statistics (Root Mean Square Error) are in order to assess performance.

We have selected a 3 different triangles which represent the cumulative paid loss along 5 different Lines of Business, resulting in a total of 15 triangles that will be analysed. These triangles are presented in Appendix B, and their key characteristics will be explained in Chapter 5.

2 | Predicting Loss Reserves

Contents

2.1 An example of a Claim Timeline **7**
2.2 The Run-off Triangle **8**
2.3 Predicting the Loss Reserve **8**
2.3.1 Chain Ladder 9
2.3.2 Bornheutter-Ferguson 10
2.3.3 Recent Research for Loss Reserve prediction 13

Central in this thesis is the prediction of the loss reserve of a non-life insurer. In order to clarify this, this chapter will first discuss the typical timeline of an insurance claim. After that, the historical developments of claims will be discussed and how they are usually presented in a spreadsheet. This is subsequently used as input for predicting the reserve.

2.1 | An example of a Claim Timeline

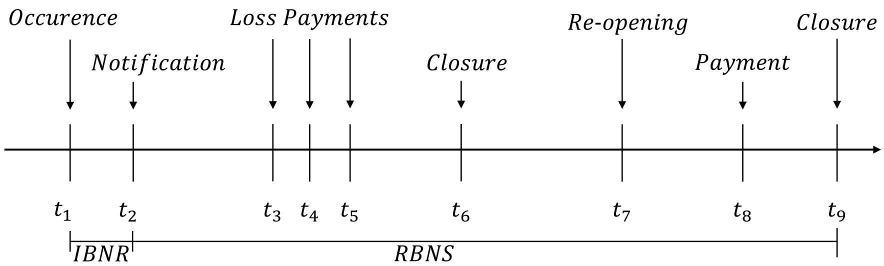


Figure 2.1: A typical timeline of a claim. Based on Antonio and Plat (2014)

The typical timeline of a claim is displayed schematically in Figure 2.1. At a certain moment in time, the claim (damage) occurs (t_1). There is a certain delay when this is reported to the insurer, which occurs at moment t_2. When the notification is made, (several) payments have to be made (t_3, t_4, t_5) until the dossier is closed (t_6). After this, it could occur that a dossier is re-opened in case new information becomes available. In that case, a final payment has to be made until it is closed again (t_7, t_8, t_9).

A claim is *Incurred, but not Reported* (IBNR) during the time span between the occurrence and notification of a claim to the insurer. After that, it is RBNS: *Reported, but not Settled* - as payments still have to be made on the claim. Moreover, it can be IBNER: *Incurred, but not enough reported*, for instance when an initial assessment of the claim results in an insufficient amount reserved for payments.

2.2 | The Run-off Triangle

As described in the previous section, payments or notifications of claims occur at different points in time, even if they originate from the same moment in time. An overview of the cumulative payment size can be generated, based on their moment in time that the claim originated (i.e. was incurred), as well as the moment in time when the insurer has made payments on this claim. Furthermore, we can define a specific calendar year at which the observed loss occurred in time. Properties of the run-off triangle are determined as:

$i =$ Incurral year

$j =$ Development lag

$k =$ Calendar Year $= i + j$

$n =$ No. of development years[1] = max(j)

$Z_{i,j} =$ Observed loss

$\hat{Z}_{i,j} =$ Estimated loss

$\hat{R}_i =$ Loss reserve for Incurral year i

$\hat{R} =$ Total loss reserve $= \sum_{i=1}^{n} \hat{R}_i$

With i as the incurral year - starting at 1, j as the development lag and $Z_{i,j}$ as the loss at that specific moment, a spreadsheet of these claims can be created. This is commonly known as a loss triangle or run-off triangle. An example is given in Example 2.1, and a more elaborate example is given in Appendix B.

Example 2.1: Run-off Triangle

	Development lag (j)				
Incurral Year (i)	1	2	3	4	5
1997 (1)	1,188,675	3,446,584	4,141,821	4,308,633	4,400,762
1998 (2)	1,235,402	4,485,415	5,135,343	5,346,687	·
1999 (3)	2,209,850	5,928,544	6,746,912	·	·
2000 (4)	2,662,546	6,149,580	·	·	·
2001 (5)	2,457,265	·	·	·	·

Example of a cumulative run-off triangle. Source: Frees (2009)

The loss paid in 1997, claimed in that year is defined as $Z_{1,1}$ = 1,188,675, and so forth.

2.3 | Predicting the Loss Reserve

Prediction of the loss reserve is done by completing the run-off triangle into a rectangle with adequate estimations. Referring to Example 2.1, the data points that are marked with a $(\cdot)$ are to be estimated. The summation of the differences between the estimated cumulative ultimate loss (at $j = n$) and the most recent observation for all incurral years ($j = n + 1 - i$) is the total loss reserve ($\hat{R}$):

$$\hat{R}_i = \begin{cases} 0 & \text{if } i = 1 \\ \hat{Z}_{i,n} - Z_{i,n+1-i} & \text{if } i > 1 \end{cases} \tag{2.1}$$

[1] The number of development years is usually identical to the number of incurral years. i.e.: $max(i) = max(j)$.

2.3.1 Chain Ladder

The most common and well-known method of predicting the Loss Reserve is the Chain Ladder. It has been popularised by the papers of Mack (1993, 1999) and is widely described in various books and papers. Using i as incurral year and j as development lag, historical development factors can be determined, which can be used to make estimations of future losses and the loss reserve[2]:

$$\begin{aligned} DF_j &= \frac{\sum_{i=1}^{n+1-j} Z_{i,j}}{\sum_{i=1}^{n+1-j} Z_{i,j-1}} \\ \hat{Z}_{i,j} &= Z_{i,n+1-i} \prod_{l=n-i+2}^{j} DF_l \end{aligned} \tag{2.2}$$

Example 2.2: Chain Ladder

The following (fictive) run-off triangle is given:

	Development lag (j)		
Incurral Year (i)	1	2	3
2010 (1)	1,000	1,500	1,750
2011 (2)	1,250	1,700	·
2012 (3)	1,400	·	·

The Chain Ladder method as given in (2.2) is applied. First, DF_2 and DF_3 are calculated:

$$\begin{aligned} DF_3 &= \frac{\sum_{i=1}^{3+1-3} Z_{i,3}}{\sum_{i=1}^{3+1-3} Z_{i,2}} = \frac{1,750}{1,500} \approx 1.17 \\ DF_2 &= \frac{\sum_{i=1}^{3+1-2} Z_{i,2}}{\sum_{i=1}^{3+1-2} Z_{i,1}} = \frac{1500+1700}{1000+1250} \approx 1.42 \end{aligned}$$

These development factors are used to calculate estimates for future losses:

$$\begin{aligned} \hat{Z}_{2,3} &= Z_{2,2} \prod_{l=3}^{3} DF_l \approx 1,700 * 1.17 && \approx 1,983 \\ \hat{Z}_{3,2} &= Z_{3,1} \prod_{l=2}^{2} DF_l \approx 1,400 * 1.42 && \approx 1,991 \\ \hat{Z}_{3,3} &= Z_{3,1} \prod_{l=2}^{3} DF_l \approx 1,400 * 1.42 * 1.17 && \approx 2,323 \end{aligned}$$

The estimated losses and corresponding reserve (by following Equation 2.1) are thus:

	Development lag (j)			
Incurral Year (i)	1	2	3	Reserve ($\hat{R}_i$)
2010 (1)	1,000	1,500	1,750	0
2011 (2)	1,250	1,700	**1,983**	283
2012 (3)	1,400	**1,991**	**2,323**	923
Development Factors		1.42	1.17	1,206 (= $\hat{R}$)

[2]Both $\hat{Z}_{i,j}$ and DF_j are indexed on j. To avoid confusion, a temporary variable l is introduced in (2.2)

2.3.2 Bornheutter-Ferguson

The Bornheutter-Ferguson (BF) method is also commonly used for determining the loss reserve, and is named after the authors of the paper in which it is first described (Bornhuetter & Ferguson, 1972). The BF method relies on both an estimated ultimate loss and estimated development through time, and uses these in order to predict the loss reserve. These estimates can be based on data inside the triangle (therefore only consisting of previously observed losses) or external data, such as the number of claims or premium volume. Considering that the method is heavily reliant on these estimators, the quality and accuracy of these estimators is crucial to result in adequate predictions.

First, definitions used throughout this section will be given:

$$\begin{aligned}\hat{\alpha}_i &= \text{Estimated ultimate loss in Incurral Year } i\\ DF_j &= \text{Chain Ladder development factor of development lag } j\\ \hat{\gamma}_j &= \text{Estimated cumulative development parameter up to development lag } j\\ \hat{\vartheta}_j &= \text{Estimated incremental development parameter of development lag } j\end{aligned}$$

Definition 2.1: Development Pattern

There exist parameters $\gamma_1, \gamma_2, \ldots, \gamma_n$ with $\gamma_n = 1$ such that the identity

$$\frac{\mathbb{E}[Z_{i,j}]}{\mathbb{E}[Z_{i,n}]} = \gamma_j$$

holds for all $j \in 1, 2, \ldots, n$ and for all $i \in 1, 2, \ldots, n$

Source: Radtke et al. (2016), with adaptations to match notation in this thesis.

The Bornheutter-Ferguson method implies (similar to the Chain Ladder method) that at the final development lag observed in the triangle, all claims are incurred[3]. As such: $\hat{\gamma}_n = 1$, as also mentioned in Definition 2.1.

We can estimate losses at a specific moment in time (i, j) by estimating the development that still has to occur, multiplying that with the estimated ultimate loss and adding that to the most recently observed loss:

$$\hat{Z}_{i,j} = Z_{i,n-i+1} + \Big(\sum_{l=n-i+2}^{j} \hat{\vartheta}_l\Big) * \hat{\alpha}_i \tag{2.3}$$

It should be noted that are several methods available to make estimations for both $\hat{\alpha}$ and $\hat{\vartheta}$. It goes beyond the scope of this research to mention all of them and their pros and cons. Instead, we will elaborate on the methods applied in this research - which is the original version of the BF method (Radtke et al., 2016).

Estimation of the cumulative and incremental development parameters $\hat{\gamma}$ and $\hat{\vartheta}$ can be performed using the Chain Ladder development factors[4]:

[3] If this assumption does not hold, a tail factor can be introduced. However, this is out of scope for this thesis.
[4] Both DF and γ are indexed on j. To avoid confusion, a temporary variable m is introduced in (2.4)

$$\hat{\gamma}_j = \begin{cases} 1 & \text{if } j = n \\ (\prod_{m=j+1}^{n} DF_m)^{-1} & \text{elsewhere} \end{cases}$$
$$\hat{\vartheta}_j = \begin{cases} \hat{\gamma}_j & \text{if } j = 1 \\ \hat{\gamma}_j - \hat{\gamma}_{j-1} & \text{elsewhere} \end{cases} \tag{2.4}$$

Example 2.3: Bornheutter-Ferguson

Consider the same triangle as in Example 2.2, but this is expanded with estimated losses:

Incurral Year (i)	Development lag (j) 1	2	3	Est. Ultimate Loss ($\hat{\alpha}_i$)
2010 (1)	1,000	1,500	1,750	1,750
2011 (2)	1,250	1,700	·	2,039
2012 (3)	1,400	·	·	2,209
Chain Ladder Factors (DF_j)		1.42	1.17	
Cum. Development Parameters	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	
Inc. Development Parameters	$\hat{\vartheta}_1$	$\hat{\vartheta}_2$	$\hat{\vartheta}_3$	

Using the Chain Ladder Factors and (2.4), the development parameters can be estimated:

$$\begin{aligned} \hat{\gamma}_3 &= 1 \\ \hat{\gamma}_2 &= (1.17)^{-1} \approx 0.85 \\ \hat{\gamma}_1 &= (1.42 * 1.17)^{-1} \approx 0.60 \\ \hat{\vartheta}_1 &= 0.60 \\ \hat{\vartheta}_2 &= 0.85 - 0.60 \approx 0.25 \\ \hat{\vartheta}_3 &= 1 - 0.85 \approx 0.15 \end{aligned}$$

Which can be used to estimate the future losses:

$$\begin{aligned} \hat{Z}_{2,3} &= 1,700 + 0.15 * 2,039 && \approx 2,006 \\ \hat{Z}_{3,2} &= 1,400 + 0.25 * 2,209 && \approx 1,952 \\ \hat{Z}_{3,3} &= 1,400 + (0.25 + 0.15) * 2,209 && \approx 2,283 \end{aligned}$$

Resulting in the following triangle and loss reserve by applying (2.1):

Incurral Year (i)	Development lag (j) 1	2	3	$\hat{\alpha}_i$	$\hat{R}_i$
2010 (1)	1,000	1,500	1,750	1,750	0
2011 (2)	1,250	1,700	**2,006**	2,039	306
2012 (3)	1,400	**1,952**	**2,283**	2,209	883
Chain Ladder Factors (DF_j)		1.42	1.17		1,189
Cum. Development Parameters	0.60	0.85	1		
Inc. Development Parameters	0.60	0.25	0.15		

It should be noted that, while counterintuitive, the estimated loss and the calculated ultimate loss need not be the same, i.e. it is possible that $\hat{Z}_{i,n} \neq \hat{\alpha}_i$. This is inherent to the BF-method, as $\hat{\alpha}_i$ is only used for the undeveloped part of the incurral year i when determining $\hat{Z}_{i,n}$.

A method to estimate the ultimate loss ($\hat{\alpha}$) is now derived. For this, data outside of the triangle is used, in the form of premiums received (net of reinsurance). Using these premiums, the Loss Ratio (LR) can be determined at every observed loss, and the expected ultimate Loss Ratio can be determined, based on the expected development that still has to occur:

$$\begin{aligned} P_i &= \text{Premium received in incurral year } i\text{, net of reinsurance} \\ \hat{LR}_{i,j} &= \frac{\frac{Z_{i,j}}{\hat{\gamma}_j}}{P_i} \end{aligned} \tag{2.5}$$

Using these Loss Ratio's, the expected ultimate loss $\hat{\alpha}_i$ is calculated after determining an estimated $\hat{LR}_i$ for each incurral year:

$$\hat{\alpha}_i = P_i * \hat{LR}_i \tag{2.6}$$

Example 2.4: Estimating the ultimate loss

Once again, the triangle given in Example 2.2 is considered, but premiums received are included.

	Development lag (j)			
Incurral Year (i)	1	2	3	(P_i)
2010 (1)	1,000	1,500	1,750	2,060
2011 (2)	1,250	1,700	·	2,400
2012 (3)	1,400	·	·	2,600
Cum. Dev. Parameters ($\hat{\gamma}_j$)	0.60	0.85	1	

By applying (2.5), the estimated Loss Ratio is determined for every observation. This is explicitly calculated for $\hat{LR}_{1,1}$, and (2.5) is also applied on the other observed losses:

$$\hat{LR}_{1,1} = \frac{\frac{1{,}000}{0.60}}{2,060} \approx 0.81$$

	Development lag (j)			
Incurral Year (i)	1	2	3	P_i
2010 (1)	0.81	0.85	0.85	2,060
2011 (2)	0.87	0.83	·	2,400
2012 (3)	0.89	·	·	2,600
Cum. Dev. Parameters ($\hat{\gamma}_j$)	0.60	0.85	1	

Based on the average of these Estimated Loss Ratio's, the ultimate Loss Ratio is approximated at 0.85 for all incurral years. (2.6) is applied to estimate the ultimate losses $\hat{\alpha}_i$:

$$\hat{\alpha}_1 = P_1 * \hat{LR}_i = 2,060 * 0,85 = 1,750$$

$$\hat{\alpha}_2 = P_2 * \hat{LR}_i = 2,400 * 0,85 \approx 2,039$$

$$\hat{\alpha}_3 = P_3 * \hat{LR}_i = 2,600 * 0,85 \approx 2,209$$

Using the estimated ultimate loss, (2.3) can be applied to determine the expected loss at a given moment, and (2.1) to determine the Loss Reserve for i. As can be seen in Examples 2.3 and 2.4, having adequate estimators for both the premium and development is key for the performance of this method.

2.3.3 | Recent Research for Loss Reserve prediction

The Chain Ladder as described in Section 2.3.1 is a deterministic method. It is able to make an estimation, but can give no indication of the error that it may produce. In the papers of Mack (1993, 1999), methods have been described to determine the distribution-free standard error produced by the CL method, as to give an indication for the uncertainty in the estimation.

In recent research, most focus has been on incorporating external data (i.e. information not captured in the triangle) in order to improve the accuracy of the predictions. For instance, a Double Chain Ladder method has been proposed, which does not only take observed aggregated losses into account, but also reported count data to infer both the IBNR and RBNS claims (Martinez-Miranda, Nielsen, & Verrall, 2012). This method results in a comparable best estimate for the loss reserve as the Chain Ladder, but a lower Root Mean Square Error (RMSE).

A different approach has been proposed by Kuang, Nielsen, and Perch Nielsen (2011), where the Chain Ladder has been extended with a calendar-year trend, which can be used in combination with non-stationary time-series forecasting. This method has an improved accuracy of predictions if the data has an unstable calendar year trend.

Zhang (2010) has described an implementation of a Multivariate Chain Ladder method, where correlation between triangles has been taken into account by applying an unrelated regression method. It is used to model intercepts of paid and incurred triangles or different lines of business, resulting in improved predictions.

Another implementation of correlation between triangle dependence is researched by Shi and Frees (2011), where a copula regression model with bootstrapping is implemented, resulting in both a Best Estimate and the uncertainty underlying the model prediction. Incorporating dependence along multiple insurers in an identical Line of Business has been subsequently researched (Shi, 2017). Both researches have given similar Best Estimates (BE) of the loss reserve, but can more accurately determine the uncertainty underlying this estimate.

There are also methods that are focused on the increased granularity of available data in an insurance company. For instance, models have been proposed to derive the loss reserves on an individual claims level, as opposed to the aggregate losses or claims that is currently used in actuarial practice (Antonio & Plat, 2014; Maciak, Okhrin, & Pešta, 2018). Most notably, Machine Learning in the form of Regression Trees has been applied by Wüthrich (2018), by learning policy and/or claims characteristics that have an influence on the loss incurred. While the Best Estimate of Wüthrich (2018) was comparable to current practice, the uncertainty of prediction is not yet calculated or implemented.

Another Machine Learning application on aggregated data has been researched, in the form of a Gaussian Process (GP) regression (Lally & Hartman, 2018). It enables capturing of trends in the data and learn those, solely based on the data presented in the run-off triangle. This model results in a comparable or better prediction of the BE of the Loss Reserve. However, the volatility and uncertainty in the model is still large.

As the Gaussian Process model by Lally and Hartman (2018) will be the main focus for this research, it will be described in more detail in Chapter 3.

3 A Gaussian process model to predict loss reserves

Contents

3.1 Introduction on the model **15**

3.2 Covariance functions **16**

3.2.1 Squared Exponential 17

3.2.2 Matérn 3/2 and Matérn 5/2 17

3.2.3 Signal and Noise 18

3.3 Input Warping **19**

3.4 Input data preparation **19**

3.5 Hyperparameters and prior distributions **21**

3.6 Model limitations **22**

In this chapter, the Gaussian process model to predict loss reserves is explained, as introduced in the paper by Lally and Hartman (2018). A brief introduction on the model will be given, after which the model design and hyperparameters are explained. Then, we will elaborate on the transformations of the input data to make the data suitable to be used in Gaussian process model. We will conclude by discussing the limitations of the current model.

3.1 Introduction on the model

A hierarchical Gaussian process model to predict loss reserves is introduced by Lally and Hartman (2018). The actuarial problem of predicting loss reserves is discussed in Section 2, and a brief introduction on Gaussian process regression is given in Section 1.3. The application of a GP regression on run-off triangles is new, while it is more common in e.g. geostatistics (where it is known as Kriging) (Gelfand & Schliep, 2016). In geostatistics, spatial or spatio-temporal data is analysed, usually on $\mathbb{R}^2$. Measurements (observations) are then used to predict missing points of interest. An example of an application in geostatistics is to examine the spatial variation of relative risk of a disease, given several observations at specific locations, or making an interpolation of radioactivity, given a limited number of observations (Diggle, Tawn, & Moyeed, 1998).

3.2 | Covariance functions

As described in Section 1.3, a Gaussian process regression model is mostly defined by its' covariance function, also referred to as the kernel function. A kernel function is valid if it results in a positive semidefinite (PSD) correlation matrix (Rasmussen & Williams, 2006). Three different kernel functions will be applied on the run-off triangle, effectively resulting in three different Gaussian process models. All kernel functions considered are both stationary and isotropic, of which definitions are given in Definitions 3.1 and 3.2.

Definition 3.1: Stationary kernel

We define x as the input for a one-dimensional Gaussian process, and x_p, x_q as two locations on x.
A kernel $k(x_p, x_q)$ is stationary if the kernel depends only on the separation $x_p - x_q$. That is: $k(x_p, x_q) = k(x_p - x_q)$

Sources: Barber (2012), with adaptations to match notation.

Definition 3.2: Isotropic covariance

A covariance function is isotropic if it is a function only of the distance $d = |x_p - x_q|$. Such covariance functions are, by construction, rotationally invariant.

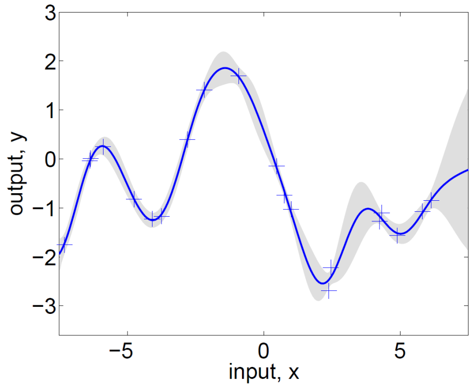
Source: Barber (2012), with adaptations to match notation.

As will be described in Section 3.4, the euclidean distance between two data points will be calculated and used as input, based on vectors x_1 (development lag) and x_2 (incurral year). This distance will be the input for the kernel functions. As such, the covariance functions considered are both stationary and isotropic kernel functions.

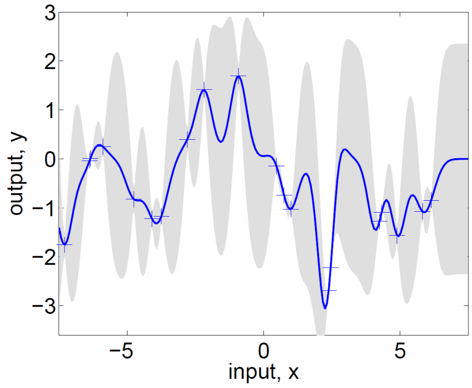
Moreover, a characteristic length-scale (ℓ) is defined to apply in our covariance functions. Loosely speaking, the length-scale defines how far one needs to move along a particular axis for the function values to become uncorrelated (Rasmussen & Williams, 2006). A visual one-dimensional example of this is given in Figure 3.1, where the length-scale is varied between $\ell = 1$, $\ell = 0.3$ and $\ell = 3$.

Lally and Hartman (2018) introduce a bandwidth parameter $\psi = \frac{1}{2\ell^2}$ to implement the characteristic length-scale in the GP model. Hence, larger values of ψ result in a shorter length-scale and vice versa. They define two separate bandwidth parameters for both x_1 and x_2, defined as respectively ψ_1 and ψ_2. As such, all covariance functions considered are defined as a function of r^2:

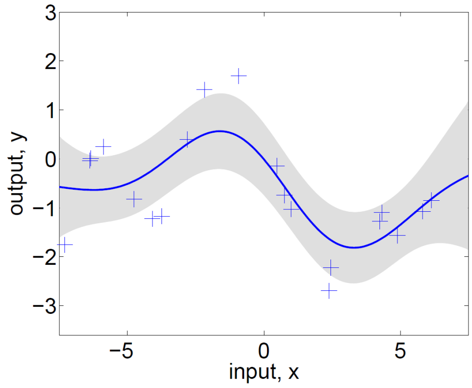
$$\begin{aligned} x_p &= (x_{1,p}, x_{2,p}) \\ x_q &= (x_{1,q}, x_{2,q}) \\ r^2_{p,q} &= \psi_1(x_{1,p} - x_{1,q})^2 + \psi_2(x_{2,p} - x_{2,q})^2 \end{aligned} \tag{3.1}$$



(a), $\ell = 1$



(b), $\ell = 0.3$



(c), $\ell = 3$

Figure 3.1: Three examples of a varying length-scale for a one-dimensional Gaussian process regression. Data marked with a "+" are observations from a GP with $\ell = 1$, the shaded grey area is the 95% CI for the underlying process $f(x)$. Source: Rasmussen and Williams (2006)

3.2.1 | Squared Exponential

The Squared Exponential (SE) covariance function, also known as the Radial Basis Function, the Exponentiated Quadratic function or the Gaussian Kernel, is defined in (3.2):

$$k(x_p, x_q) = \exp(-r_{p,q}^2) \tag{3.2}$$

The SE covariance function is infinitely differentiable, and thus generally results in smooth predictions (Rasmussen & Williams, 2006). As this might not be a realistic representation of the actual process, we also include two covariance functions that result in rougher predictions.

3.2.2 | Matérn 3/2 and Matérn 5/2

The Matérn-class of covariance functions results in rougher functions than the Squared Exponential function. They are parametrised by one value, and the values $\frac{3}{2}$ and $\frac{5}{2}$ are most commonly used for Machine Learning purposes (Rasmussen & Williams, 2006), where the Matérn $\frac{3}{2}$ function is rougher than the Matérn $\frac{5}{2}$ function. Matérn $\frac{3}{2}$ is given in (3.3) and Matérn $\frac{5}{2}$ in (3.4).

$$k(x_p, x_q) = (1 + \sqrt{3}r_{p,q})\exp(-\sqrt{3}r_{p,q}) \tag{3.3}$$

$$k(x_p, x_q) = (1 + \sqrt{5}r_{p,q} + \tfrac{5}{3}r_{p,q}^2)\exp(-\sqrt{5}r_{p,q}) \tag{3.4}$$

In order to visualise the roughness of these covariance functions, random draws from one-dimensional Gaussian processes with these covariance functions have been plotted in Figure 3.2.

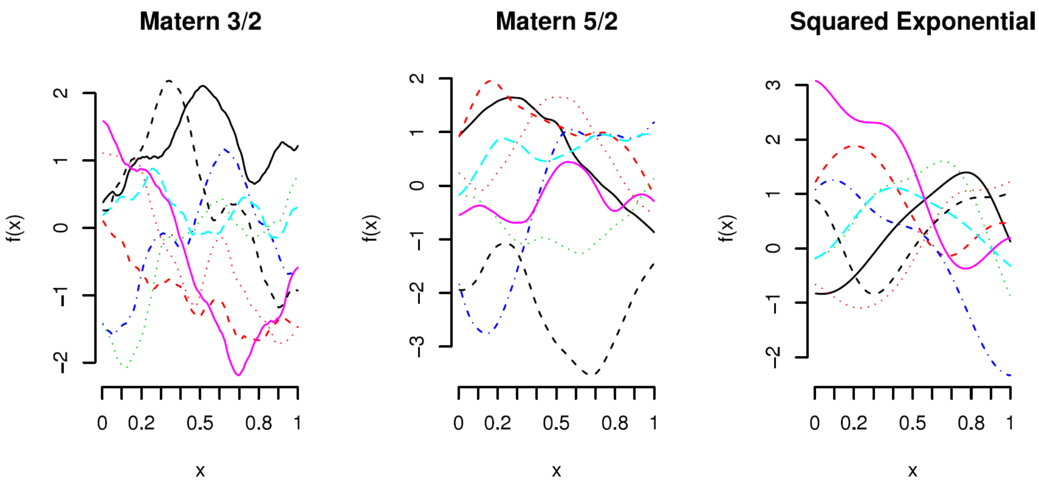


Figure 3.2: Random draws from one-dimensional Gaussian processes with different kernel functions, with $\psi = 10$. Source: Lally and Hartman (2018)

3.2.3 | Signal and Noise

The covariance functions that we will consider are in normalised form. Hence, if $r_{p,q}^2 = 0$ (i.e. the distance between two points is zero), then $k(x, x') = 1$ for all functions. This assumes that the underlying process has a variance of 1. In order to relax this assumption, Lally and Hartman (2018) introduce a parameter η^2 for all covariance functions, by which the covariance functions will be multiplied.

Moreover, observations might be noisy, as such that they are not entirely accurate. For actual observed losses, it can be assumed that the data quality is sufficient enough to be noise-free. However, if estimators would be introduced to the model (for example, in the case of Bornheutter-Ferguson), these estimations can contain noise, as their true value is unknown. Henceforth, we will add a parameter $\sigma^2\delta_{pq}$ to the covariance functions, where δ_{pq} is the Kronecker delta (which equals 1 if $p = q$, and 0 otherwise) (Lally & Hartman, 2018).

As such, our Squared Exponential, Matérn $\frac{3}{2}$ and Matérn $\frac{5}{2}$ functions are modified and the resulting functions are given in (3.5), where the subscript of r_{pq} is suppressed for notation purposes. The functions as given in (3.5) are the final form of the covariance functions that will be considered in the research.

$$\begin{aligned}
k_{SE}(x_p, x_q) &= \eta^2 \exp(-r^2) + \sigma^2\delta_{pq} \\
k_{M\frac{3}{2}}(x_p, x_q) &= \eta^2(1 + \sqrt{3}r)\exp(-\sqrt{3}r) + \sigma^2\delta_{pq} \\
k_{M\frac{5}{2}}(x_p, x_q) &= \eta^2(1 + \sqrt{5}r + \tfrac{5}{3}r^2)\exp(-\sqrt{5}r) + \sigma^2\delta_{pq}
\end{aligned} \tag{3.5}$$

3.3 Input Warping

Input Warping is a methodology that can be applied in order improve results when a non-stationary process is modelled with a stationary covariance function (as defined in Definition 3.1) (Snoek, Swersky, S. Zemel, & P. Adams, 2014). Commonly, processes that we wish to model are non-stationary.

Lally and Hartman (2018) have established that the application of a Gaussian process on the prediction of loss reserves require such a correction, and have applied Input Warping in their model. Input Warping is introduced in the paper of Snoek et al. (2014). It translates the non-stationary process to a stationary one, by changing the covariance function from $k(x_p, x_q)$ to $k(w(x_p), w(x_q))$. The function $w(x)$ is the warping function that modifies the input vectors to one that is stationary (Snoek et al., 2014).

In theory, a multitude of warping functions can be applied. However, Snoek et al. (2014) recommend to use the Beta cumulative distribution function (CDF), as the Beta-function can take on a multitude of forms (e.g. linear, exponential, logarithmic, sigmoidal) by varying the two parameters α and β. As the Beta-distribution can only take values on $[0, 1]$, the input vectors of our data is normalised to this interval, as described in Section 3.4.

For our model, we will warp the input vectors x_1 and x_2 with their own warping function. Our covariance functions thus change accordingly. We therefore redefine the $r_{p,q}$ term that we have defined in (3.1) to be warped:

$$\begin{aligned} w_1(x) &= \text{BetaCDF}(x, \alpha_1, \beta_1) \\ w_2(x) &= \text{BetaCDF}(x, \alpha_2, \beta_2) \\ r_{p,q}^2 &= \psi_1(w_1(x_{1,p}) - w_1(x_{1,q}))^2 + \psi_2(w_2(x_{2,p}) - w_2(x_{2,q}))^2 \end{aligned} \tag{3.6}$$

Where BetaCDF is the Beta Cumulative Distribution Function. From (3.6), we can see that an unique warping function is defined for each axis, with their own α and β parameters.

3.4 Input data preparation

The Gaussian Process model by Lally and Hartman (2018) defines two input dimensions for every data point that is in the data set: the development lag (x_1) and incurral year (x_2). Using these two input parameters, we can determine the euclidean distance between two data points. This distance is then used as parameter for the covariance function.

Example 3.1: Indexing every data point in the triangle

Consider the triangle presented in Example 2.2:

Incurral Year (i)	Development lag (j) 1	2	3
2010 (1)	1,000	1,500	1,750
2011 (2)	1,250	1,700	$\circ_2$
2012 (3)	1,400	$\circ_1$	$\circ_3$

For every Z in the triangle, x_1 (DL) and x_2 (IY) is defined, with identical indices:

	1	2	3	4	5	6	7	8	9
Z	1,000	1,250	1,400	1,500	1,700	1,750	$\circ_1$	$\circ_2$	$\circ_3$
x_1	1	1	1	2	2	3	2	3	3
x_2	1	2	3	1	2	1	3	2	3

The input vectors x_1 and x_2 will be normalised on the interval $[0, 1]$ in order to be able to apply Input Warping, as described in Section 3.3. Normalisation is performed by applying (3.7) on both vectors x_1 and x_2, to transform them to the normalised vectors x'_1 and x'_2:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{3.7}$$

Example 3.2: Normalising the input vectors

Consider the data as given in Example 3.1:

	1	2	3	4	5	6	7	8	9
Z	1,000	1,250	1,400	1,500	1,700	1,750	$\circ_1$	$\circ_2$	$\circ_3$
x_1	1	1	1	2	2	3	2	3	3
x_2	1	2	3	1	2	1	3	2	3

(3.7) is now applied on the vectors x_1 and x_2, to transform them to the interval $[0, 1]$:

	1	2	3	4	5	6	7	8	9
Z	1,000	1,250	1,400	1,500	1,700	1,750	$\circ_1$	$\circ_2$	$\circ_3$
x'_1	0	0	0	0.5	0.5	1	0.5	1	1
x'_2	0	0.5	1	0	0.5	0	1	0.5	1

Furthermore, the observations Z will be standardised such that it has a mean of zero, and a standard deviation of 1. These transformations improve the predictions that can be made by the Gaussian process model, in such a way that one configuration of the model suffices regardless of which run-off triangle we use as input. This also allows us to set the mean of the Gaussian process model to zero. Standardisation is performed by applying (3.8) to our observations Z_i, with $\bar{Z}$ as the mean of Z, and s as the standard deviation of Z.

$$Z'_i = \frac{Z_i - \bar{Z}}{s} \tag{3.8}$$

Example 3.3: Standardising the observations

Considering the data from Example 3.2, the *observations* of Z are to be standardized:

	1	2	3	4	5	6	7	8	9
Z	1,000	1,250	1,400	1,500	1,700	1,750	$\circ_1$	$\circ_2$	$\circ_3$
x'_1	0	0	0	0.5	0.5	1	0.5	1	1
x'_2	0	0.5	1	0	0.5	0	1	0.5	1

The mean of Z is $1{,}433\frac{1}{3}$ and the standard deviation is approximately 282.25. Z can now be standardised by applying (3.8), resulting in the following data:

	1	2	3	4	5	6	7	8	9
Z'	-1.535	-0.650	-0.118	0.236	0.945	1.122	$\circ_1$	$\circ_2$	$\circ_3$
x'_1	0	0	0	0.5	0.5	1	0.5	1	1
x'_2	0	0.5	1	0	0.5	0	1	0.5	1

The resulting vector Z' now has a mean of zero, and a SD of 1.

Throughout this thesis, an identical setup is followed as applied by Lally and Hartman (2018). The data transformations and programming is performed using the R Project (R Core Team, 2018). We sample from the posterior distribution by applying Markov Chain Monte Carlo (MCMC) sampling, implemented using the STAN package, which is available as a plugin for R (Carpenter et al., 2017).

MCMC allows us to sample directly from the posterior distribution. The No-U-Turn Sampler (NUTS) is applied (Hoffman & Gelman, 2014), and we setup 4 chains of 2,000 iterations - of which the first 1,000 iterations are regarded as a warm-up period and are thus discarded. The remaining 1,000 iterations (of 4 chains, so 4,000 in total) are samples from our Bayesian posterior distribution. We derive the mean of this posterior distribution as our Best Estimate, and can determine the uncertainty in this prediction directly from the samples of the posterior distribution.

3.5 Hyperparameters and prior distributions

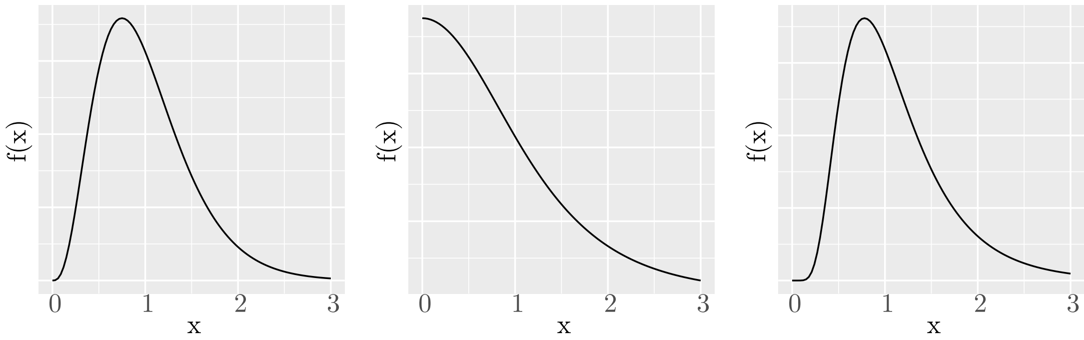
All hyperparameters that will be applied in the model have been discussed. In order to adequately make inferences of each hyperparameter, a prior distribution for every hyperparameter is defined from which we can sample using MCMC. In the model of Lally and Hartman (2018), weakly informative priors have been defined for every hyperparameter. The hyperparameters and their prior distributions are summarised in Table 3.1.

The prior gamma$(4,4)$ is the Gamma distribution and is parametrised with a shape and rate value of 4. It approximately has a mean of 1 and a variance of 0.25 (Lally & Hartman, 2018). The $\mathcal{T}^+(4,0,1)$ is the Student's T-distribution, with four degrees of freedom, a mean of zero and a variance of 1. Finally, $\ln\mathcal{N}(0,0.5)$ is the lognormal distribution, with a mean of zero and a standard deviation of 0.5.

All prior distributions are constrained to be ≥ 0. These distributions are flexible enough to take on extreme values if the data suggests to do so. We have included plots of the probability density functions (PDF) of all prior distributions in Figure 3.3 on the interval [0,3].

Table 3.1: Overview of hyperparameters and prior distributions in the Gaussian process model

Parameter	Description	Prior Distribution
ψ_1	Bandwidth-parameter $\frac{1}{2\ell^2}$ of the DL-axis	$\text{gamma}(4, 4)$
ψ_2	Bandwidth-parameter $\frac{1}{2\ell^2}$ of the IY-axis	$\text{gamma}(4, 4)$
η^2	Signal parameter of the entire Gaussian process	$\mathcal{T}^+(4, 0, 1)$
σ^2	Noise parameter of the entire Gaussian process	$\mathcal{T}^+(4, 0, 1)$
α_1	Alpha-parameter of the warping function of the DL-axis	$\ln \mathcal{N}(0, 0.5)$
β_1	Beta-parameter of the warping function of the DL-axis	$\ln \mathcal{N}(0, 0.5)$
α_2	Alpha-parameter of the warping function of the IY-axis	$\ln \mathcal{N}(0, 0.5)$
β_2	Beta-parameter of the warping function of the IY-axis	$\ln \mathcal{N}(0, 0.5)$



(a) $\text{gamma}(4, 4)$ (b) $\mathcal{T}^+(4, 0, 1)$ (c) $\ln \mathcal{N}(0, 0.5)$

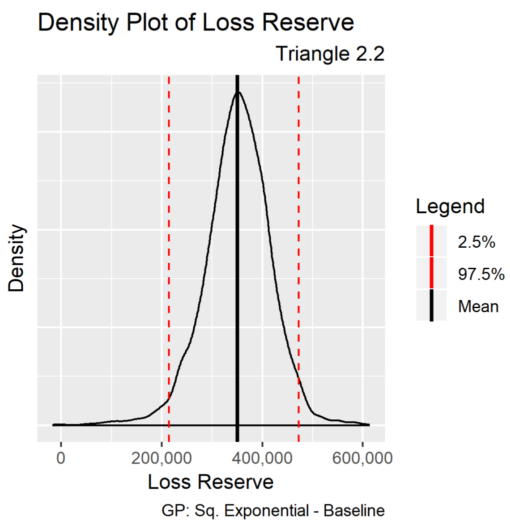
Figure 3.3: Plots of the PDF of prior distributions on the interval [0,3]

3.6 | Model limitations

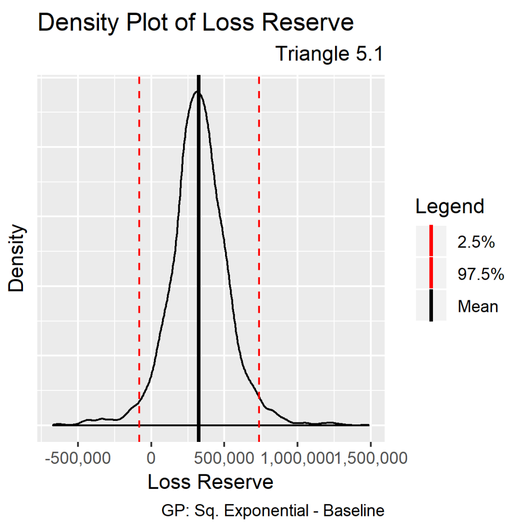
The model presented in this section generally performs well in predicting a best estimate of the loss reserve, which is the mean of the posterior distribution. The results are similar to other methods currently used in actuarial practice and recent research, but the density of the posterior distribution is wider than can be expected of the actual uncertainty. As such, it is not currently applicable into actuarial practice, and might require changes to adequately reflect this risk.

In order to give an indication of the uncertainty of the model results, two density plots of the model by Lally and Hartman (2018) are given in Figure 3.4. While the prediction of Triangle 2.2 (Figure 3.4a) is wide (95% Confidence Interval: $[215k, 473k]$), its' Best Estimate is adequate (351k, actual observed losses 354k).

However, the model is not able to capture all trends accordingly, as becomes obvious for Triangle 5.1. The density plot of Triangle 5.1 (Figure 3.4b) gives no conclusive results whatsoever (95% CI: $[-80k, 736k]$), and the Best Estimate that the model produces for Triangle 5.1 is 325k, while the actual observed losses are 46k. As such, the performance of the model has room for improvement. How we intend to improve the model is explained in Chapter 4.



(a) Density plot of Triangle 2.2, Sq. Exponential



(b) Density plot of Triangle 5.1, Sq. Exponential

Figure 3.4: Density plots of predicted Loss Reserves, by the Gaussian process model of Lally and Hartman (2018).

4 | Research setup

Contents

4.1 Methodology . . . 26

4.2 Prior distribution of hyperparameters . . . 26

4.2.1 Bandwidth parameter . . . 26

4.2.2 Signal and noise parameters . . . 27

4.2.3 Warping parameters . . . 27

4.3 Extending the model with premium information . . . 29

4.3.1 Transform observations to Loss Ratio's . . . 29

4.3.2 Supply Premiums as input instead of Accident Years . . . 29

4.3.3 Add Estimators of the Ultimate Loss to the model . . . 30

This chapter will cover the research that we will perform on the Gaussian process model as described in Chapter 3, based on the limitations identified and the research questions given in Section 1.4. We will perform several analyses on a larger data set than the paper of Lally and Hartman (2018) in order to validate the accuracy and applicability of the model.

Furthermore, we wish to extend the model by supplying it with external information in order to improve our predictions. Our data set has external data in the form of received premiums in an incurral year. As such, we want to investigate if the model can be adapted in such a way to include this data to improve the best estimate prediction of the loss reserve. Investigating the appropriate method to implement this is important, because simply assuming relatedless of data and learning them together can be detrimental for performance of a Machine Learning model (Bonilla, Chai, & Williams, 2008).

We recall our research questions:

Can we improve the Hierarchical Gaussian Process model
to predict the loss reserve of a non-life insurer?

With the following sub-questions:

- In order to validate the design choices:
  - Is the model applicable on a more extensive data set?
  - Are the prior distributions on the hyperparameters adequately chosen?
- In order to improve the Best Estimate and/or reduce the confidence interval:
  - Can relevant, out-of-triangle information be supplied to the model?
  - Can the GP model be extended with a Bornheutter-Ferguson estimation method?

4.1 | Methodology

We will elaborate on the data set that we will apply in Chapter 5. Performance of the models will be analysed based on both the Best Estimate of the Loss Reserve, and the Root Mean Square Error of prediction (RMSE) of this Best Estimate - as defined in Equation 4.1 - where Z_t is a data point that we wish to estimate, and T is the total number of data points to estimate.

Given that our data set consists of both the upper and lower triangle, we have a natural split in a training and testing set for validation. Furthermore, we can determine the 95% Confidence Interval of the posterior distribution by analysing the MCMC samples, and the standard deviation of the 4,000 samples. For the Chain Ladder method, we can determine the Standard Error of the prediction as described in the paper of Mack (1993).

$$RMSE = \frac{\sum_{t=1}^{T}(\hat{Z}_t - Z_t)^2}{T} \tag{4.1}$$

Before conducting our research, we will calculate baseline measurements consisting of both the Chain Ladder predictions and the Gaussian process model as given in the paper of Lally and Hartman (2018) to investigate the performance of the model on a larger data set. This will give an answer to our first sub-question: *"Is the model applicable on a more extensive data set?"*

4.2 | Prior distribution of hyperparameters

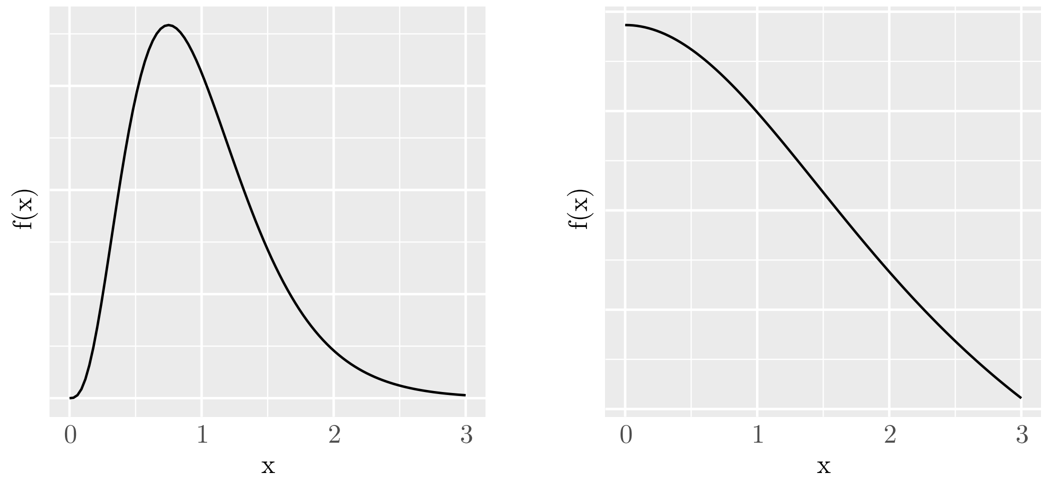
With this analysis, we aim to deduct a new combination of prior parameters and measure it against the baseline measurements performed earlier. In these researches, we will only vary one parameter, while keeping the others identical to the baseline GP model. As such, we can strictly identify the influence of the prior distribution on this specific parameter. Using these analyses, we can determine an 'optimal' prior distribution for every triangle of each hyperparameter, in order to draw a conclusion of the influence of these choices. These analyses will be used to answer our second sub-question: *"Are the prior distributions on the hyperparameters adequately chosen?"*

4.2.1 | Bandwidth parameter

Flaxman et al. (2016) have described a fast inference method for hierarchical Gaussian processes. In this manuscript, they argue to use a Cauchy-distribution for the "inverse length-scale" (i.e. the bandwidth as defined earlier), with a location of 0 and a scale of 2.5 and constrained to be positive. This argument is also supported by a different research focused on using the Cauchy-distribution, which comes to this conclusion based on a frequentist-risk analysis (Polson & Scott, 2012). This prior distribution is notably different than the gamma distribution used in the model by Lally and Hartman (2018), and we thus investigate its' performance on run-off triangles. We summarise in Table 4.1 and Figure 4.1.

Table 4.1: Overview of researched prior distributions for the bandwidth parameters

Parameter	Prior Distribution (Original)	Prior Distribution (Researched)
ψ_1	$\text{gamma}(4, 4)$	$\text{Cauchy}(0, 2.5)$
ψ_2	$\text{gamma}(4, 4)$	$\text{Cauchy}(0, 2.5)$



(a) gamma$(4, 4)$

(b) Cauchy$(0, 2.5)$

Figure 4.1: Plots of the PDF of prior distributions of the bandwidth parameters, on the interval [0,3]

4.2.2 | Signal and noise parameters

In the manuscript of Flaxman et al. (2016), a $\ln\mathcal{N}(0, 1)$ distribution is applied for the signal and noise parameters, mentioning that these parameters should resemble the scale of the data. As we standardise our input to have a mean of zero, and a standard deviation of one, the $\ln\mathcal{N}(0, 1)$ could also be an appropriate prior for our data; given that we also constrain this prior to be positive. This is summarised in Table 4.2 and Figure 4.2.

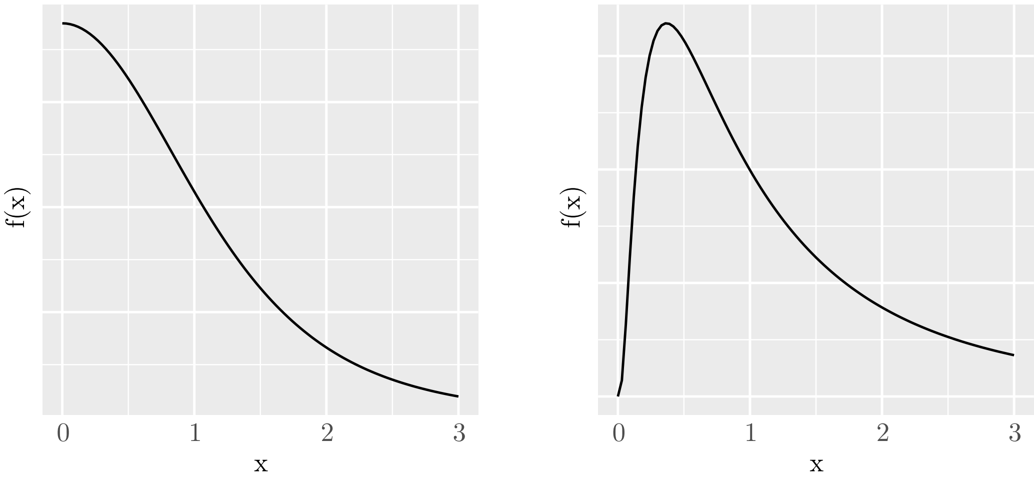
Table 4.2: Overview of researched prior distributions for signal and noise parameters

Parameter	Prior Distribution (Original)	Prior Distribution (Researched)
η^2	$\mathcal{T}^+(4, 0, 1)$	$\ln\mathcal{N}(0, 1)$
σ^2	$\mathcal{T}^+(4, 0, 1)$	$\ln\mathcal{N}(0, 1)$

4.2.3 | Warping parameters

In the paper of Snoek et al. (2014), where Input Warping is first described, several prior distributions for α and β are given in order to get several functional forms of the Beta-distribution. For our application, Lally and Hartman (2018) conclude that an Exponential warping function performs best. An exponential warping function is recommended by Snoek et al. (2014) to be modelled by different priors than the priors applied by Lally and Hartman (2018). As such, we investigate the influence of this choice on model performance.

The parameters and their distributions are given in Table 4.3, and their PDF's are plotted in Figure 4.3.

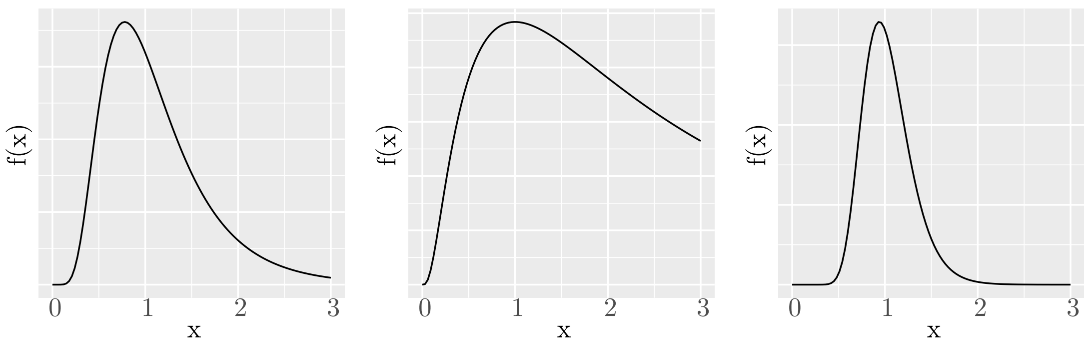


(a) $\mathcal{T}^{+}(4, 0, 1)$ (b) $\ln\mathcal{N}(0, 1)$

Figure 4.2: Plots of the PDF of prior distributions of the signal and noise parameters, on the interval [0,3]

Table 4.3: Overview of researched prior distributions for Input Warping

Parameter	Prior Distribution (*Lally*)	Prior Distribution (*Snoek*)
α_1	$\ln\mathcal{N}(0, 0.5)$	$\ln\mathcal{N}(1, 1)$
β_1	$\ln\mathcal{N}(0, 0.5)$	$\ln\mathcal{N}(0, 0.25)$
α_2	$\ln\mathcal{N}(0, 0.5)$	$\ln\mathcal{N}(1, 1)$
β_2	$\ln\mathcal{N}(0, 0.5)$	$\ln\mathcal{N}(0, 0.25)$



(a) $\ln\mathcal{N}(0, 0.5)$ (b) $\ln\mathcal{N}(1, 1)$ (c) $\ln\mathcal{N}(0, 0.25)$

Figure 4.3: Plots of the PDF of prior distributions of the warping parameters on the interval [0,3]

4.3 | Extending the model with premium information

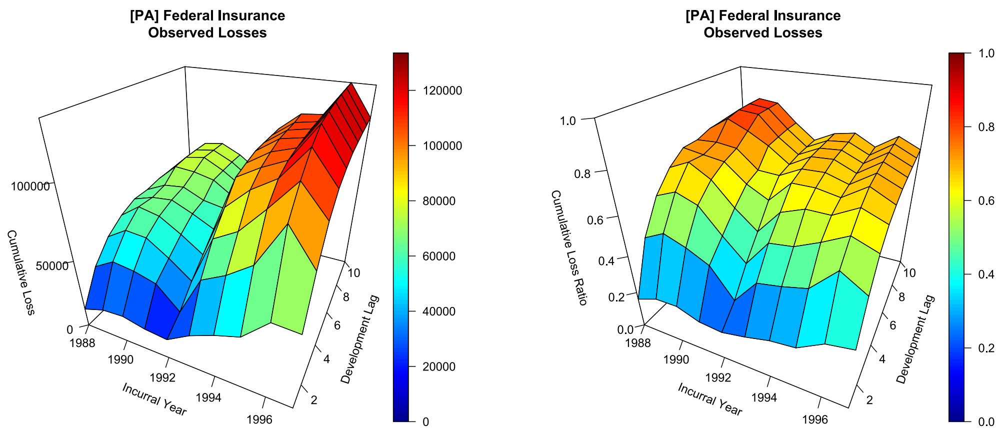
Parallel to an analysis of prior distributions, we will try several configurations to include the received premium. As we intend to determine the feasibility of such a method, we will restrict these setups to the model with the prior distributions as initially described by Lally and Hartman (2018) for equal comparison. We will, however, apply all three covariance functions on the model. By doing this, we wish to answer the following two sub-questions:

In order to improve the Best Estimate and/or reduce the confidence interval:

- *Can relevant, out-of-triangle information be supplied to the model?*
- *Can the GP model be extended with a Bornheutter-Ferguson estimation method?*

4.3.1 | Transform observations to Loss Ratio's

One of the methods to implicitly give information on the premiums is by pre-transforming the input data. We do so by dividing the observed losses at a specific moment in time ($Z_{i,j}$) by the net premiums received in that year (P_i). Using this standardisation, the Development-lag run-off will not materially change, as they are all divided by an identical constant (for example: the Chain Ladder factors will remain identical). However, over the incurral-year axis, different trends might become visible. We wonder if the model's behaviour can cope with these changes, and if they could potentially improve model predictions. We visualise the transformation in Figure 4.4, where we give an example of observed losses and observed loss ratio's.



(a) Observed cumulative losses (Triangle 3.2) (b) Observed cumulative LR (Triangle 3.2)

Figure 4.4: Plots of observed cumulative losses and loss ratio of Triangle 3.2

We transform the predictions by the GP model back by multiplying it with the same constants, and then determine the Loss Reserve and error metrics. We will compare this with the Chain Ladder and the Hierarchical GP model by Lally and Hartman (2018) without adaptations that might arise from the prior distributions analysis.

4.3.2 | Supply Premiums as input instead of Accident Years

In the data set we consider, both the known observations and the net premiums received in that year are available. As we will present in Section 5, both premiums and ultimate losses in that year fluctuate heavily in some run-off triangles. As such, we want to investigate if using net premiums, as opposed to the incurral year, could be a more meaningful input parameter.

As premiums might fluctuate throughout the years, the model could find more support for relatively distant observations due to a reshuffling of data points, reducing the standard error of the model and potentially narrowing the confidence interval. An example of such a revised spreadsheet is given in Table 4.4.

Table 4.4: Triangle 5.3: Upper Triangle. Sorted on premium received.

Premium	1	2	3	4	5	6	7	8	9	10
177k	22k	61k	85k	100k	109k	115k	119k	122k	123k	125k
201k	27k	78k	106k	122k	133k	139k	143k	146k	147k	·
245k	25k	·	·	·	·	·	·	·	·	·
246k	33k	100k	135k	158k	169k	178k	183k	188k	·	·
286k	39k	114k	157k	181k	197k	209k	213k	·	·	·
287k	29k	66k	·	·	·	·	·	·	·	·
338k	36k	83k	111k	·	·	·	·	·	·	·
340k	42k	126k	165k	189k	204k	214k	·	·	·	·
366k	41k	100k	132k	151k	· ·	·	·	·	·	·
419k	46k	117k	155k	179k	194k	·	·	·	·	·

As can be seen in Table 4.4, due to the reshuffling of data, the prediction of losses or a general trend might be more difficult in the Incurral-year axis. As such, we do not expect this method to give sufficient results on every triangle that we analyse. For triangles that have few shocks, however, this method could potentially give a more accurate result. We will, nonetheless, analyse on all triangles.

4.3.3 | Add Estimators of the Ultimate Loss to the model

The Bornheutter-Ferguson method is reliant on estimated ultimate losses for each incurral year. As our model has a significant uncertainty in the most distant predictions, we attempt to mitigate this by adding estimated observations as extra input to our model based on the observed premium.

We apply somewhat identical logic to our estimations as the BF-method described in 2.3.2. Of a triangle, we will determine the Estimated Loss Ratio of every observed point in the triangle, based on the development parameters as described in Section 2.3.2. Our estimation method is given in (4.2) and is built up out of two parts: the development to date, and the development that still has to occur. The development to date is modeled by multiplying the most recent Estimated Loss Ratio with its' cumulative development parameter ($\hat{\gamma}_l$), effectively returning the most recent observed Loss Ratio. For the development that still has to occur, we will apply Gaussian Noise. We generate this noise by defining a new random variable X, which has a normal distribution. It has a mean equal to the mean of the estimated Loss Ratio's in the triangle, and the standard deviation equal to the standard deviation of these estimations. Therefore: if a triangle has few shocks, its' standard deviation will be smaller and estimators will thus be less volatile than when a triangle has more shocks.

$$X \sim \mathcal{N}(\bar{LR}, \mathrm{SD}(LR))$$
$$l = n + 1 - i \quad (4.2)$$
$$\hat{\alpha}_i = ((\hat{\gamma}_l * LR_{i,l}) + (1 - \hat{\gamma}_l) * X) * P_i$$

Example 4.1: Estimators based on Bornheutter-Ferguson

We recall the estimated Loss Ratio's as calculated in Example 2.4:

Incurral Year (i)	Development year (j) 1	2	3	P_i
2010 (1)	0.81	0.85	0.85	2,060
2011 (2)	0.87	0.83	·	2,400
2012 (3)	0.89	·	·	2,600
Cum. Dev. Parameters ($\hat{\gamma}_j$)	0.60	0.85	1	

We analyse the six estimated Loss Ratio's. Of these, the mean is approx. 0.85 and the standard deviation is approx. 0.03. We want to make a estimation of the ultimate losses of $i = 2$ and $i = 3$. We do this by applying Equation 4.2:

$$X \sim \mathcal{N}(0.85, 0.03)$$
$$\hat{\alpha}_2 = (0.85 * 0.83 + (1 - 0.85) * X) * 2,400$$
$$\hat{\alpha}_3 = (0.60 * 0.89 + (1 - 0.60) * X) * 2,600$$

Given that a X is a random variable, we can not give a deterministic outcome in this example.

The estimations are supplied as extra observations of the model at $j = n$, but we model this with a different noise-parameter σ^2_{est}, which will have an identical prior distribution as σ^2. All other parameters will be kept equal. We will give an example of renewed input in Table 4.5, where the observations marked in bold are the estimators applied - and thus modeled with their own noise-parameter.

Table 4.5: Upper Triangle of 5.2, with added BF-estimators

IY	Development Lag 1	2	3	4	5	6	7	8	9	10
1988	19,016	44,632	59,804	66,052	70,115	72,219	73,565	74,273	75,112	75,655
1989	17,346	42,058	59,686	64,821	67,313	69,036	69,942	70,428	70,846	**71,192**
1990	12,212	28,087	42,719	46,564	48,016	49,030	49,700	49,994	·	**50,817**
1991	9,490	19,697	32,062	38,698	40,369	41,220	41,970	·	·	**42,739**
1992	7,605	14,874	21,105	29,016	35,208	36,884	·	·	·	**39,511**
1993	5,596	11,527	14,677	17,073	20,813	·	·	·	·	**23,007**
1994	4,885	10,118	13,103	14,570	·	·	·	·	·	**18,781**
1995	4,056	8,435	10,439	·	·	·	·	·	·	**17,490**
1996	4,213	8,768	·	·	·	·	·	·	·	**17,604**
1997	5,258	·	·	·	·	·	·	·	·	**25,048**

The noise-parameter allows the GP to take the uncertainty of the estimators into account when making predictions. Because there are now extra observations, the Gaussian process model now has more observations to make predictions with, and it can now perform an interpolation as opposed to only extrapolating, and as such we expect the confidence interval to take on more realistic values.

5 | Data set used

Contents

5.1 Medical Malpractice (MM) 33
5.1.1 Physicians Reciprocal Insurers (Triangle 1.1) 34
5.1.2 Promutual Group (Triangle 1.2) 34
5.1.3 Scpie Indemnity Company (Triangle 1.3) 36
5.2 Commercial Automobile (CA) 36
5.2.1 Farmers' Automobile (Triangle 2.1) 36
5.2.2 State Farm (Triangle 2.2) 38
5.2.3 The Ins Co (Triangle 2.3) 38
5.3 Personal Automobile (PA) 40
5.3.1 Farmers' Automobile (Triangle 3.1) 40
5.3.2 Federal Insurance Company (Triangle 3.2) 40
5.3.3 State Farm (Triangle 3.3) 42
5.4 Product Liability (PL) 42
5.4.1 Allstate Insurance Company (Triangle 4.1) 42
5.4.2 Federal Insurance Company (Triangle 4.2) 44
5.4.3 Federated Mutual Group (Triangle 4.3) 44
5.5 Workers' Compensation (WC) 46
5.5.1 Allstate Insurance Company (Triangle 5.1) 46
5.5.2 Lumbermen's Underwriting Alliance (Triangle 5.2) 46
5.5.3 State Farm (Triangle 5.3) 48

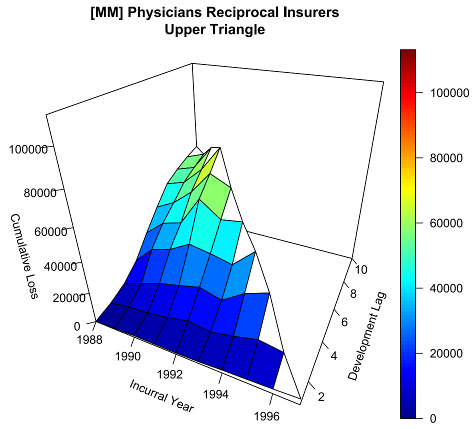
In order to test and improve our model, we have selected a total of 15 triangles from different Lines of Business. 3D-plots of both the upper triangle (which is supplied to the model as input) and the actual observed losses are presented (which the model attempts to estimate), along with a summary table containing the yearly premiums, ultimate losses and the corresponding loss ratio. We will highlight the most important challenges these data sets pose for our model. The actual data underlying these plots are given in Appendix C. Furthermore, all LoB's will be briefly addressed.

5.1 | Medical Malpractice (MM)

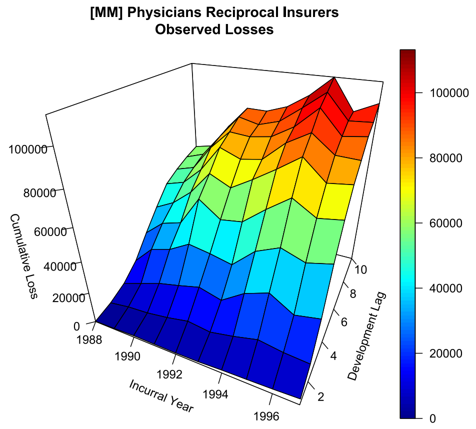
Medical Malpractice policies protect a customer (healthcare professional) against claims from a patient. These products usually have a relatively long run-off period, as these procedures (such as lawsuits) can have a considerable duration. Furthermore, payments can be made over a longer period of time.

5.1.1 | Physicians Reciprocal Insurers (Triangle 1.1)

In Figure 5.1, both the upper triangle and observed losses are presented for this insurer. Payments are made gradually along the development lag-axis, and we can also observe a growth along the incurral year-axis. The triangle is relatively stable, with incurral year 1995 having a larger claims volume than surrounding observations. We expect the models to be able to capture these properties adequately.



(a) Upper Triangle



(b) Observed Losses

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	59k	61k	77k	90k	90k	94k	102k	113k	96k	102k
Premium	73k	76k	86k	87k	92k	93k	97k	107k	107k	111k
Loss Ratio	0.80	0.81	0.90	1.04	0.98	1.01	1.05	1.06	0.89	0.92

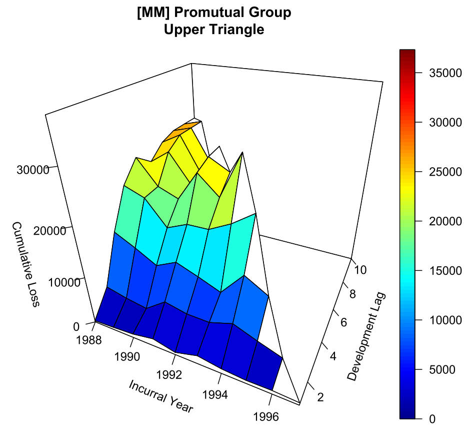
(c) Overview of losses and premiums

Figure 5.1: Data set of Physicians Reciprocal Insurers (Medical Malpractice)

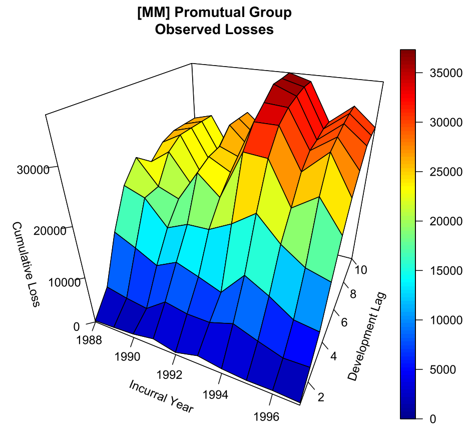
5.1.2 | Promutual Group (Triangle 1.2)

The Promutual Group details are given in Figure 5.2. Comparing this to the Triangle 1.1, we see a more rough pattern emerging, while the payments along the development lag-axis appear to stabilise after $DL = 6$.

We also see a smaller volume of both claims and premiums in comparison with Triangle 1.1. When looking at the underlying data in Table C.2 (Appendix C), we see that the ultimate claims level is fairly stable, while the premium volume has declined. As such, we expect that models that take Loss Ratio's and/or Premiums into account will have considerable difficulty with this triangle.



(a) Upper Triangle



(b) Observed Losses

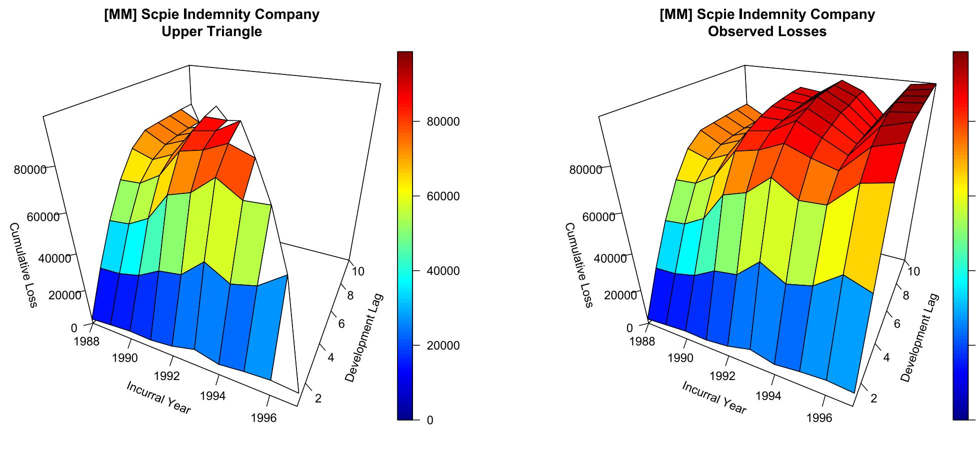
	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	26k	27k	20k	29k	26k	37k	36k	29k	33k	29k
Premium	53k	61k	52k	44k	29k	26k	25k	26k	27k	23k
Loss Ratio	0.49	0.44	0.39	0.65	0.87	1.41	1.43	1.10	1.20	1.29

(c) Overview of losses and premiums

Figure 5.2: Data set of Promutual Group (Medical Malpractice)

5.1.3 | Scpie Indemnity Company (Triangle 1.3)

The plots of Scpie Indemnity Company are displayed in Figure 5.3. Here, we see that the payments along the development lag-axis appear to take on substantial size between $DL = 1$ and $DL = 3$, after which they stabilise after $DL = 5$. Premium volumes are declining from 1988-1991, after which they gradually grow. Loss Ratio's from the early incurral years are considerably lower than that of the later incurral years. As such, we also expect that the LR/Premium model might underperform. We do expect that the normal Gaussian Process models are able to reflect the general trend observed, as it has already been applied on this triangle (Lally & Hartman, 2018).



(a) Upper Triangle (b) Observed Losses

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	78k	72k	75k	90k	89k	96k	92k	82k	97k	99k
Premium	135k	112k	99k	96k	99k	99k	97k	102k	102k	108k
Loss Ratio	0.57	0.64	0.76	0.94	0.90	0.97	0.95	0.80	0.96	0.91

(c) Overview of losses and premiums

Figure 5.3: Data set of Scpie Indemnity Company (Medical Malpractice)

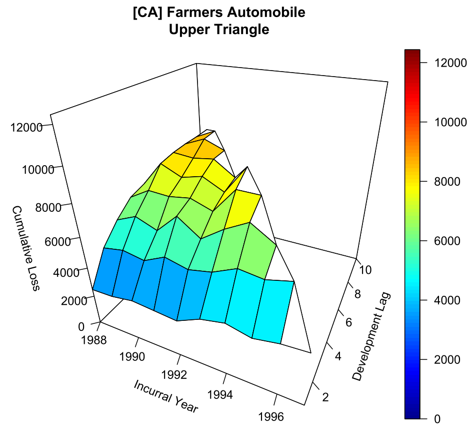
5.2 | Commercial Automobile (CA)

Automobile insurance products, both Commercial and Personal, in general have shorter tails and thus a quicker development than other insurance products (Shi, 2017).

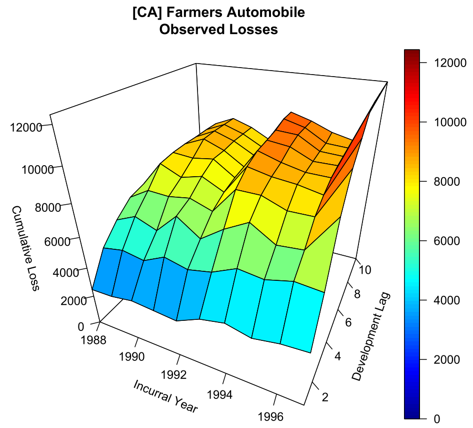
5.2.1 | Farmers' Automobile (Triangle 2.1)

The Farmers' Automobile triangle is plotted in Figure 5.4a and the observed losses are plotted in Figure 5.4b. We see that the pattern along the development-lag is relatively stable and flattens out at approximately $DL = 5$. Along the Incurral Year-axis, we see that it is fairly rough, both in the upper triangle and the observed losses. We expect that Matèrn-covariance functions are thus better at capturing this.

Some errors are to be expected, as there is an outlier in Incurral Year 1997 that is not captured by the first observation of IY 1997, nor by an increase in the net premium received.



(a) Upper Triangle



(b) Observed Losses

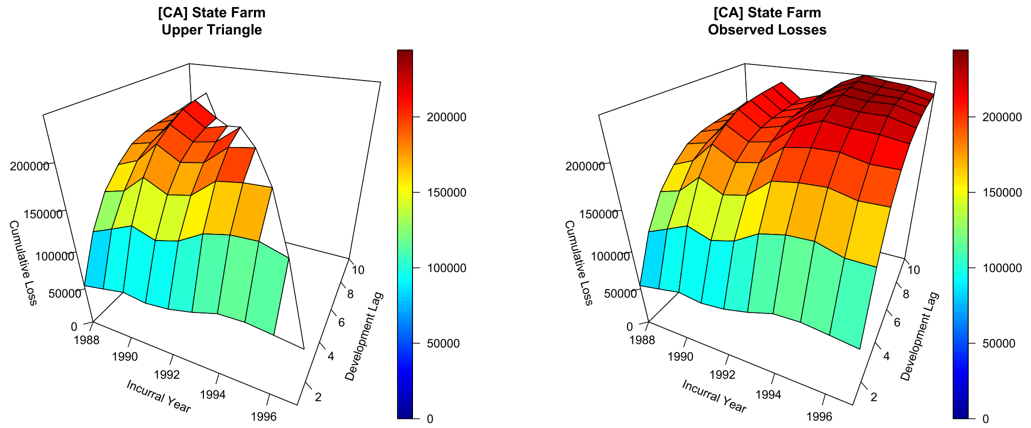
	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	7,115	8,425	8,991	8,549	7,853	9,980	9,583	9,038	8,719	12k
Premium	9,147	9,854	11k	11k	12k	13k	13k	14k	14k	14k
Loss Ratio	0.78	0.85	0.81	0.75	0.67	0.80	0.71	0.64	0.61	0.89

(c) Overview of losses and premiums

Figure 5.4: Data set of Farmers' Automobile (Commercial Automobile)

5.2.2 | State Farm (Triangle 2.2)

State Farm is the largest Commercial Automobile insurer (based on net premiums) that we analyse in this research. Its data is given in Figure 5.5 and Table C.5. We see a relatively stable (and thus predictable) loss development along all axis. When looking at the premiums, we see an increase that is not necessarily reflected in the claim volume. As such, the Loss Ratio's have declined over gradually over time. Models applying Premium information should be able to capture this behaviour.



(a) Upper Triangle (b) Observed Losses

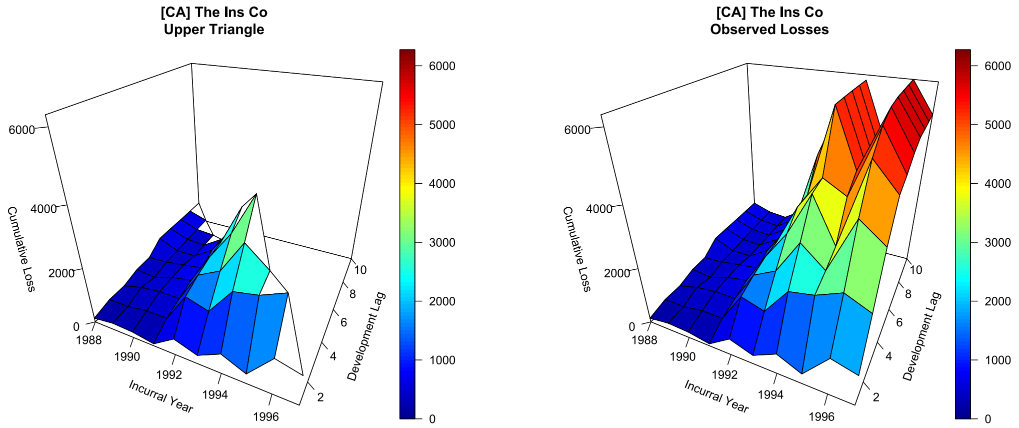
	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	193k	200k	225k	207k	215k	233k	244k	240k	237k	231k
Premium	286k	309k	326k	333k	342k	356k	380k	399k	407k	407k
Loss Ratio	0.68	0.65	0.69	0.62	0.63	0.66	0.64	0.60	0.58	0.57

(c) Overview of losses and premiums

Figure 5.5: Data set of State Farm (Commercial Automobile)

5.2.3 | The Ins Co (Triangle 2.3)

Upon inspection of the data of The Ins Co (Figure 5.6, we see a highly volatile claims development, most notably in Incurral Years from 1992 onward. We also see a significant growth in premium volume. Claims development in the incurral years up to and including 1991 might therefore be less relevant. Model performance on such an unstable triangle is interesting. The upper triangle on losses does show indicators of enlarged claims development, but which model is able to capture and estimate these characteristics best is difficult to hypothesise.



(a) Upper Triangle

(b) Observed Losses

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	791	553	479	936	2,254	4,458	6,117	4,480	6,273	5,322
Premium	1,300	1,697	1,553	1,425	3,252	5,617	9,536	5,674	7,389	6,378
Loss Ratio	0.61	0.33	0.31	0.66	0.69	0.79	0.64	0.79	0.85	0.83

(c) Overview of losses and premiums

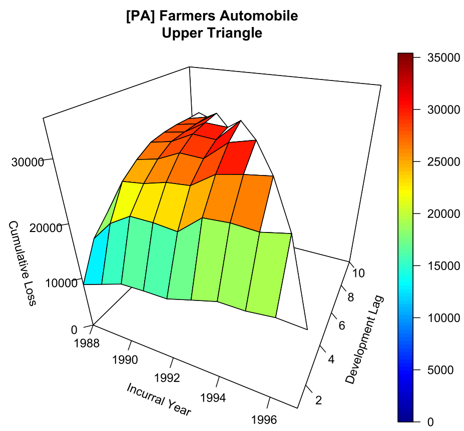
Figure 5.6: Data set of The Ins Co (Commercial Automobile)

5.3 | Personal Automobile (PA)

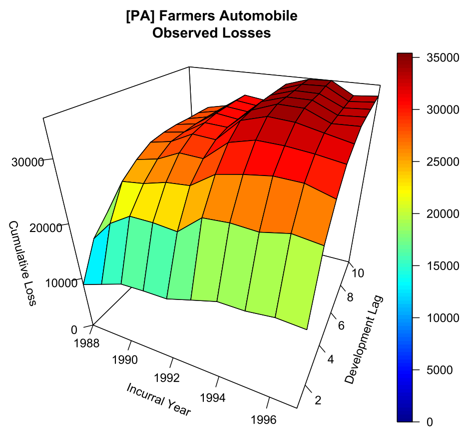
Personal Automobile insurances are much alike Commercial Automobile in regards to claims development (Shi, 2017). However, paid claims appear to be a bit slower: having a more gradual cumulative development in the early development-lags, and levelling out around $DL = 6$.

5.3.1 | Farmers' Automobile (Triangle 3.1)

The Farmers' Automobile triangle and losses are displayed in Figure 5.7. It is not very volatile throughout time, as no significant shocks can be observed. Both premium and claims volume are aligned with regards to growth. As such, the models should be able to predict these losses adequately.



(a) Upper Triangle



(b) Observed Losses

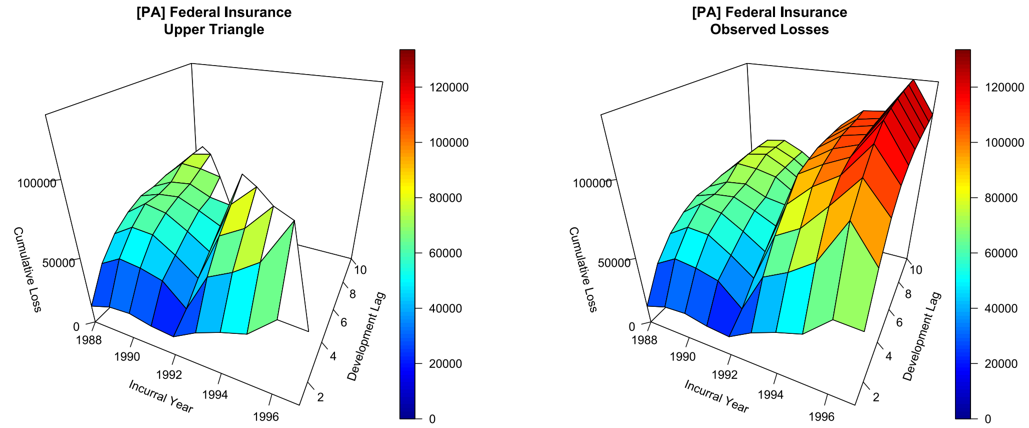
	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	21k	28k	29k	31k	31k	34k	35k	35k	33k	34k
Premium	28k	31k	37k	40k	43k	43k	44k	45k	45k	47k
Loss Ratio	0.73	0.89	0.79	0.78	0.72	0.79	0.81	0.80	0.74	0.71

(c) Overview of losses and premiums

Figure 5.7: Data set of Farmers' Automobile (Commercial Automobile)

5.3.2 | Federal Insurance Company (Triangle 3.2)

In Figure 5.8b, we see that on the incurral year-axis a growth of claims volume has occurred, while the development-lag axis has few fluctuations. The Gaussian process model might run into over-estimation with regards to the losses in incurral year 1997, caused by the larger volume in the previous incurral year. Premium volumes, however, are aligned with this. As such, we hypothesise that this extra information will aid the model.



(a) Upper Triangle (b) Observed Losses

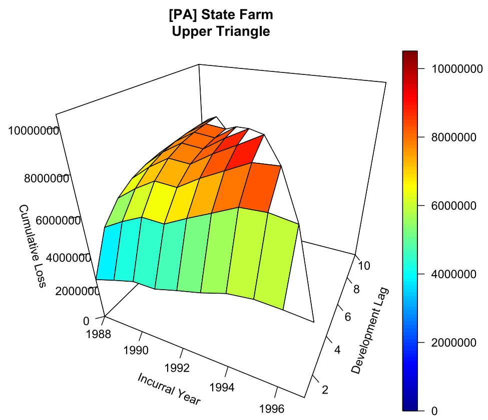
	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	64k	77k	79k	73k	62k	98k	110k	112k	134k	113k
Premium	83k	92k	96k	99k	96k	139k	152k	168k	181k	165k
Loss Ratio	0.76	0.84	0.83	0.73	0.65	0.70	0.72	0.66	0.74	0.69

(c) Overview of losses and premiums

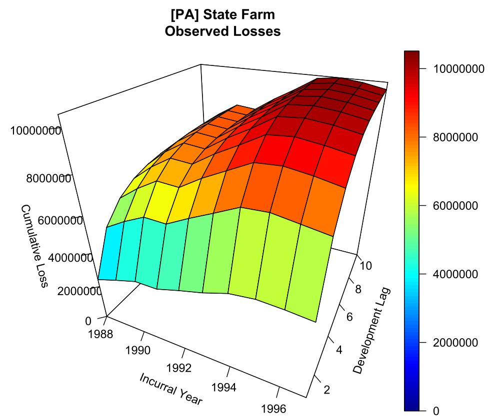
Figure 5.8: Data set of Federal Insurance Company (Commercial Automobile)

5.3.3 State Farm (Triangle 3.3)

The State Farm Personal Automobile insurance is the largest triangle with regards to claims volume. As such, it has few outliers and appears to have a fairly smooth development over time, as is displayed in Figure 5.9. We expect the models to be able to reflect this behaviour.



(a) Upper Triangle



(b) Observed Losses

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	6.8M	7.7M	8.4M	8.3M	9.0M	9.7M	10.4M	10.5M	10.4M	10.2M
Premium	7.8M	8.7M	9.8M	10.6M	11.5M	12.2M	13.3M	14.1M	14.7M	14.9M
Loss Ratio	0.87	0.88	0.86	0.78	0.79	0.79	0.78	0.74	0.71	0.68

(c) Overview of losses and premiums

Figure 5.9: Data set of State Farm (Personal Automobile)

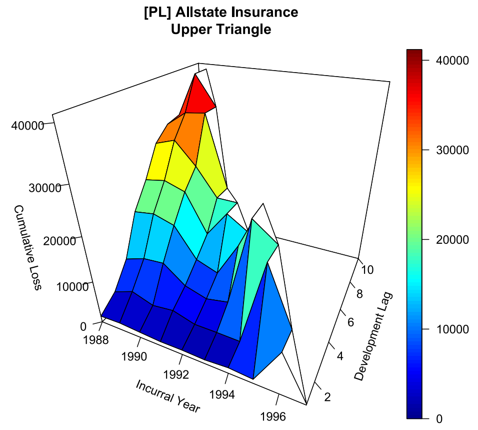
5.4 Product Liability (PL)

Product Liability is a type of insurance policy that protects companies that develop and/or create products from liability claims caused by those products, for instance in the case of a defect. We see in general that the development of these claims are more volatile than the other triangles considered so far, and they are thus challenging for any model to capture adequately.

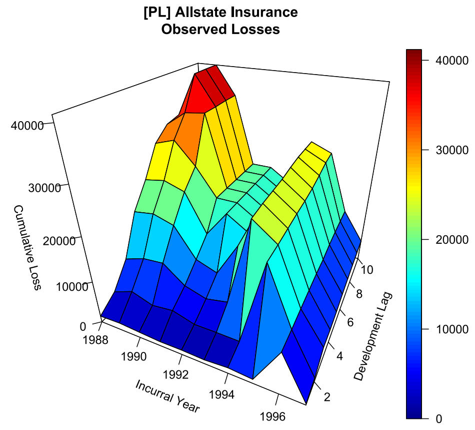
5.4.1 Allstate Insurance Company (Triangle 4.1)

The Allstate Insurance Company is one of the insurers considered for this Line of Business. Upon inspection of the 3D-plots, we observe that the run-off period for all Incurral Years is relatively flat in the development lag-axis. Furthermore, IY 1997 has a low claims volume, which is reflected in the net premium received, but might result in over-estimation if this is unknown.

The cause of this flat run-off period might be a diagonal shock that we observe in Calendar Year 1996 (not to be confused with incurral year 1996). We verify this by making a plot of the incurral claims, which is given in Figure 5.11, where such a diagonal shock can be observed. Also taking the significantly reduced claims and premiums in 1997 into account, we hypothesise that a business action could be underlying these changes. We do not expect any model to be able to adequately reflect this characteristic.



(a) Upper Triangle



(b) Observed Losses

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	36k	41k	35k	19k	19k	16k	27k	25k	12k	5,049
Premium	49k	47k	39k	32k	22k	24k	34k	27k	23k	4,450
Loss Ratio	0.75	0.88	0.89	0.61	0.86	0.67	0.79	0.95	0.51	1.13

(c) Overview of losses and premiums

Figure 5.10: Data set of Allstate Insurance Company (Product Liability)

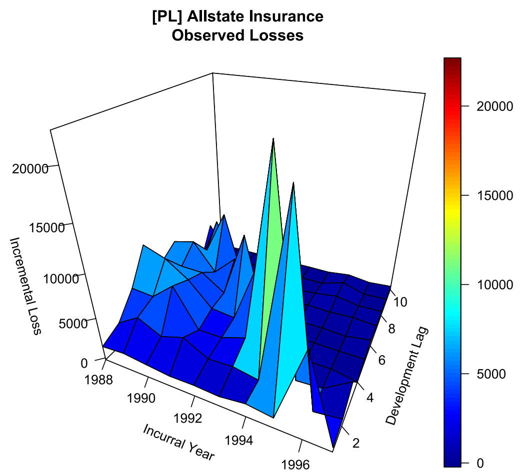
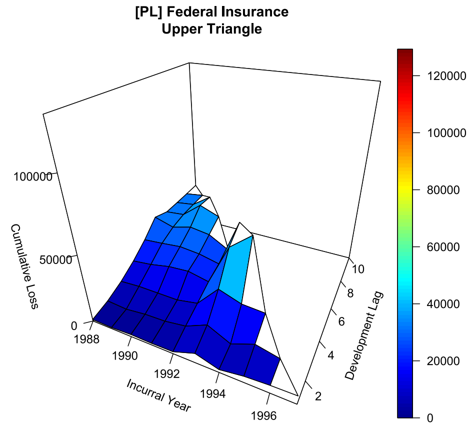


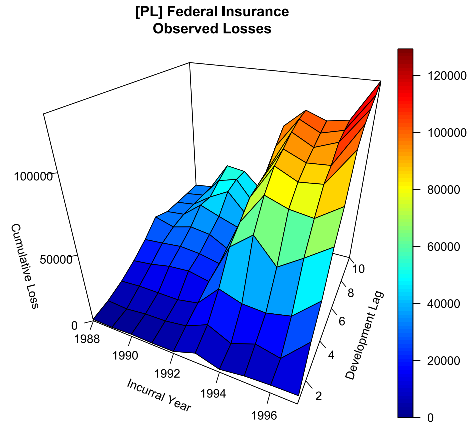
Figure 5.11: Incremental losses of Allstate Insurance Company (Product Liability)

5.4.2 | Federal Insurance Company (Triangle 4.2)

This triangle, as opposed to Triangle 4.1, has a positive development in the incurral year-axis, especially from 1993 onwards. This is partially reflected in the information in the upper triangle, but considering most of the development takes place in the lower triangle, we wonder if any model is able to capture this behaviour accordingly. Furthermore, this behaviour is not adequately captured in the premium information.



(a) Upper Triangle



(b) Observed Losses

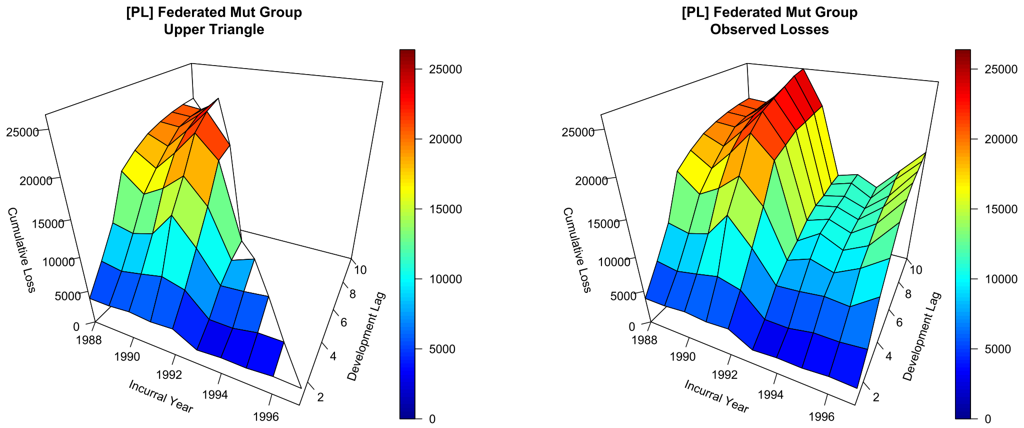
	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	32k	34k	56k	54k	41k	94k	106k	101k	104k	129k
Premium	102k	99k	100k	104k	105k	114k	124k	140k	150k	150k
Loss Ratio	0.32	0.35	0.55	0.52	0.39	0.82	0.86	0.72	0.69	0.86

(c) Overview of losses and premiums

Figure 5.12: Data set of Federal Insurance Company (Product Liability)

5.4.3 | Federated Mutual Group (Triangle 4.3)

Finally in the Product Liability Line of Business, we consider the Federated Mutual Group. This insurance company has seen a descent in claims volume with regards to the incurral year axis, while having a relatively stable development pattern, especially when comparing this to its' peers. This decline in claims volume is not reflected in the premium information, and we thus hypothesise that all models will have considerable difficulty in making predictions.



(a) Upper Triangle (b) Observed Losses

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	21k	21k	23k	26k	22k	12k	12k	11k	14k	17k
Premium	30k	28k	28k	28k	29k	24k	26k	28k	30k	32k
Loss Ratio	0.72	0.76	0.82	0.94	0.75	0.48	0.47	0.38	0.47	0.54

(c) Overview of losses and premiums

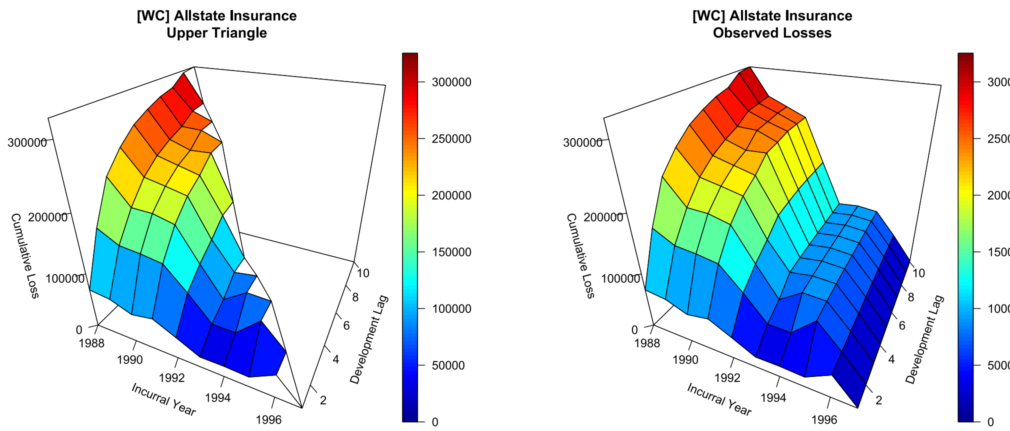
Figure 5.13: Data set of Federated Mutual Group (Product Liability)

5.5 | Workers' Compensation (WC)

Finally, we look at Workers' Compensation lines of business, which insures the policy holders' income and treatment in the case of a medical injury occured at their work.

5.5.1 | Allstate Insurance Company (Triangle 5.1)

The most notable trend that can be seen in Figure 5.14 is the decline of claims volumes along the incurral year-axis, also reflected in the premium volumes. Furthermore, we see that this triangle is relatively stable from $DL \geq 4$, and that the reduced claims volumes is forebode by the reduced claims in the first development lag period of 1997. We thus expect the models to be able to estimate this appropriately.



(a) Upper Triangle (b) Observed Losses

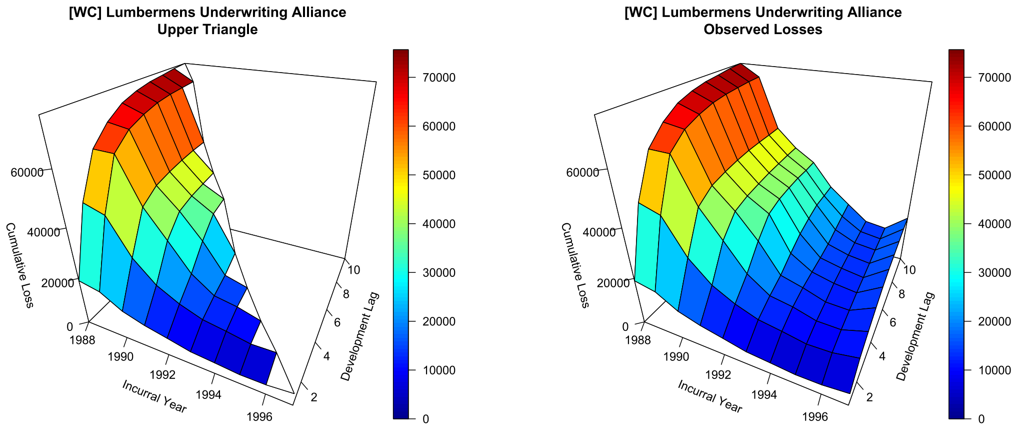
	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	325k	278k	263k	248k	169k	91k	95k	91k	49k	2,909
Premium	395k	374k	280k	314k	253k	201k	174k	146k	93k	7,651
Loss Ratio	0.82	0.74	0.94	0.79	0.67	0.45	0.54	0.62	0.53	0.38

(c) Overview of losses and premiums

Figure 5.14: Data set of Allstate Insurance Company (Workers' Compensation)

5.5.2 | Lumbermen's Underwriting Alliance (Triangle 5.2)

Much like Triangle 5.1, we see that the losses from the Lumbermen's Underwriting Alliance (LUA) have a declining trend with regards to the incurral year-axis. Furthermore, most of the development has occured prior to $DL = 5$. As such, it shows a fairly similar pattern to the AIC triangle, and the models should show roughly similar results.



(a) Upper Triangle (b) Observed Losses

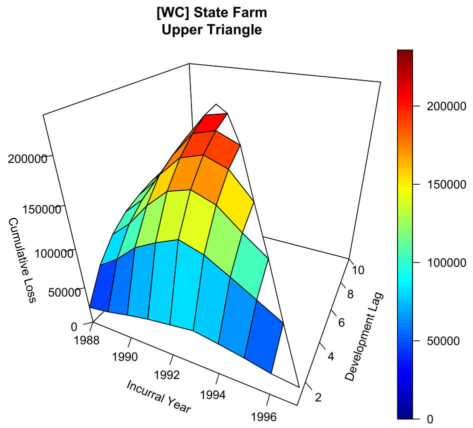
	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	76k	71k	50k	43k	38k	28k	21k	15k	14k	21k
Premium	94k	92k	68k	60k	62k	71k	62k	68k	50k	49k
Loss Ratio	0.80	0.77	0.74	0.72	0.60	0.39	0.34	0.22	0.27	0.42

(c) Overview of losses and premiums

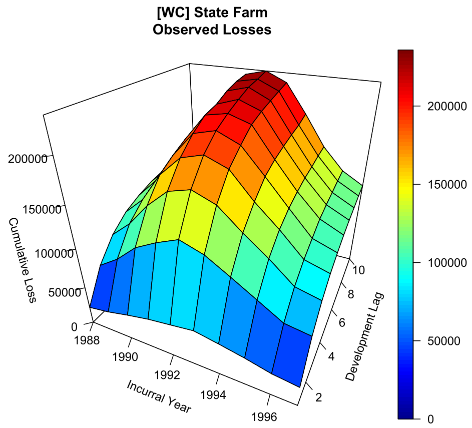
Figure 5.15: Data set of Lumbermen's Underwriting Alliance (Workers' Compensation)

5.5.3 | State Farm (Triangle 5.3)

The State Farm triangle has been used in the paper of Lally and Hartman (2018) to validate the applicability of the Gaussian Process model on predicting Loss Reserves. When looking at the triangle, we once again see a clear trend emerging on the incurral year-axis, while the development stabilises from $DL \geq 5$, much like the two peers that we will analyse in this Line of Business. The Loss Ratio is gradually declining, and models incorporating premium information could thus benefit from this information.



(a) Upper Triangle



(b) Observed Losses

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	125k	149k	196k	230k	236k	225k	191k	153k	126k	112k
Premium	177k	201k	246k	286k	340k	419k	366k	338k	287k	245k
Loss Ratio	0.71	0.74	0.80	0.80	0.69	0.54	0.52	0.45	0.44	0.45

(c) Overview of losses and premiums

Figure 5.16: Data set of State Farm (Workers' Compensation)

6 | Results

Contents

6.1 Benchmark measurements **49**

6.2 Prior distributions **51**

6.2.1 Bandwidth parameter (ψ) 51

6.2.2 Noise parameter (σ^2) 54

6.2.3 Signal parameter (η^2) 57

6.2.4 Warping parameters (α, β) 58

6.2.5 Optimal prior configuration 59

6.3 Premium Information **63**

6.3.1 Loss Ratio's 63

6.3.2 Premium as input 67

6.3.3 Bornheutter-Ferguson Estimations 71

In this chapter, we will present the results of the research performed. We will follow notation of Chapter 4, where our research is further outlined. The triangles that we will analyse on are described in detail in Chapter 5.

For each research, we will give both the predicted Loss Reserve and the Root Mean Square Error of prediction. Furthermore, we will highlight several $\mathbb{R}^3$-plots where the models give adequate predictions or where they underperform.

In the tables comparing predicted Loss Reserves, we have marked the best prediction in **bold**. For the tables comparing RMSE calculations and Standard Deviations, the lowest error or deviation has been marked in **bold**.

6.1 | Benchmark measurements

We will analyse the results of the Chain Ladder method (as described in Section 2.3.1) and the Gaussian Process model as defined by Lally and Hartman (2018) and described in Section 3.

The goal of this is threefold: we want to verify our implementation, get a better understanding of the prediction and the current error of the models, but also perform this analysis to answer the following research question:

- *Is the model applicable on a more extensive data set?*

Triangles that are already analysed by Lally and Hartman (2018) are 1.3, 3.1 and 5.3.

An overview of the predicted Loss Reserve by these models is given in Table 6.1, which is compared to the actual observed losses. In Table 6.2, an overview of the RMSE of all methods can be found. Finally, we give an overview of the Standard Deviation of all predicted Loss Reserves in Table 6.3, to give an indication of the confidence of the predictions.

Table 6.1: Benchmark Measurement: Observed and predicted Loss Reserves

#	LoB	Observed	Chain Ladder	Matérn 3/2	Matérn 5/2	Sq. Exponential
1.1	MM	407,525	740,677	288,484	**313,911**	284,687
1.2	MM	104,094	63,104	**83,937**	77,709	77,340
1.3	MM	164,633	240,423	155,306	**156,467**	153,394
2.1	CA	16,640	**17,239**	14,254	13,789	13,891
2.2	CA	353,949	410,384	329,846	345,927	**350,702**
2.3	CA	13,046	**11,377**	6,627	6,993	6,652
3.1	PA	37,397	42,833	38,901	40,278	**37,737**
3.2	PA	137,642	367,607	280,729	368,464	**212,727**
3.3	PA	11,561,327	12,586,821	10,113,582	10,996,920	**11,087,032**
4.1	PL	10,965	162,098	**90,421**	103,467	98,422
4.2	PL	422,513	**325,328**	179,658	231,733	212,941
4.3	PL	37,612	**36,863**	35,031	32,638	28,782
5.1	WC	45,916	**193,320**	310,663	334,577	325,816
5.2	WC	40,225	**34,490**	55,831	55,688	51,710
5.3	WC	307,810	**304,882**	328,945	343,456	311,081

Table 6.2: Benchmark Measurement: RMSE

#	LoB	Chain Ladder	Matérn 3/2	Matérn 5/2	Sq. Exponential
1.1	MM	60,512	13,756	**11,015**	13,416
1.2	MM	6,735	**4,039**	4,786	4,904
1.3	MM	15,544	6,675	6,153	**6,107**
2.1	CA	**896**	1,472	1,354	1,347
2.2	CA	5,781	5,010	**4,217**	4,727
2.3	CA	**629**	1,368	1,155	1,100
3.1	PA	1,685	**1,042**	1,378	1,420
3.2	PA	43,475	17,015	28,351	**12,590**
3.3	PA	328,024	224,798	**163,189**	197,843
4.1	PL	24,160	**11,898**	14,798	16,188
4.2	PL	**12,996**	26,986	19,588	20,015
4.3	PL	**1,592**	2,476	2,223	2,058
5.1	WC	**20,100**	34,701	44,033	53,562
5.2	WC	**2,413**	3,505	3,716	4,126
5.3	WC	**7,165**	14,815	14,934	10,062

When comparing Table 6.1 and Table 6.2, we notice that, with some exceptions, a better prediction of the Loss Reserve results in a lower RMSE. In general, we observe from Table 6.1 that the Gaussian Process models are able to perform better or than or comparable to the Chain Ladder model, but is also outperformed in some triangles by the Chain Ladder. The GP method is generally outperformed by the CL in the Workers' Compensation and Product Liability Lines of Business, with the notable exception of Triangle 4.1 - which has a Diagonal Shock that neither model is able to capture accordingly.

When comparing the covariance functions of the Gaussian Process model, we see that there is no significant outperformance of any one along all triangles. In general, when comparing the predictions, we see that the triangles that have a smooth pattern are better predicted by the

Squared Exponential function, and the models applying a Matérn covariance function perform better on the triangles that are more volatile.

In Appendix D, plots are given of the actual observed losses, the Chain Ladder prediction and the best performing Gaussian process model in order to visualise the predictions by these models.

Table 6.3: Benchmark Measurement: Standard Deviation of Loss Reserve

#	LoB	Chain Ladder	Matérn 3/2	Matérn 5/2	Sq. Exponential
1.1	MM	106,809	86,625	90,804	**83,825**
1.2	MM	**22,698**	33,058	27,092	26,187
1.3	MM	**30,106**	41,430	36,476	37,098
2.1	CA	**2,472**	5,458	4,712	4,128
2.2	CA	**18,221**	73,495	74,844	65,842
2.3	CA	4,760	**3,931**	4,061	4,335
3.1	PA	**2,675**	8,039	8,852	8,223
3.2	PA	**51,036**	92,948	118,434	98,847
3.3	PA	**549,869**	2,400,639	2,616,817	2,541,395
4.1	PL	86,611	57,195	**56,209**	58,213
4.2	PL	**84,093**	95,199	128,196	116,028
4.3	PL	**4,702**	21,155	21,202	18,783
5.1	WC	**49,582**	229,973	226,372	202,139
5.2	WC	**7,378**	37,775	37,127	26,968
5.3	WC	**20,364**	107,898	120,175	104,266

An overview of the Standard Deviation of the Loss Reserve prediction is given in Table 6.3. In general, the Chain Ladder has the lowest Standard Deviation, with three exceptions. In the case of all three exceptions, the predictions made by the GP models are considerably lower than the prediction by the Chain Ladder. Considering that the wide density of the Loss Reserve was a limitation of the Gaussian process model, this meets our expectations.

6.2 | Prior distributions

In this section, we will vary the prior distributions as explained in Section 4.2. We benchmark the different prior distribution of the parameter that we wish to investigate against the results of the Gaussian process models as given in Section 6.1.

The relevant research question that we attempt to answer with this analysis is:

- *Are the prior distributions on the hyperparameters adequately chosen?*

6.2.1 | Bandwidth parameter (ψ)

We have changed the prior distribution of the bandwidth parameter from a *Gamma(4,4)* to a *Cauchy(0,2.5)* distribution. Table 6.4 presents the predicted loss reserves and Table 6.5 contains the RMSE of the predictions. The results of the estimated parameters are given in Appendix E.

The Cauchy prior distribution for the bandwidth parameter has varying results: sometimes it results in better predictions and/or a lower RMSE, but its' performance is unstable - once even suggesting a negative loss reserve (marked in red). As can be seen in Tables E.1 and E.2 (Appendix E), the inferred values of ψ are remarkably high, and thus would result in more volatile functions as described in Section 3.2. Especially in the case of the Squared Exponential

Table 6.4: Loss Reserve predictions of variations on the prior distribution of ψ

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	Observed	Cauchy	Gamma	Cauchy	Gamma	Cauchy	Gamma
1.1	407,525	269,104	**288,484**	263,070	**313,911**	216,849	**284,687**
1.2	104,094	80,073	**83,937**	69,541	**77,709**	63,729	**77,340**
1.3	164,633	151,235	**155,306**	150,681	**156,467**	137,297	**153,394**
2.1	16,640	12,372	**14,254**	12,169	**13,789**	9,285	**13,891**
2.2	353,949	323,956	**329,846**	334,710	**345,927**	324,372	**350,702**
2.3	13,046	3,101	**6,627**	2,059	**6,993**	-838	**6,652**
3.1	37,397	**37,993**	38,901	**38,613**	40,278	38,126	**37,737**
3.2	137,642	**209,772**	280,729	**221,388**	368,464	**166,334**	212,727
3.3	11,561,327	**10,124,963**	10,113,582	10,645,458	**10,996,920**	10,447,237	**11,087,032**
4.1	10,965	**48,026**	90,421	**57,721**	103,467	**62,062**	98,422
4.2	422,513	99,228	**179,658**	81,173	**231,733**	19,260	**212,941**
4.3	37,612	32,338	**35,031**	**32,725**	32,638	**30,524**	28,782
5.1	45,916	**285,877**	310,663	**268,616**	334,577	**303,043**	325,816
5.2	40,225	**52,439**	55,831	**53,097**	55,688	**46,105**	51,710
5.3	307,810	**323,255**	328,945	**330,246**	343,456	302,472	**311,081**

Table 6.5: Root Mean Square Errors of variations on the prior distribution of ψ

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	LoB	Cauchy	Gamma	Cauchy	Gamma	Cauchy	Gamma
1.1	MM	15,868	**13,756**	15,934	**11,015**	18,284	**13,416**
1.2	MM	**3,598**	4,039	**4,779**	4,786	5,831	**4,904**
1.3	MM	6,863	**6,675**	6,429	**6,153**	6,658	**6,107**
2.1	CA	1,676	**1,472**	1,709	**1,354**	1,916	**1,347**
2.2	CA	5,639	**5,010**	**3,878**	4,217	**3,983**	4,727
2.3	CA	1,897	**1,368**	2,062	**1,155**	2,436	**1,100**
3.1	PA	**975**	1,042	**1,219**	1,378	**1,386**	1,420
3.2	PA	**8,818**	17,015	**10,086**	28,351	**7,001**	12,590
3.3	PA	**221,335**	224,798	**150,992**	163,189	**138,657**	197,843
4.1	PL	**6,149**	11,898	**7,446**	14,798	**9,085**	16,188
4.2	PL	38,779	**26,986**	40,347	**19,588**	47,914	**20,015**
4.3	PL	2,601	**2,476**	2,245	**2,223**	**1,965**	2,058
5.1	WC	**30,970**	34,701	**28,873**	44,033	**34,910**	53,562
5.2	WC	**3,174**	3,505	**3,267**	3,716	**2,844**	4,126
5.3	WC	**13,365**	14,815	**13,157**	14,934	**9,304**	10,062

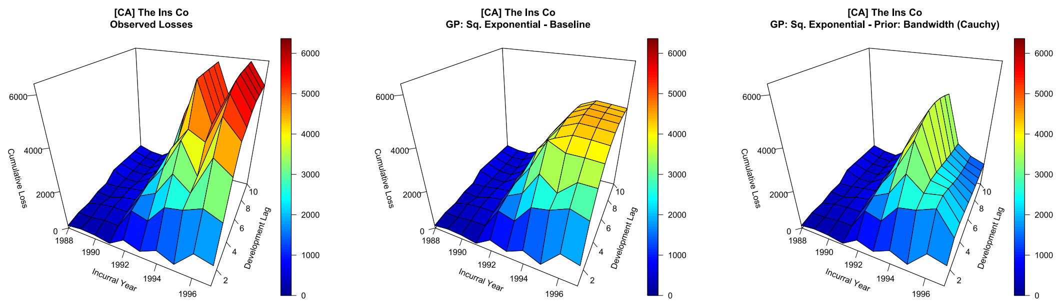
covariance function, the Cauchy prior distribution appears to get too much freedom. Performance of the Cauchy prior seems to be a better combination with the Matérn covariance functions.

We visualise results by plotting Triangle 2.3 (Squared Exponential), Triangle 4.1 (Matérn 3/2) and Triangle 4.2 (Squared Exponential) in respectively Figures 6.1, 6.2 and 6.3.

Figure 6.1c (Triangle 2.3 - Sq. Exponential) displays a weakness of the Cauchy prior: the instability of the triangle is captured for the Incurral Year of 1994, but it does not adequately capture the larger claims of 1995, 1996 and 1997, as a result of the flexibility caused by the large bandwidth parameter of the Incurral Year ($\psi_2 = 45.0727$, Table E.2). On the other hand, the $Gamma(4, 4)$ prior distribution (Figure 6.1b) is unable to capture any of the outliers in this unstable triangle.

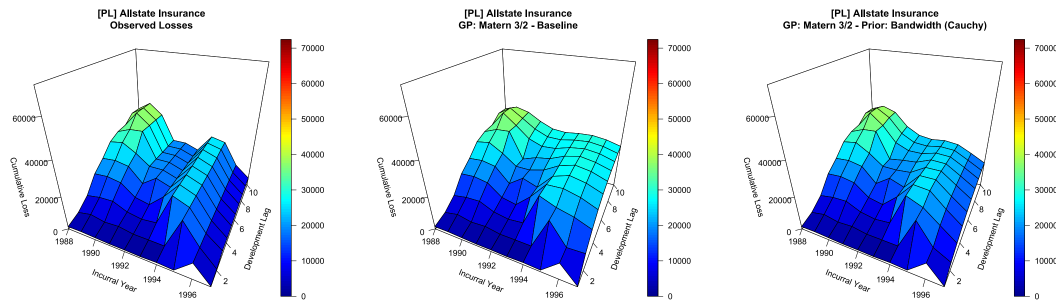
This is also visible in the results of Triangle 4.2 (Figure 6.3, where the Cauchy-distribution results in unrealistic estimates: the increased losses (starting from 1994 onward) appear to be interpreted as an outlier. As a result, the run-off for 1994 is very unrealistic. We have also included the predicted run-off in Table 6.6 - from which it is visible that the GP model predicts losses to become lower over time. We have highlighted decreasing values on the DL-axis in red, which is uncommon behaviour for a run-off triangle.

On the other hand, Figure 6.2c (Triangle 4.1, Matérn 3/2) displays how the increased flexibility of the Cauchy prior improves the model predictions when compared to the Gamma(4,4) prior (Figure 6.2b). Trends in the data are more adequately captured, resulting in a significant improvement of both the Loss Reserve prediction and the RMSE.



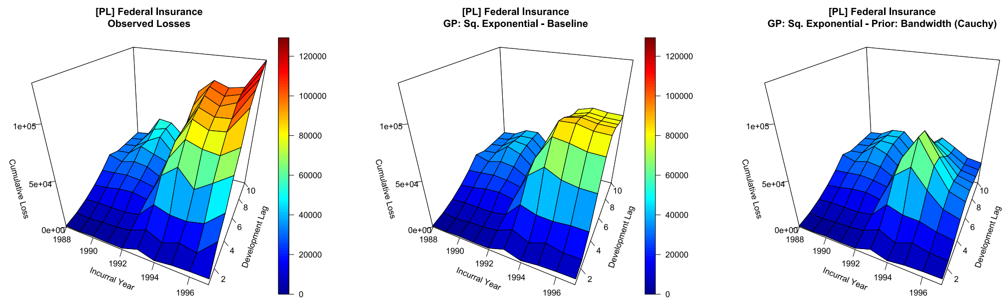
(a) Observed Losses (b) Gamma, Sq. Exp. (c) Cauchy, Sq. Exp.

Figure 6.1: Visualisation of models with variations on the bandwidth parameter of Triangle 2.3



(a) Observed Losses (b) Gamma, Matérn 3/2 (c) Cauchy, Matérn 3/2

Figure 6.2: Visualisation of models with variations on the bandwidth parameter of Triangle 4.1



(a) Observed Losses (b) Gamma, Sq. Exp. (c) Cauchy, Sq. Exp.

Figure 6.3: Visualisation of models with variations on the bandwidth parameter of Triangle 4.2

Table 6.6: Run-off prediction by the half-Cauchy distribution, applied on Triangle 4.2, Squared Exponential as displayed in Figure 6.3c

IY	Development Lag									
	1	2	3	4	5	6	7	8	9	10
1988	1,249	4,092	8,893	15,516	23,806	34,070	30,699	31,130	31,716	32,430
1989	946	2,929	9,953	17,368	25,139	29,628	32,059	32,972	33,610	34,014
1990	1,765	4,743	9,854	20,471	26,756	33,654	43,282	42,588	41,860	41,616
1991	1,408	5,226	9,768	21,637	25,339	31,574	37,738	41,942	43,012	42,786
1992	1,647	7,628	12,848	18,323	21,943	23,139	25,241	27,277	28,363	28,554
1993	7,566	14,991	29,466	49,090	57,590	53,392	44,038	37,063	33,514	32,481
1994	2,299	6,856	23,954	61,496	80,918	80,006	66,574	54,359	47,405	45,087
1995	4,959	13,541	23,003	43,597	57,732	57,995	50,237	42,725	38,201	36,736
1996	6,063	9,707	17,840	28,960	36,698	37,334	33,883	30,537	28,519	27,815
1997	6,507	8,607	14,968	24,051	30,890	32,010	29,788	27,452	26,063	25,549

6.2.2 Noise parameter (σ^2)

We will now present the results on the noise-parameter of the observations (σ^2). The predicted Loss Reserves are given in Table 6.7, and the RMSE of the mean predictions are presented in Table 6.8.

We observe that the variation of the prior distribution of σ^2 in general has a low impact. While we have accentuated the best prediction and lowest error, it can be argued if these differences are of any significance. While both prior distributions are weakly informative, the Student-T distribution in general performs better. For some triangles, however, a notable change of performance can be observed (e.g. Triangle 3.2). The estimations of σ^2 for each simulation is displayed in Table 6.9.

In general, we observe that $\hat{\sigma^2}$ tends to converge to zero. The Student-T prior distribution always results in lower estimates of $\hat{\sigma^2}$, which is as expected, considering the underlying distributions (Figure 4.2).

Assuming that the data-set is of such a quality that it consists of noise-free observations, a convergence to zero is to be expected. However, in case data quality would be worse, the log-normal distribution might perform better. The minor variations in the predictions between the two prior distributions is in line with the minor differences of the Loss Reserve and RMSE.

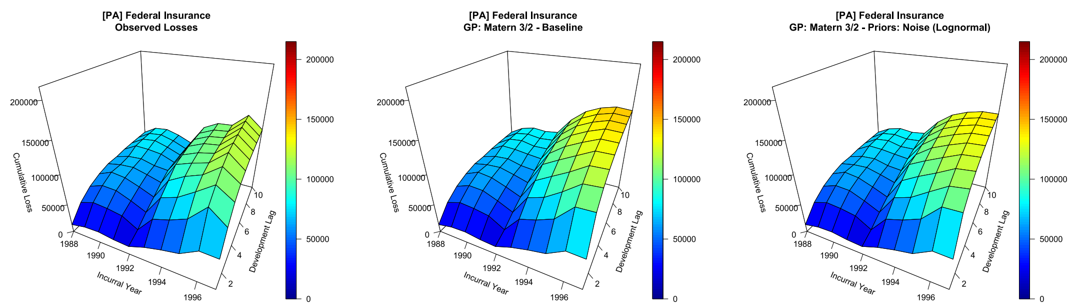
We visualise Triangle 3.2 (Matérn 3/2) and 3.3 (Sq. Exponential) in Figures 6.4 and 6.5. As can be seen, differences between both prior distributions are minimal.

Table 6.7: Estimated Loss Reserves of variations on the prior distribution of σ^2

#	Observed	Matérn 3/2		Matérn 5/2		Sq. Exponential	
		Log-normal	Student-T	Log-normal	Student-T	Log-normal	Student-T
1.1	407,525	272,601	**288,484**	298,574	**313,911**	**306,052**	284,687
1.2	104,094	74,194	**83,937**	72,700	**77,709**	71,019	**77,340**
1.3	164,633	**155,526**	155,306	**159,180**	156,467	**154,512**	153,394
2.1	16,640	12,485	**14,254**	13,471	**13,789**	13,518	**13,891**
2.2	353,949	324,416	**329,846**	344,222	**345,927**	325,100	**350,702**
2.3	13,046	5,672	**6,627**	6,777	**6,993**	**6,673**	6,652
3.1	37,397	**36,923**	38,901	**37,068**	40,278	35,782	**37,737**
3.2	137,642	**244,253**	280,729	**307,681**	368,464	**195,412**	212,727
3.3	11,561,327	9,957,777	**10,113,582**	10,629,365	**10,996,920**	10,680,209	**11,087,032**
4.1	10,965	90,796	**90,421**	**101,121**	103,467	**96,122**	98,422
4.2	422,513	174,581	**179,658**	208,862	**231,733**	200,162	**212,941**
4.3	37,612	**38,584**	35,031	**36,824**	32,638	**30,206**	28,782
5.1	45,916	377,474	**310,663**	385,721	**334,577**	347,118	**325,816**
5.2	40,225	60,455	**55,831**	59,887	**55,688**	53,423	**51,710**
5.3	307,810	342,114	**328,945**	355,063	**343,456**	313,746	**311,081**

Table 6.8: Root Mean Square Errors of variations on the prior distribution of σ^2

#	LoB	Matérn 3/2		Matérn 5/2		Sq. Exponential	
		Log-normal	Student-T	Log-normal	Student-T	Log-normal	Student-T
1.1	MM	15,345	**13,756**	12,703	**11,015**	**11,115**	13,416
1.2	MM	4,906	**4,039**	5,195	**4,786**	5,324	**4,904**
1.3	MM	6,697	**6,675**	6,158	**6,153**	6,251	**6,107**
2.1	CA	**1,461**	1,472	1,372	**1,354**	1,366	**1,347**
2.2	CA	5,855	**5,010**	4,764	**4,217**	6,565	**4,727**
2.3	CA	1,371	**1,368**	1,158	**1,155**	1,102	**1,100**
3.1	PA	1,105	**1,042**	**1,264**	1,378	**1,358**	1,420
3.2	PA	**13,707**	17,015	**21,194**	28,351	**11,500**	12,590
3.3	PA	251,794	**224,798**	204,904	**163,189**	278,615	**197,843**
4.1	PL	**11,889**	11,898	**14,522**	14,798	**15,526**	16,188
4.2	PL	27,727	**26,986**	22,047	**19,588**	21,738	**20,015**
4.3	PL	2,563	**2,476**	2,331	**2,223**	2,253	**2,058**
5.1	WC	45,989	**34,701**	52,412	**44,033**	**46,797**	53,562
5.2	WC	3,952	**3,505**	4,206	**3,716**	4,180	**4,126**
5.3	WC	16,944	**14,815**	16,780	**14,934**	10,840	**10,062**

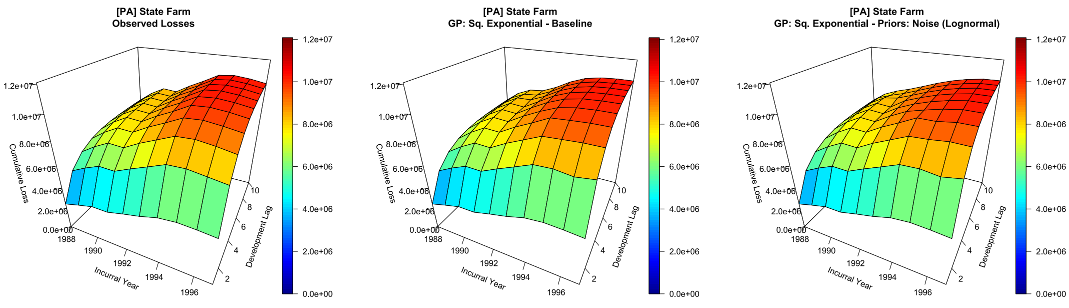


(a) Observed Losses (b) Student-T, Matérn 3/2 (c) Lognormal, Matérn 3/2

Figure 6.4: Visualisation of models with variations on the noise parameter of Triangle 3.2

Table 6.9: Estimations of σ^2, split per triangle and prior distribution

$\hat{\sigma}$		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	LoB	Log-normal	Student-T	Log-normal	Student-T	Log-normal	Student-T
1.1	MM	0.0117	0.0020	0.0142	0.0027	0.0197	0.0086
1.2	MM	0.0439	0.0228	0.0529	0.0425	0.0582	0.0484
1.3	MM	0.0083	0.0014	0.0083	0.0016	0.0090	0.0019
2.1	CA	0.0354	0.0039	0.0435	0.0348	0.0474	0.0405
2.2	CA	0.0067	0.0009	0.0068	0.0008	0.0158	0.0007
2.3	CA	0.0388	0.0075	0.0618	0.0495	0.0669	0.0584
3.1	PA	0.0085	0.0015	0.0091	0.0020	0.0098	0.0027
3.2	PA	0.0154	0.0024	0.0205	0.0029	0.0768	0.0634
3.3	PA	0.0061	0.0008	0.0065	0.0008	0.0090	0.0008
4.1	PL	0.0767	0.0641	0.0872	0.0763	0.0911	0.0794
4.2	PL	0.0192	0.0042	0.0203	0.0063	0.0465	0.0133
4.3	PL	0.0093	0.0016	0.0091	0.0017	0.0096	0.0017
5.1	WC	0.0097	0.0014	0.0113	0.0017	0.0272	0.0033
5.2	WC	0.0063	0.0009	0.0074	0.0010	0.0088	0.0019
5.3	WC	0.0050	0.0007	0.0044	0.0006	0.0041	0.0006



(a) Observed Losses (b) Student-T, Sq. Exponential (c) Lognormal, Sq. Exponential

Figure 6.5: Visualisation of models with variations on the noise parameter on Triangle 3.3

6.2.3 | Signal parameter (η^2)

In this section, we investigate the effect of the prior distribution of the Signal parameter. The estimated Loss Reserves are given in Table 6.10 and the Root Mean Square Error of these predictions in Table 6.11.

Table 6.10: Estimated Loss Reserves of variations on the prior distribution of η^2

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	Observed	Log-normal	Student-T	Log-normal	Student-T	Log-normal	Student-T
1.1	407,525	**293,655**	288,484	**321,546**	313,911	**289,649**	284,687
1.2	104,094	**88,815**	83,937	**81,078**	77,709	**80,300**	77,340
1.3	164,633	**157,072**	155,306	**161,363**	156,467	**154,665**	153,394
2.1	16,640	**14,699**	14,254	**14,468**	13,789	13,268	**13,891**
2.2	353,949	328,106	**329,846**	**351,765**	345,927	344,004	**350,702**
2.3	13,046	**7,152**	6,627	**7,774**	6,993	**7,674**	6,652
3.1	37,397	39,503	**38,901**	**39,328**	40,278	38,107	**37,737**
3.2	137,642	296,369	**280,729**	398,443	**368,464**	238,091	**212,727**
3.3	11,561,327	**10,120,266**	10,113,582	**11,123,106**	10,996,920	**11,090,064**	11,087,032
4.1	10,965	93,502	**90,421**	111,827	**103,467**	101,058	**98,422**
4.2	422,513	185,705	**179,658**	237,454	**231,733**	223,322	**212,941**
4.3	37,612	**35,638**	35,031	**33,748**	32,638	**29,072**	28,782
5.1	45,916	313,856	**310,663**	**315,743**	334,577	332,444	**325,816**
5.2	40,225	**54,835**	55,831	**50,707**	55,688	**50,550**	51,710
5.3	307,810	341,215	**328,945**	346,722	**343,456**	**305,868**	311,081

Table 6.11: RMSE of variations on the prior distribution of η^2

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	LoB	Log-normal	Student-T	Log-normal	Student-T	Log-normal	Student-T
1.1	MM	**13,515**	13,756	**10,091**	11,015	**12,693**	13,416
1.2	MM	**3,753**	4,039	**4,594**	4,786	**4,741**	4,904
1.3	MM	**6,520**	6,675	**6,099**	6,153	**6,087**	6,107
2.1	CA	**1,458**	1,472	**1,315**	1,354	1,367	**1,347**
2.2	CA	5,112	**5,010**	**4,006**	4,217	**4,687**	4,727
2.3	CA	**1,312**	1,368	**1,084**	1,155	**991**	1,100
3.1	PA	1,073	**1,042**	**1,335**	1,378	1,472	**1,420**
3.2	PA	18,659	**17,015**	31,302	**28,351**	15,287	**12,590**
3.3	PA	**221,738**	224,798	**154,732**	163,189	199,666	**197,843**
4.1	PL	11,928	**11,898**	15,316	**14,798**	16,813	**16,188**
4.2	PL	**26,209**	26,986	**19,254**	19,588	**18,971**	20,015
4.3	PL	2,487	**2,476**	**2,173**	2,223	**1,990**	2,058
5.1	WC	35,009	**34,701**	**39,683**	44,033	54,109	**53,562**
5.2	WC	**3,433**	3,505	**3,423**	3,716	**4,037**	4,126
5.3	WC	15,665	**14,815**	15,050	**14,934**	**9,887**	10,062

Minor differences can be observed in the variations of the prior distribution of the signal-parameter, analogous to the noise-parameter. The log-normal distribution appears to be a better combination with the Matérn 5/2 kernel function, while for both the Squared Exponential and the Matérn 3/2 kernel functions both prior distributions appear to be adequate. In all cases, the variation in prior distribution gives similar visual results (i.e. no difference in trends modeled are visible). Therefore, no further visualisations are given.

6.2.4 | Warping parameters (α, β)

The parameters for the Input Warping ensure that the stationary covariance function is able to make adequate estimations, even when the underlying process need not be stationary. We have investigated suggestions of these hyperpriors given in the original paper describing Input Warping (Snoek et al., 2014), against the benchmark of the model by Lally and Hartman (2018). The estimated Loss Reserve of these models are given in Table 6.12, and Table 6.13 gives the RMSE of all models. The researched setup of the parameters in these tables are indicated by the names of the respective authors. The inferred values of $\alpha_1, \beta_1, \alpha_2, \beta_2$ are given in Appendix E.

Table 6.12: Estimated Loss Reserves of variations on the prior distribution of Input Warping parameters

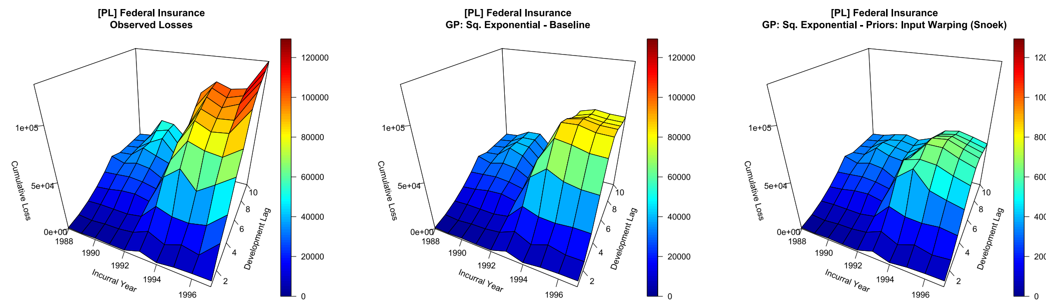
		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	Observed	Snoek	Lally	Snoek	Lally	Snoek	Lally
1.1	407,525	**290,902**	288,484	**318,213**	313,911	**299,256**	284,687
1.2	104,094	64,756	**83,937**	64,732	**77,709**	65,200	**77,340**
1.3	164,633	145,554	**155,306**	139,773	**156,467**	115,130	**153,394**
2.1	16,640	**14,946**	14,254	**14,236**	13,789	13,873	**13,891**
2.2	353,949	**342,116**	329,846	**360,514**	345,927	302,151	**350,702**
2.3	13,046	**7,115**	6,627	**7,231**	6,993	**6,910**	6,652
3.1	37,397	**37,544**	38,901	**35,779**	40,278	31,976	**37,737**
3.2	137,642	292,426	**280,729**	392,496	**368,464**	**178,848**	212,727
3.3	11,561,327	**10,374,343**	10,113,582	10,809,192	**10,996,920**	9,779,046	**11,087,032**
4.1	10,965	**89,576**	90,421	**103,009**	103,467	**95,305**	98,422
4.2	422,513	170,923	**179,658**	224,991	**231,733**	115,097	**212,941**
4.3	37,612	**39,693**	35,031	**40,824**	32,638	**33,318**	28,782
5.1	45,916	318,854	**310,663**	**314,423**	334,577	396,805	**325,816**
5.2	40,225	67,647	**55,831**	61,629	**55,688**	**50,727**	51,710
5.3	307,810	361,863	**328,945**	371,890	**343,456**	314,450	**311,081**

Table 6.13: RMSE of variations on the prior distribution of Input Warping parameters

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	LoB	Snoek	Lally	Snoek	Lally	Snoek	Lally
1.1	MM	14,046	**13,756**	**10,956**	11,015	**10,839**	13,416
1.2	MM	6,247	**4,039**	6,192	**4,786**	5,699	**4,904**
1.3	MM	7,071	**6,675**	6,793	**6,153**	7,373	**6,107**
2.1	CA	1,532	**1,472**	**1,352**	1,354	1,354	**1,347**
2.2	CA	6,419	**5,010**	**3,380**	4,217	6,569	**4,727**
2.3	CA	**1,349**	1,368	1,182	**1,155**	1,121	**1,100**
3.1	PA	1,148	**1,042**	**1,375**	1,378	**1,411**	1,420
3.2	PA	17,556	**17,015**	30,168	**28,351**	**10,221**	12,590
3.3	PA	**207,115**	224,798	**138,204**	163,189	231,661	**197,843**
4.1	PL	**10,927**	11,898	**13,410**	14,798	**15,002**	16,188
4.2	PL	29,900	**26,986**	23,135	**19,588**	33,694	**20,015**
4.3	PL	2,827	**2,476**	2,247	**2,223**	**1,661**	2,058
5.1	WC	**34,089**	34,701	**36,685**	44,033	54,253	**53,562**
5.2	WC	4,091	**3,505**	4,077	**3,716**	**3,893**	4,126
5.3	WC	16,603	**14,815**	16,158	**14,934**	**9,470**	10,062

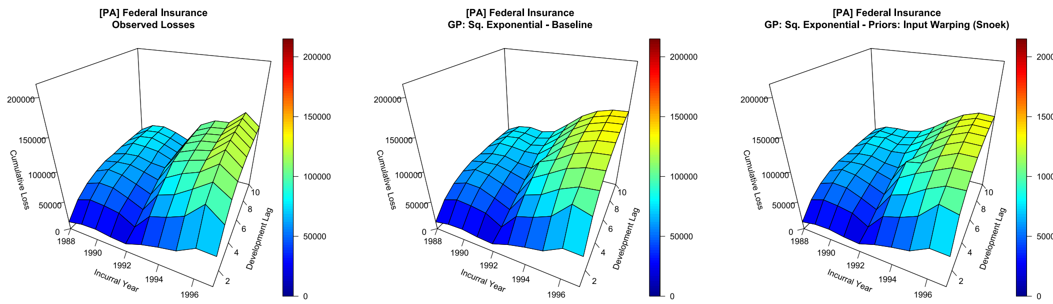
In general, we observe that the Prior distribution as used in the model by Lally and Hartman (2018) is adequate for most triangles. In the case that the setup by Snoek et al. (2014) gives better results, the difference in general is fairly minor.

The effect of these variations are visualised by $\mathbb{R}^3$-plots of Triangles 4.2 (Matérn 3/2) in Figure 6.6 and Triangle 3.2 (Sq. Exponential) in Figure 6.7. The priors by Snoek appear to be unable to model the growth witnessed in Triangle 4.2 in the Incurral Year-axis - while the trend of Triangle 2.3 (which is similar, not visualised) is captured accordingly by Snoek's prior distribution. For all other triangles, including Triangle 3.2 visualised in Figure 6.7, the general trend remains identical.



(a) Observed Losses (b) Lally, Sq. Exponential (c) Snoek, Sq. Exponential

Figure 6.6: Visualisation of models applying Snoek's Input Warping prior distributions on Triangle 4.2



(a) Observed Losses (b) Lally, Sq. Exponential (c) Snoek, Sq. Exponential

Figure 6.7: Visualisation of models applying Snoek's Input Warping prior distributions on Triangle 3.2

6.2.5 Optimal prior configuration

Given the results from the previous sections on the prior distributions, we can identify an optimal combination of priors per triangle and covariance function, based on the RMSE of the previous results. These configurations are given in Tables 6.14, 6.15 and 6.16, where variations on the model by Lally and Hartman (2018) are marked in **bold**.

In three cases, the configuration by Lally and Hartman (2018) gave the lowest RMSE for the GP model. In 15 cases, only one different prior distribution than the configuration by Lally and Hartman (2018) gave the lowest RMSE. For the 28 other combinations, we ran the model with the optimal configuration. The predicted Loss Reserves are given in Table 6.17 and the corresponding RMSE is given in Table 6.18, with the best performing method overall marked in green.

Table 6.14: Optimal prior configuration of Matérn 3/2 covariance function

Triangle	Bandwidth (ψ)	Noise (σ^2)	Signal (η^2)	Warping (α, β)
1.1	Gamma	Student-T	**Log-normal**	Lally
1.2	**Cauchy**	Student-T	**Log-normal**	Lally
1.3	Gamma	Student-T	**Log-normal**	Lally
2.1	Gamma	**Log-normal**	**Log-normal**	Lally
2.2	Gamma	Student-T	Student-T	Lally
2.3	Gamma	Student-T	**Log-normal**	**Snoek**
3.1	**Cauchy**	Student-T	Student-T	Lally
3.2	**Cauchy**	**Log-normal**	Student-T	Lally
3.3	**Cauchy**	Student-T	**Log-normal**	**Snoek**
4.1	**Cauchy**	**Log-normal**	Student-T	**Snoek**
4.2	Gamma	Student-T	**Log-normal**	Lally
4.3	Gamma	Student-T	Student-T	Lally
5.1	**Cauchy**	Student-T	Student-T	**Snoek**
5.2	**Cauchy**	Student-T	**Log-normal**	Lally
5.3	**Cauchy**	Student-T	Student-T	Lally

Table 6.15: Optimal prior configuration of Matérn 5/2 covariance function

Triangle	Bandwidth (ψ)	Noise (σ^2)	Signal (η^2)	Warping (α, β)
1.1	Gamma	Student-T	**Log-normal**	**Snoek**
1.2	**Cauchy**	Student-T	**Log-normal**	Lally
1.3	Gamma	Student-T	**Log-normal**	Lally
2.1	Gamma	Student-T	**Log-normal**	**Snoek**
2.2	**Cauchy**	Student-T	**Log-normal**	**Snoek**
2.3	Gamma	Student-T	**Log-normal**	Lally
3.1	**Cauchy**	**Log-normal**	**Log-normal**	**Snoek**
3.2	**Cauchy**	**Log-normal**	Student-T	Lally
3.3	**Cauchy**	Student-T	**Log-normal**	**Snoek**
4.1	**Cauchy**	**Log-normal**	Student-T	**Snoek**
4.2	Gamma	Student-T	**Log-normal**	Lally
4.3	Gamma	Student-T	**Log-normal**	Lally
5.1	**Cauchy**	Student-T	**Log-normal**	**Snoek**
5.2	**Cauchy**	Student-T	**Log-normal**	Lally
5.3	**Cauchy**	Student-T	Student-T	Lally

We observe that tweaking the prior distribution can have a positive impact on the model performance, especially for the Matérn covariance functions. We visualise the model predictions of Triangles 3.2 and 4.1, Squared Exponential, in Figures 6.8 and 6.9.

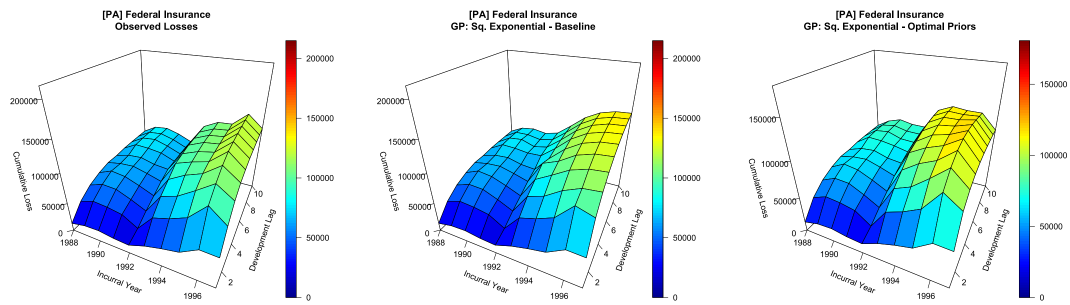
For Triangle 3.2, we see that the optimal priors are able to capture the more volatile trend, resulting in a lower RMSE and more accurate Best Estimate. In Figure 6.9, it can be observed that the optimal priors are also able to capture the more volatile trend, most notably the declined claims in Incurral Year 1997.

Table 6.16: Optimal prior configuration of Squared Exponential covariance functions

Triangle	Bandwidth (ψ)	Noise (σ^2)	Signal (η^2)	Warping (α, β)
1.1	Gamma	**Log-normal**	**Log-normal**	**Snoek**
1.2	Gamma	Student-T	**Log-normal**	Lally
1.3	Gamma	Student-T	**Log-normal**	Lally
2.1	Gamma	Student-T	Student-T	Lally
2.2	**Cauchy**	Student-T	**Log-normal**	Lally
2.3	Gamma	Student-T	**Log-normal**	Lally
3.1	**Cauchy**	**Log-normal**	Student-T	**Snoek**
3.2	**Cauchy**	**Log-normal**	Student-T	**Snoek**
3.3	**Cauchy**	Student-T	Student-T	Lally
4.1	**Cauchy**	**Log-normal**	Student-T	**Snoek**
4.2	Gamma	Student-T	**Log-normal**	Lally
4.3	**Cauchy**	Student-T	**Log-normal**	**Snoek**
5.1	**Cauchy**	**Log-normal**	Student-T	Lally
5.2	**Cauchy**	Student-T	**Log-normal**	**Snoek**
5.3	**Cauchy**	Student-T	**Log-normal**	**Snoek**

Table 6.17: Estimated Loss Reserves of models with optimal prior configuration

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	Observed	Optimal	Lally	Optimal	Lally	Optimal	Lally
1.1	407,525	**293,655**	288,484	**340,579**	313,911	**312,492**	284,687
1.2	104,094	**84,920**	83,937	74,270	**77,709**	**80,300**	77,340
1.3	164,633	**157,072**	155,306	**161,363**	156,467	**154,665**	153,394
2.1	16,640	13,330	**14,254**	**14,903**	13,789	13,891	13,891
2.2	353,949	329,846	329,846	333,871	**345,927**	338,740	**350,702**
2.3	13,046	**7,351**	6,627	**7,774**	6,993	**7,674**	6,652
3.1	37,397	**37,993**	38,901	31,861	**40,278**	28,429	**37,737**
3.2	137,642	**187,778**	280,729	**193,522**	368,464	**135,655**	212,727
3.3	11,561,327	**10,584,904**	10,113,582	10,517,807	**10,996,920**	10,447,237	**11,087,032**
4.1	10,965	**64,614**	90,421	**64,792**	103,467	**56,876**	98,422
4.2	422,513	**185,705**	179,658	**237,454**	231,733	**223,322**	212,941
4.3	37,612	**35,031**	**35,031**	**33,748**	32,638	**30,712**	28,782
5.1	45,916	**306,089**	310,663	**299,159**	334,577	347,729	**325,816**
5.2	40,225	**51,995**	55,831	**54,563**	55,688	60,724	**51,710**
5.3	307,810	**323,255**	328,945	**330,246**	343,456	**304,559**	311,081

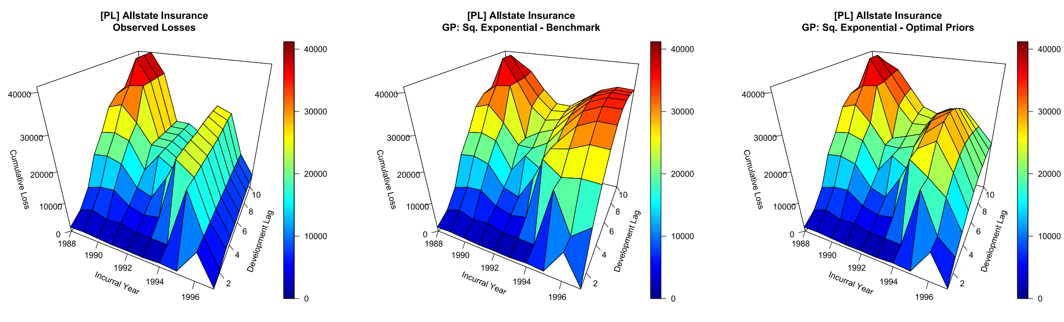


(a) Observed Losses (b) Lally, Sq. Exponential (c) Optimal, Sq. Exponential

Figure 6.8: Visualisation of models with optimal priors on Triangle 3.2

Table 6.18: RMSE of variations of models with optimal prior configuration

	Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	Optimal	Lally	Optimal	Lally	Optimal	Lally
1.1	**13,515**	13,756	**8,389**	11,015	**9,596**	13,416
1.2	**3,223**	4,039	**4,504**	4,786	**4,741**	4,904
1.3	**6,520**	6,675	**6,099**	6,153	**6,087**	6,107
2.1	**1,427**	1,472	**1,320**	1,354	1,347	1,347
2.2	5,010	5,010	5,231	**4,217**	**3,255**	4,727
2.3	**1,353**	1,368	**1,084**	1,155	**991**	1,100
3.1	**975**	1,042	**1,325**	1,378	**1,418**	1,420
3.2	**7,454**	17,015	**7,849**	28,351	**8,912**	12,590
3.3	**191,528**	224,798	**143,372**	163,189	**138,657**	197,843
4.1	**7,706**	11,898	**8,117**	14,798	**8,298**	16,188
4.2	**26,209**	26,986	**19,254**	19,588	**18,971**	20,015
4.3	2,476	2,476	**2,173**	2,223	2,259	**2,058**
5.1	**31,736**	34,701	**30,305**	44,033	**44,363**	53,562
5.2	**3,139**	3,505	**3,365**	3,716	**4,049**	4,126
5.3	**13,365**	14,815	**13,157**	14,934	**8,315**	10,062



(a) Observed Losses (b) Lally, Sq. Exponential (c) Optimal, Sq. Exponential

Figure 6.9: Visualisation of models with optimal priors on Triangle 4.1

6.3 | Premium Information

In this section, the results of adding premium information to the model are presented. The research setup is described in Section 4.3. The methods of implementing this information researched is either a transformation to Loss Ratio's, using the premium volume as input as opposed to the incurral years or by adding a Bornheutter-Ferguson estimation to the model. In all cases, we use identical prior distribution to the model as initially described by Lally and Hartman (2018). We recall the research questions that we attempt to answer:

In order to improve the Best Estimate and/or reduce the confidence interval:
- *Can relevant, out-of-triangle information be supplied to the model?*
- *Can the GP model be extended with a Bornheutter-Ferguson estimation method?*

The triangles which have a smooth variation in premium volume over time are triangles 1.1, 2.1, 2.2, 3.1, 3.3 and 5.2 . The triangles with more abrupt changes in volume are triangles 1.2, 2.3, 3.2, 4.1, 4.2, 5.1 and 5.2. Especially for the latter triangles, we intend to improve the Best Estimate predicted by the model by supplying the model with extra information - these triangles are highlighted in all results tables. For all triangles, we with to reduce the confidence interval of the loss reserve.

6.3.1 | Loss Ratio's

Prediction and Error of Best Estimate

First, we touch upon the transformation of the entire triangle to Loss Ratio's. We give the predicted Loss Reserves in Table 6.19, and the RMSE of these predictions in Table 6.20.

Table 6.19: Estimated Loss Reserves of Loss Ratio data against Cumulative Loss data

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	Observed	LR	Cum. Loss	LR	Cum. Loss	LR	Cum. Loss
1.1	407,525	**338,250**	288,484	**356,174**	313,911	**351,119**	284,687
1.2	104,094	83,921	**83,937**	**96,231**	77,709	**104,655**	77,340
1.3	164,633	**166,147**	155,306	**164,002**	156,467	**153,824**	153,394
2.1	16,640	**16,727**	14,254	**17,227**	13,789	**17,226**	13,891
2.2	353,949	**362,388**	329,846	362,530	**345,927**	**354,148**	350,702
2.3	13,046	**7,707**	6,627	**7,551**	6,993	**7,189**	6,652
3.1	37,397	44,361	**38,901**	44,721	**40,278**	42,930	**37,737**
3.2	137,642	285,633	**280,729**	**320,787**	368,464	323,791	**212,727**
3.3	11,561,327	**12,372,458**	10,113,582	13,010,468	**10,996,920**	12,327,887	**11,087,032**
4.1	10,965	**39,420**	90,421	**50,066**	103,467	**54,592**	98,422
4.2	422,513	**221,350**	179,658	**266,676**	231,733	**254,007**	212,941
4.3	37,612	**37,744**	35,031	**36,698**	32,638	28,409	**28,782**
5.1	45,916	**116,491**	310,663	**172,023**	334,577	**146,763**	325,816
5.2	40,225	**49,738**	55,831	**43,021**	55,688	**35,880**	51,710
5.3	307,810	332,075	**328,945**	**302,573**	343,456	255,733	**311,081**

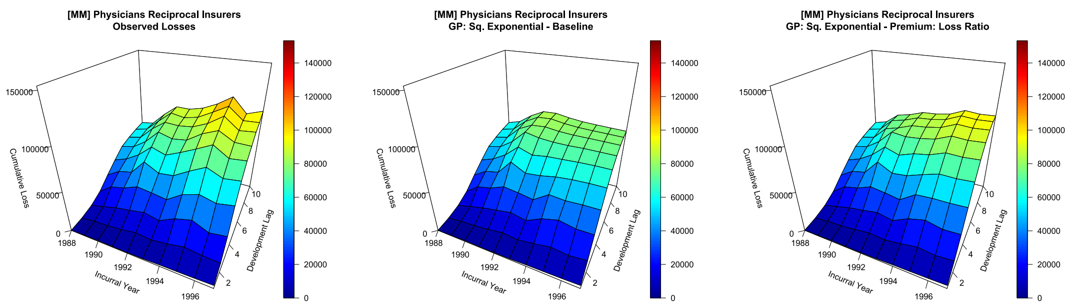
The transformation to Loss Ratio's gives a substantial improvement over the application on cumulative loss figures, most notably in the cases where the models on Cumulative Loss data fail to capture important trends. The model with the lowest RMSE does not always align with the

Table 6.20: RMSE of Loss Ratio data against Cumulative Loss data

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	LoB	LR	Cum. Loss	LR	Cum. Loss	LR	Cum. Loss
1.1	MM	**8,400**	13,756	**6,322**	11,015	**5,759**	13,416
1.2	MM	**3,346**	4,039	**4,347**	4,786	6,022	**4,904**
1.3	MM	**5,593**	6,675	**5,420**	6,153	**5,431**	6,107
2.1	CA	**1,395**	1,472	**1,295**	1,354	**1,288**	1,347
2.2	CA	**4,394**	5,010	7,789	**4,217**	7,761	**4,727**
2.3	CA	**1,047**	1,368	**1,009**	1,155	**977**	1,100
3.1	PA	1,920	**1,042**	2,169	**1,378**	2,171	**1,420**
3.2	PA	**16,585**	17,015	**21,392**	28,351	23,539	**12,590**
3.3	PA	**211,368**	224,798	365,734	**163,189**	433,816	**197,843**
4.1	PL	**5,726**	11,898	**6,929**	14,798	**7,702**	16,188
4.2	PL	**20,695**	26,986	**16,037**	19,588	**16,490**	20,015
4.3	PL	**1,701**	2,476	**1,534**	2,223	**1,675**	2,058
5.1	WC	**9,406**	34,701	**17,765**	44,033	**17,588**	53,562
5.2	WC	**3,270**	3,505	**3,244**	3,716	**3,166**	4,126
5.3	WC	**9,230**	14,815	**7,595**	14,934	**7,465**	10,062

model that has the better Loss Reserve estimation, indicating that some models might still not be able to capture all trends accordingly (e.g. Triangle 2.2, Squared Exponential covariance).

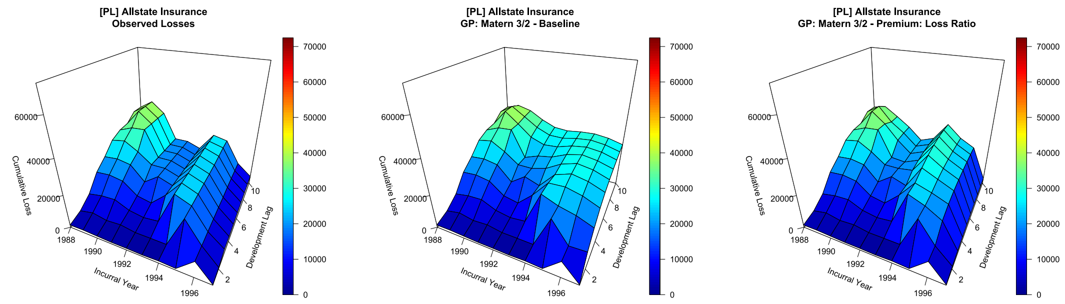
We visualise Triangles 1.1 (Squared Exponential), 4.1 (Matérn 3/2) and 1.2 (Squared Exponential) in, respectively, Figures 6.10, 6.11 and 6.12.



(a) Observed Losses (b) Cumul. Losses, Sq. Exp. (c) Loss Ratio's, Sq. Exp.

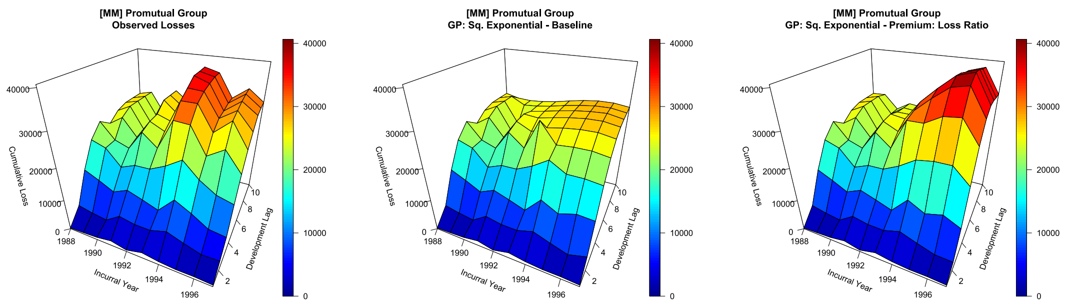
Figure 6.10: Visualisation of models applying Loss Ratio's as input on Triangle 1.1

Both Figures 6.10 and 6.11 show an improvement caused by the transformation to Loss Ratio's. Both models are able to predict the trends in the data adequately, resulting in both a lower RMSE and a more accurate Best Estimate of the Loss Reserve. However, Figure 6.12 shows that, while the Best Estimate of the loss reserve is adequate (as presented in Table 6.19), the model does not capture the relevant trends accordingly - as reflected in the RMSE.



(a) Observed Losses (b) Cumul. Losses, Matérn 3/2 (c) Loss Ratio's, Matérn 3/2

Figure 6.11: Visualisation of models applying Loss Ratio's as input on Triangle 4.1

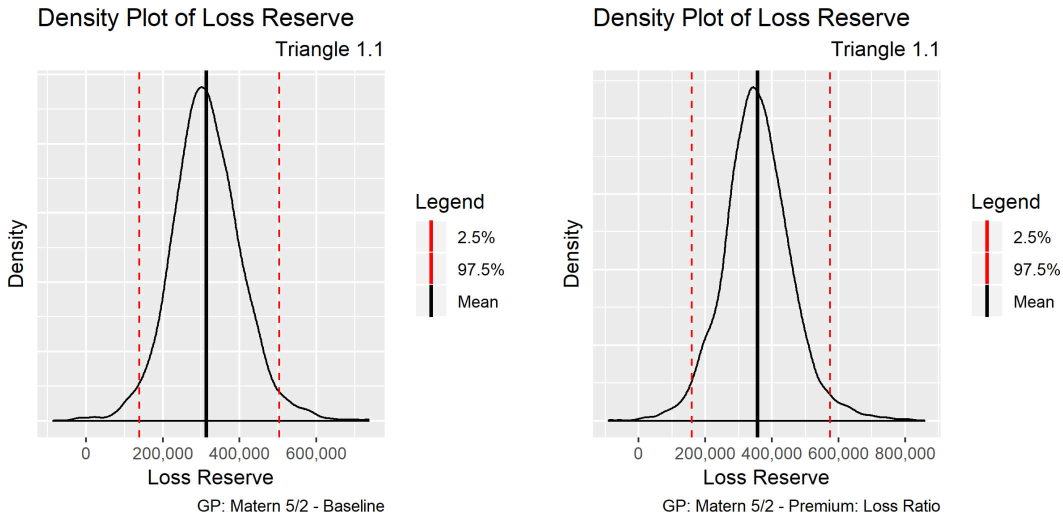


(a) Observed Losses (b) Cumul. Losses, Sq. Exp. (c) Loss Ratio's, Sq. Exp.

Figure 6.12: Visualisation of models applying Loss Ratio's as input on Triangle 1.2

Confidence of prediction

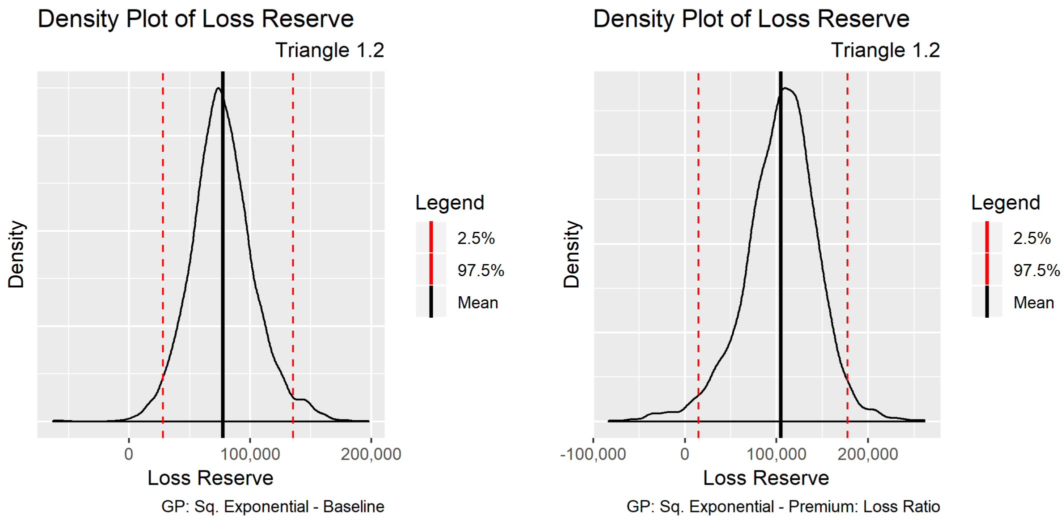
In Table 6.21, the standard deviations of the predicted Loss Reserves are given. We observe no substantial improvement in this regard when comparing the Loss Ratio implementation against the cumulative loss implementation. We visualise this by two density plots given in Figures 6.13 and 6.14 (resp. Triangle 1.1 (Matérn 5/2) and 1.2 (Sq. Exponential), where no improvement can be observed.



(a) Cum. Loss

(b) Loss Ratio

Figure 6.13: Density of Triangle 1.1 (Matérn 5/2)



(a) Cum. Loss

(b) Loss Ratio

Figure 6.14: Density of Triangle 1.2 (Sq. Exponential)

Table 6.21: Standard Deviation of Loss Ratio data against Cumulative Loss data

#	LoB	Matérn 3/2		Matérn 5/2		Sq. Exponential	
		LR	Cum. Loss	LR	Cum. Loss	LR	Cum. Loss
1.1	MM	109,023	**86,625**	102,753	**90,804**	**78,349**	83,825
1.2	MM	**30,975**	33,058	38,165	27,092	40,184	**26,187**
1.3	MM	**40,472**	41,430	38,080	36,476	41,184	**37,098**
2.1	CA	6,574	**5,458**	4,821	4,712	4,292	**4,128**
2.2	CA	98,457	**73,495**	96,002	74,844	95,817	**65,842**
2.3	CA	6,092	**3,931**	5,134	**4,061**	4,703	**4,335**
3.1	PA	9,569	**8,039**	9,400	**8,852**	10,210	**8,223**
3.2	PA	115,785	**92,948**	120,643	**118,434**	101,241	**98,847**
3.3	PA	3,729,540	**2,400,639**	3,885,112	2,616,817	3,394,312	**2,541,395**
4.1	PL	**30,759**	57,195	**30,920**	56,209	**30,956**	58,213
4.2	PL	129,409	**95,199**	158,818	**128,196**	144,230	**116,028**
4.3	PL	**19,890**	21,155	**19,281**	21,202	**16,752**	18,783
5.1	WC	**75,294**	229,973	**130,722**	226,372	**124,941**	202,139
5.2	WC	**36,321**	37,775	**34,618**	37,127	28,393	**26,968**
5.3	WC	142,339	**107,898**	125,286	**120,175**	**101,408**	104,266

6.3.2 | Premium as input

Prediction and Error of Best Estimate

In this section, we present the the RMSE and predicted Loss Reserves by supplying the net premium as input to the model, as opposed to the Incurral Year. The predicted Loss Reserves are presented in Table 6.22 and the corresponding RMSE is given in Table 6.23.

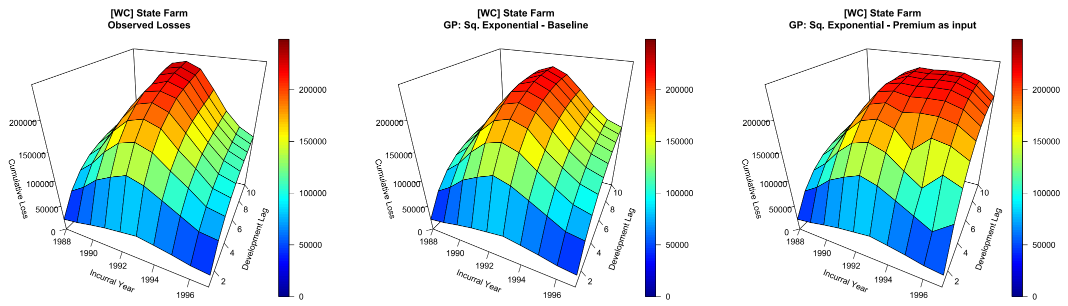
Table 6.22: Estimated Loss Reserves of applying net premium as input

#	Observed	Matérn 3/2		Matérn 5/2		Sq. Exponential	
		Premium	IY	Premium	IY	Premium	IY
1.1	407,525	255,061	**288,484**	283,618	**313,911**	**306,237**	284,687
1.2	104,094	74,258	**83,937**	75,054	**77,709**	75,283	**77,340**
1.3	164,633	70,174	**155,306**	72,442	**156,467**	61,084	**153,394**
2.1	16,640	**15,021**	14,254	**15,281**	13,789	**14,623**	13,891
2.2	353,949	**343,038**	329,846	**353,379**	345,927	360,940	**350,702**
2.3	13,046	6,032	**6,627**	6,144	**6,993**	3,966	**6,652**
3.1	37,397	31,301	**38,901**	33,049	**40,278**	31,755	**37,737**
3.2	137,642	**106,143**	280,729	**132,115**	368,464	**132,415**	212,727
3.3	11,561,327	**10,216,463**	10,113,582	10,918,190	**10,996,920**	**11,225,306**	11,087,032
4.1	10,965	**74,960**	90,421	**81,369**	103,467	**74,415**	98,422
4.2	422,513	141,330	**179,658**	172,116	**231,733**	173,763	**212,941**
4.3	37,612	58,486	**35,031**	57,983	**32,638**	55,911	**28,782**
5.1	45,916	349,603	**310,663**	364,351	**334,577**	335,710	**325,816**
5.2	40,225	210,484	**55,831**	211,814	**55,688**	210,551	**51,710**
5.3	307,810	505,340	**328,945**	536,028	**343,456**	541,134	**311,081**

We observe that this method provides no advantage over the regular model. Triangles 3.2 and 4.1 have an improvement in RMSE, but the other results show a detrimental effect, most notably the results of Triangles 5.2 and 5.3. When analysing the estimated loss reserves, we see that performance of this model is considerably worse than applying it on cumulative losses, even for the triangles with a volatile premium development. The results are visualised by plotting Triangle 5.3 (Squared Exponential) in Figure 6.15 and Triangle 3.2 (Matérn 5/2) in Figure 6.16.

Table 6.23: RMSE of applying net premium as input

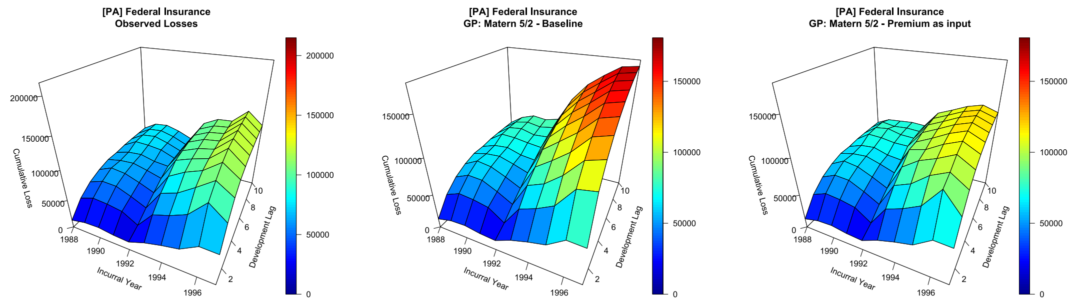
		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	LoB	Premium	Incurral Year	Premium	Incurral Year	Premium	Incurral Year
1.1	MM	15,151	**13,756**	12,230	**11,015**	**9,754**	13,416
1.2	MM	5,293	**4,039**	5,242	**4,786**	5,208	**4,904**
1.3	MM	15,428	**6,675**	15,196	**6,153**	15,388	**6,107**
2.1	CA	**1,296**	1,472	**1,301**	1,354	**1,340**	1,347
2.2	CA	**4,795**	5,010	5,198	**4,217**	5,658	**4,727**
2.3	CA	**1,160**	1,368	1,165	**1,155**	1,375	**1,100**
3.1	PA	1,551	**1,042**	1,562	**1,378**	1,620	**1,420**
3.2	PA	**8,970**	17,015	**6,341**	28,351	**5,793**	12,590
3.3	PA	**215,809**	224,798	163,488	**163,189**	**179,581**	197,843
4.1	PL	**7,547**	11,898	**7,991**	14,798	**7,397**	16,188
4.2	PL	29,154	**26,986**	24,706	**19,588**	24,623	**20,015**
4.3	PL	5,561	**2,476**	5,545	**2,223**	5,523	**2,058**
5.1	WC	36,324	**34,701**	**43,624**	44,033	**50,766**	53,562
5.2	WC	22,485	**3,505**	22,518	**3,716**	22,385	**4,126**
5.3	WC	54,862	**14,815**	55,407	**14,934**	55,580	**10,062**



(a) Observed Losses (b) Incurral Year, Sq. Exponential (c) Net Premium, Sq. Exponential

Figure 6.15: Visualisation of models having supplied premium as input on Triangle 5.3

In Figure 6.15, we see a clear underperformance of the model. Due to the changed input, the model is unable to grasp the general trend in the model, resulting in detrimental results. On the other hand, we see an improvement of the model in Figure 6.16. While the regular model gives remarkable results (the Matérn 3/2 and Sq. Exponential covariance functions perform better for this triangle), an improvement can be seen. For this triangle, the amount of net premium received appears to be an adequate indicator for the size of the loss reserve.



(a) Observed Losses (b) Incurral Year, Matérn 5/2 (c) Net Premium, Matérn 5/2

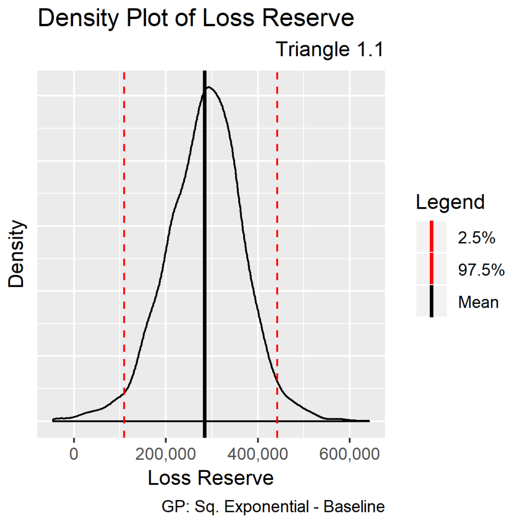
Figure 6.16: Visualisation of models having supplied premium as input on Triangle 3.2

Confidence of Prediction

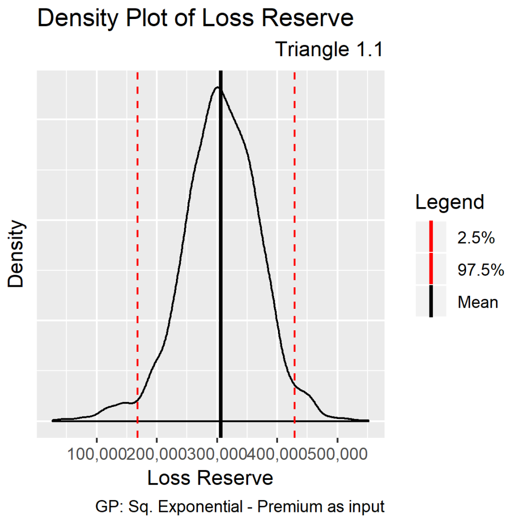
Concluding, we present the results of the standard deviation of the loss reserve in Table 6.24. While there is no improvement of Best Estimate predictions, the volatility in the predictions appears to have declined. We visualise this in Figures 6.17 and 6.18, where density plots are given of Triangles 1.1 and 4.3 (Both Squared Exponential). We can see a decrease of the Confidence Interval of the prediction, even though the predictions (as described earlier) are generally of worse quality.

Table 6.24: Standard Deviation of Loss Reserve Predictions: applying net premium as input

	Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	Premium	Incurral Year	Premium	Incurral Year	Premium	Incurral Year
1.1	**76,069**	86,625	**66,152**	90,804	**64,015**	83,825
1.2	**27,077**	33,058	**23,145**	27,092	**22,218**	26,187
1.3	46,582	**41,430**	42,858	**36,476**	45,505	**37,098**
2.1	5,469	**5,458**	5,007	**4,712**	4,328	**4,128**
2.2	**72,223**	73,495	79,675	**74,844**	75,385	**65,842**
2.3	**3,576**	3,931	**3,809**	4,061	**3,883**	4,335
3.1	9,058	**8,039**	8,860	**8,852**	10,855	**8,223**
3.2	**84,720**	92,948	**77,845**	118,434	**82,931**	98,847
3.3	**2,256,542**	2,400,639	**2,265,738**	2,616,817	2,271,498	**2,541,395**
4.1	**48,327**	57,195	**48,411**	56,209	**48,396**	58,213
4.2	**91,648**	95,199	**94,520**	128,196	**106,841**	116,028
4.3	**11,370**	21,155	**10,196**	21,202	**10,029**	18,783
5.1	237,872	**229,973**	244,785	**226,372**	**191,188**	202,139
5.2	63,896	**37,775**	58,815	**37,127**	58,343	**26,968**
5.3	117,000	**107,898**	**108,588**	120,175	112,918	**104,266**

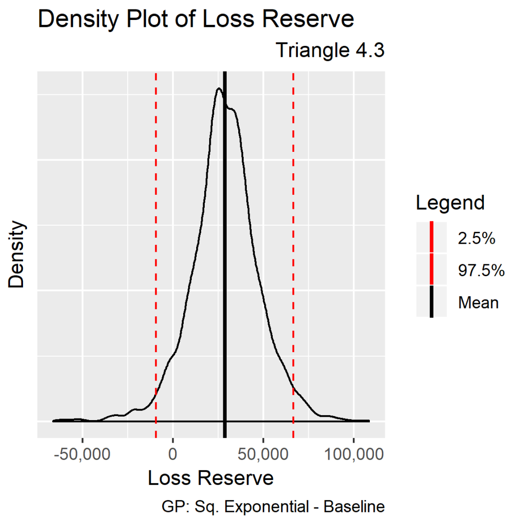


(a) Cum. Loss

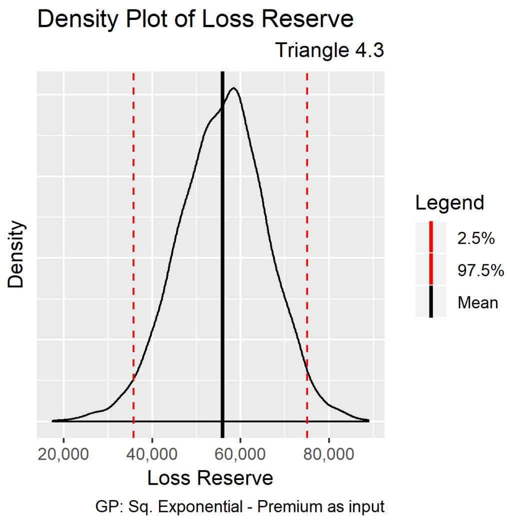


(b) Premium

Figure 6.17: Density of Triangle 1.1 (Sq. Exponential)



(a) Cum. Loss



(b) Premium

Figure 6.18: Density of Triangle 4.3 (Sq. Exponential)

6.3.3 | Bornheutter-Ferguson Estimations

In this section, we attempt to both improve the predictions of triangles by supplying information on the net premium, and get a more accurate prediction of the Loss Reserve and reduce the uncertainty of that prediction.

The estimations made by the Bornheutter-Ferguson method directly result in the Best Estimate for the Loss Reserve, as these estimators are supplied to the model as if they are noisy observations at the final development lag. The GP model is able to estimate the amount of noise in these observations. While the model is thus able to estimate the uncertainty of the Loss Reserve by estimating the noise parameter, the Best Estimate is always equal to the supplied estimators. As such, we do not compare models on their Loss Reserve predictions, but solely on error statistics.

The BF-estimators, compared to their actual ultimate loss, are given in Appendix F.

Error of prediction

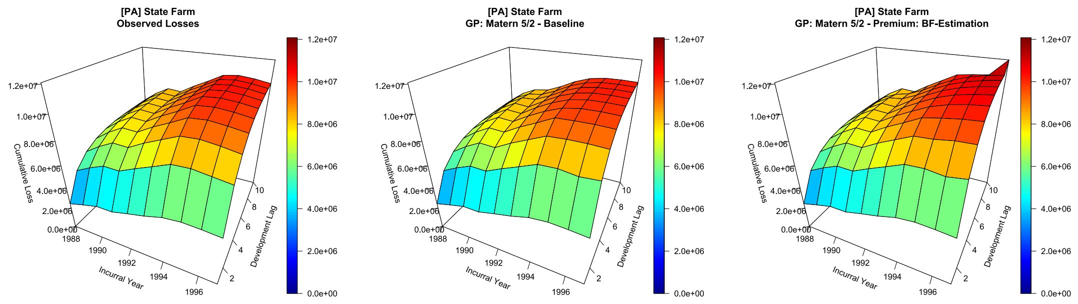
First, we analyse the RMSE of the predictions, given in Table 6.25. It should be noted that this is the RMSE of all missing observations as predicted by the Gaussian process model (i.e. not including Development Lag 10). The BF-observations, as generated by a seperate process, have not been taken in account. For equal comparison, we give the benchmark RMSE for the identical number of estimations (i.e. also excluding DL=10).

Table 6.25: RMSE of applying Bornheutter-Ferguson estimation

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	LoB	BF	Cumul. Loss	BF	Cumul. Loss	BF	Cumul. Loss
1.1	MM	**6,206**	12,738	**5,619**	10,114	**5,628**	12,108
1.2	MM	4,350	**4,019**	4,827	**4,709**	4,996	**4,834**
1.3	MM	8,239	**6,954**	6,652	**6,371**	6,306	**6,287**
2.1	CA	**1,199**	1,527	**1,193**	1,389	**1,205**	1,382
2.2	CA	9,630	**4,927**	10,051	**4,205**	9,985	**4,824**
2.3	CA	**1,025**	1,415	**977**	1174	**938**	1,102
3.1	PA	1,494	**1,063**	1,606	**1,413**	1,649	**1,459**
3.2	PA	25,380	**16,442**	**25,980**	27,359	24,343	**12,543**
3.3	PA	513,303	**220,553**	495,037	**165,984**	597,705	**207,054**
4.1	PL	12,240	**11,988**	**14,568**	15,090	**16,193**	16,757
4.2	PL	**22,421**	25,111	19,725	**17,580**	18,449	**17,512**
4.3	PL	2,532	**2,496**	2,612	**2,238**	2,727	**2,061**
5.1	WC	**14,788**	33,806	**15,684**	44,063	**21,630**	54,903
5.2	WC	**2,904**	3,533	**3,171**	3,753	**3,986**	4,193
5.3	WC	22,230	**14,378**	20,501	**14,505**	18,305	**9,713**

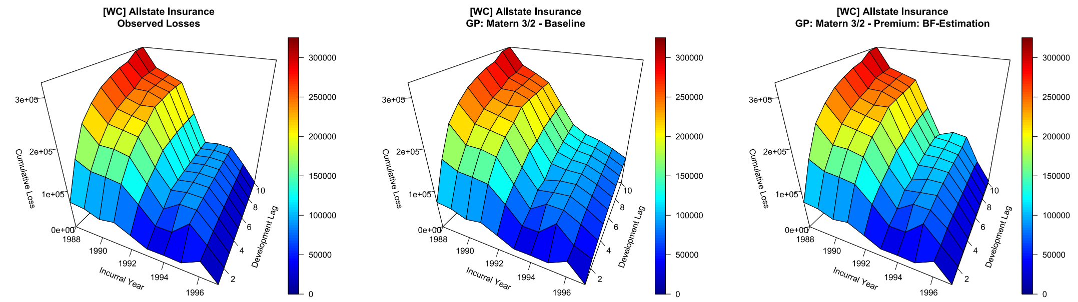
Table 6.25 gives no conclusive preference to any method. The RMSE of the predictions varies between triangles. While the model is able to capture that there is noise in the estimation, this does not necessarily result in better predictions.

The quality of the estimator might have an influence on the results of the interpolation. However, when we zoom in on Triangle 3.3 and its' estimators, we see that even while the estimators are fairly adequate, the errors of the model are significantly increased. This triangle is visualised in Figure 6.19, with Matérn 5/2 as covariance function. Triangle 5.1 (Matérn 3/2) is visualised in Figure 6.20, which has a better performance than the regular model, most notably caused by the estimator of Incurral Year 1997, which is giving the model adequate guidance.



(a) Observed Losses (b) Regular model, Matérn 5/2 (c) BF-Estimation, Matérn 5/2

Figure 6.19: Visualisation of models with a Bornheutter-Ferguson estimator, applied on Triangle 3.3



(a) Observed Losses (b) Regular model, Matérn 3/2 (c) BF-Estimation, Matérn 3/2

Figure 6.20: Visualisation of models with a Bornheutter-Ferguson estimator, applied on Triangle 5.1

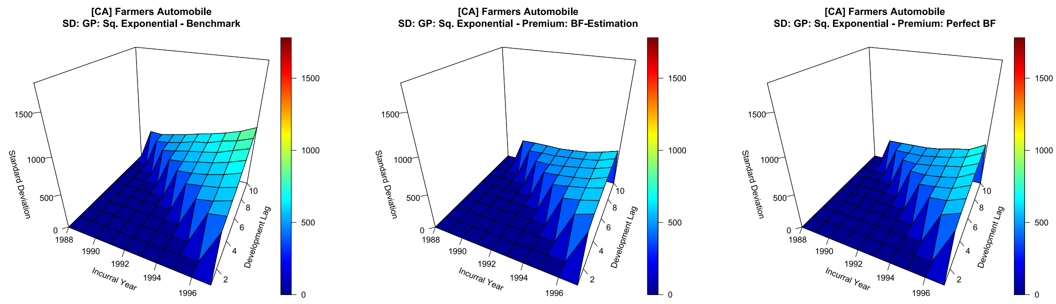
Perfect BF-estimators

In order to quantify the influence of the quality of the estimators, we have also performed a model run by supplying the model with perfect estimations, i.e. supplying it with the actual observed losses at DL $= 10$. The primary goal of this is to rule out the effect of the quality of the estimators, and solely assessing the model performance. The RMSE of the estimations is given in Table 6.26. As expected, providing the model with perfect information drastically reduces the errors of the model.

In contrast to earlier results, we are unable to give a density or confidence interval of the loss reserve due to the implementation that we have performed. However, we can visualise the standard deviation per predicted point, based on the 4,000 MCMC samples for each prediction. This is visualised in Figures 6.21, 6.22 and 6.23. An improvement of all BF models over the Regular GP model can be seen. For most triangles, the perfect estimators reduce the spread. However, for Triangle 2.1 (Figure 6.21), this is not the case.

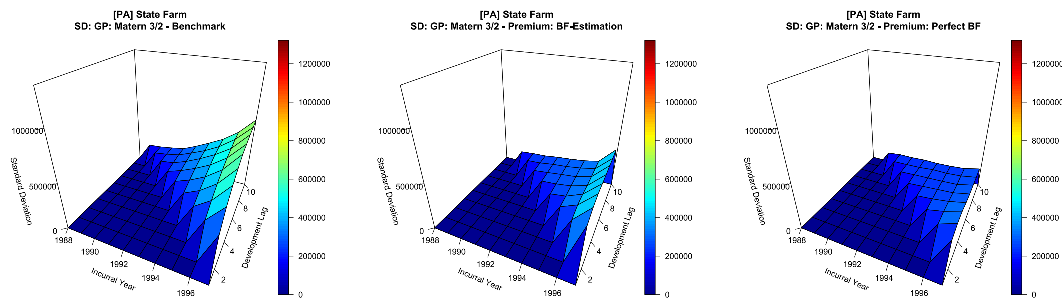
Table 6.26: RMSE of "Perfect" Bornheutter-Ferguson estimation

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
#	LoB	Perfect	BF	Perfect	BF	Perfect	BF
1.1	MM	**5,304**	6,206	**5,206**	5,619	**5,358**	5,628
1.2	MM	**3,123**	4,350	**4,096**	4,827	**4,237**	4,996
1.3	MM	**5,951**	8,239	**6,082**	6,652	**6,200**	6,306
2.1	CA	1,203	**1,199**	1,249	**1,193**	1,247	**1,205**
2.2	CA	**3,127**	9,630	**3,853**	10,051	**4,406**	9,985
2.3	CA	**772**	1,025	**772**	977	**779**	938
3.1	PA	**733**	1,494	**1,055**	1,606	**1,260**	1,649
3.2	PA	**6,045**	25,380	**7,824**	25,980	**8,441**	24,343
3.3	PA	**139,165**	513,303	**151,615**	495,037	**182,248**	597,705
4.1	PL	**4,728**	12,240	**7,782**	14,568	**9,596**	16,193
4.2	PL	**9,778**	22,421	**9,935**	19,725	**10,708**	18,449
4.3	PL	**1,517**	2,532	**1,664**	2,612	**1,680**	2,727
5.1	WC	**6,196**	14,788	**8,037**	15,684	22,860	**21,630**
5.2	WC	**2,549**	2,904	**2,780**	3,171	**3,502**	3,986
5.3	WC	**2,846**	22,230	**2,734**	20,501	**2,617**	18,305



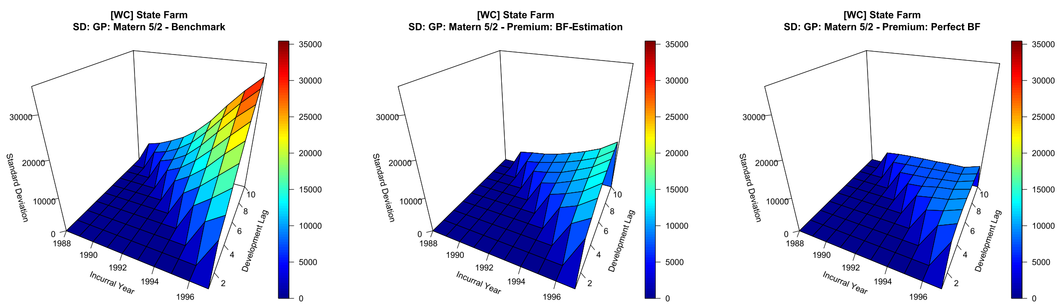
(a) Regular GP (b) BF Estimation (c) Perfect Estimation

Figure 6.21: Visualisation of Standard Deviations of Triangle 2.1, Sq. Exponential



(a) Regular GP (b) BF Estimation (c) Perfect Estimation

Figure 6.22: Visualisation of Standard Deviations of Triangle 3.3, Matérn 3/2



(a) Regular GP (b) BF Estimation (c) Perfect Estimation

Figure 6.23: Visualisation of Standard Deviations of Triangle 5.3, Matérn 5/2

7 | Conclusion

Contents

7.1 Data applicability 75
7.2 Prior distributions and model design 75
7.3 Adding premium information to the model 76

In this chapter, we will draw conclusions from the results as presented in Chapter 6. Our research questions are:

Can we improve the Hierarchical Gaussian Process model
to predict the loss reserve of a non-life insurer?

With the following sub-questions:

- In order to validate the design choices:
  - Is the model applicable on a more extensive data set?
  - Are the prior distributions on the hyperparameters adequately chosen?
- In order to improve the Best Estimate and/or reduce the confidence interval:
  - Can relevant, out-of-triangle information be supplied to the model?
  - Can the GP model be extended with a Bornheutter-Ferguson estimation method?

7.1 | Data applicability

The model performance has been tested on a more extensive data set in order to validate applicability of the model to multiple characteristics. From this, we conclude that the Gaussian Process model is able to make comparable or better predictions than the current industry practice. However, all models have problems when making estimations on a data set that is less predictable: such as having a diagonal shock, an increase in received premiums or other influences that cause a volatile claims development.

7.2 | Prior distributions and model design

We have applied a multitude of prior distributions in order to validate design choices made in the model. From this analysis, we conclude that distributions and/or covariance functions that perform best are dependant on the triangle. For the entire model, we can not conclude a preferred setting that performs best, regardless of the data.

The signal and noise parameters and their prior distributions have minimal impact on the model performance: while the model should be able to have the freedom to infer the adequate values for these parameters, most weakly informative priors will give adequate results. Furthermore, considering that the noise parameter tends to converge to zero for nearly all models of this dataset, one can consider fixing the noise to zero - but only if the data is of sufficient quality. The prior distribution on the Input Warping parameters have an impact on model performance of some triangles - but the priors by Snoek et al. (2014) sometimes fail to capture important trends.

We have observed that the prior distribution on the bandwidth parameter has the most influence on model performance. The Cauchy distribution (constrained to be positive) allows for more extreme values, and thus offers more flexibility to the model. For some triangles (with a more volatile development) this results in better predictions, while the increased flexibility sometimes converges to less realistic values - once even predicting a negative loss reserve. Especially in the case of a combination with the Squared Exponential covariance function, this results in too much freedom of the model. By applying the Cauchy-distribution, the model tends to consider realistic values as outliers, and thus fails to capture a trend in the data. As such, we do not recommend applying the Cauchy-distribution for the bandwidth parameter.

Concluding, while the prior distributions should be adequately chosen, weakly informative priors seem to give adequate results.

7.3 | Adding premium information to the model

Supplying premium information to the model is researched. Based on our results, we conclude that a transformation of the model to Loss Ratio's appear to give the most adequate results. Technically speaking, this is not *adding* premium information to the model, but rather pre-transforming the data to remove premium influences out of the supplied data. With some exceptions, this transformation gives good results for nearly all triangles, for both the Best Estimate and the RMSE.

Supplying the premiums as input does not generally result in better predictions of the loss reserve. The standard deviation of the predictions (and thus, the uncertainty) has dropped in the majority of triangles. However, given that both the Best Estimates and error statistics of the loss reserves have worsened, we do not recommend this implementation.

Finally, we have added Bornheutter-Ferguson estimators to the model. While the model is able to correct for noisy estimators, the prediction of the loss reserve is (with our implementation) heavily depending on the actual estimators. The quality of the estimators is important for the model performance: better estimators naturally result in better predictions and lower errors.

For improving the Best Estimate predictions, the Loss Ratio transformation produces better results. However, this transformation does not result in a substantially narrower confidence interval. For this, adding estimators on the most distant Development Lag has proper potential. Validation on supplying the actual observed values (naturally) produces good results. Our basic Bornheutter-Ferguson implementation sometimes produces adequate results, but the model is heavily dependant on the quality of these estimators. Therefore, estimations should be done carefully - and we suggest a different method than a draw from a random variable to perform this estimation.

We conclude that a transformation to Loss Ratio's improves model performance with regard to the Best Estimate. Adding information to the model in the form of estimators reduce the density of predictions, which could result in more realistic confidence intervals.

8 Discussion and Recommendations

Contents

8.1 **Data set** **77**

8.2 **Priors on Loss Ratio's** **78**

8.3 **Bornheutter-Ferguson Estimation** **78**

8.4 **Implementation in practice** **78**

8.5 **Technical possibilities and limitations** **79**

In this chapter, we will discuss the results generated by our implemented models and limitations that we have identified. Furthermore, we will give recommendations for future research.

8.1 Data set

The advantages of the data set used are that it is publicly available, and that the run-off of these triangles is known. However, some of these triangles are highly volatile. In this research, we have applied the model to 15 triangles with different characteristics. Some of these characteristics are known without knowledge of the lower triangle: such as increased premiums, or a diagonal shock in the upper triangle.

The model could thus be explicitly tuned to each triangle. For instance, when it is known beforehand that a diagonal shock has occurred for business reasons, we could modify (for instance) the kernel function to capture this trend, for instance by adding a periodical kernel function. A Gaussian process model can be tuned explicitly to the data and corresponding expectations by modifications of the kernel functions - which is a recommendation for future research.

A disadvantage of this data set is that it contains historical data of 1988-2006. It is possible that this data set is less representative for the current actuarial practice. For instance, digitalisation of claims handling could have sped up the process. Applying the model on a more recent data set is therefore recommended.

Furthermore, this research has been performed on either cumulative losses or cumulative Loss Ratio's of paid triangles. It is possible that applying this model on incremental triangles yields different and/or better results. Furthermore, an application on Incurred triangles could be considered, as incurred claims develop differently than paid claims. We recommend both the application of incremental losses, loss ratio's and incurred triangles (both cumulative and incremental) for further research.

In this research, we have attempted to add premium information to the model as it was the only out-of-triangle data available in the data set. In literature on estimating loss reserves, different external data is considered, such as the amount of outstanding claims at a specific moment in time, the amount of written policies, or for instance by combining triangles amongst the same insurer or different insurers, but identical Lines of Business. Subsequent research could focus on identifying different out-of-triangle data that can be taken into account.

8.2 | Priors on Loss Ratio's

We have researched transformation of cumulative losses to cumulative loss ratio's, in order to implement premium information. For this, we have used identical priors as the model on cumulative losses. It could very well be that different priors give better results, such as adapting the Input Warping prior distributions in the Incurral Year-axis; as this need not be exponential.

8.3 | Bornheutter-Ferguson Estimation

The Bornheutter-Ferguson estimation applied takes the volatility into account of the entire triangle for the random variable used for the unknown portion of the incurral year. This method sometimes results in unrealistic estimators, which could be fine-tuned based on, for instance, expect judgement. Also, different methods are be available to produce estimators that have not been considered in this research, such as the Cape Cod method - which also applies a volume measure (Radtke et al., 2016).

Even though our Gaussian process model is able to cope with noisy observations, more accurate estimations result in lower errors and as such should be determined adequately, depending on triangle characteristics.

The technical implementation of these estimators is that they are modeled as seperate observations at the final Development Lag. This implementation is partially hindered by the Input Warping methodology, as this requires all input to be normalised on the interval $[0, 1]$. It is possible that a different implementation yields better results.

8.4 | Implementation in practice

As indicated in our Results and Conclusion, drawing general conclusions about model design is difficult. There are a multitude of variations available, and depending on characteristics in the data, a Gaussian process model can be tuned to give the best results, either by means of adapting the kernel function or the hyperparameters. Tuning the model to the data (and application) should be done with adequate care. This introduces Expert Judgement to the implementation, which comes with new risks. In our research, we were able to tune the model to the actual observed losses and their error statistics, but this is -naturally- not possible in practice. Therefore, identifying the best configuration for the model is difficult, as tuning the model is required for some triangles to get adequate results.

From a regulatory point of view, the model tends to become more and more of a 'black box', especially if it is to be tuned even further to specific triangles. The current implementation with Markov Chain Monte Carlo sampling to infer parameters, and applying Input Warping already requires adequate expertise in this field, which introduces new risks on a governance level. Considering that a lot of Expert Judgement will be required in tuning the model, a single, human point of failure is introduced.

Also, our results still show room for improvement with regards to the limitation of the model that we intended to research: narrowing the confidence interval of the predicted loss reserves. From a regulatory point of view, the uncertainty of these predictions could be a huge barrier.

8.5 | Technical possibilities and limitations

The kernel function is one of the most important design choices that can be made in a Gaussian process model. As such, future research could work on identifying different kernel functions based on triangle characteristics. Valid kernel functions can be combined to a new kernel function by either multiplication or addition. For example, our currently analysed covariance functions can be combined with a periodical kernel in order to model periodically inflated claims in the months of December and January of an *supplementary healthcare* insurance - which is an observed phenomenon in practice

Moreover, we have modeled identical covariances along both the Development Lag-axis and Incurral Year-axis. This need not be the case, for instance when along the Incurral Year-axis a more rough pattern is witnessed than along the Development Lag-axis, where the former could then be modelled with a Matérn 3/2 covariance function and the latter with a Squared Exponential covariance function. We were, however, unable to get the sampler to accept such a construction, and thus recommend this for further research.

Finally, the model takes between 4 and 20 minutes to generate results of one model run, depending on the complexity of the triangle. The Computational Complexity of any Gaussian process is $\mathcal{O}(N^3)$ (Barber, 2012), and as such larger dataset become increasingly complicated. In this case, we have only considered yearly run-off triangles consisting of 10 years. If one were to model on monthly run-off triangles over a period of, e.g., 3 years, the time complexity of this would render the model unusable. An approximation method of the Gaussian process is described by Flaxman et al. (2016), but it would require further validation on the application on predicting loss reserves.

References

Antonio, K., & Plat, R. (2014). Micro-level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*, *2014*(7), 649-669. Retrieved from https://doi.org/10.1080/03461238.2012.755938 doi: 10.1080/03461238.2012.755938

Barber, D. (2012). *Bayesian Reasoning and Machine Learning.* Cambridge University Press.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637–654. Retrieved from http://www.jstor.org/stable/1831029

Bonilla, E. V., Chai, K. M., & Williams, C. (2008). Multi-task Gaussian Process Prediction. In J. C. Platt, D. Koller, Y. Singer, & S. T. Roweis (Eds.), *Advances in Neural Information Processing Systems 20* (pp. 153–160). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/3189-multi-task-gaussian-process-prediction.pdf

Bornhuetter, R., & Ferguson, R. (1972). The actuary and IBNR. In *Proceedings of the Casualty Actuarial Society.*

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software, Articles*, *76*(1), 1–32. Retrieved from https://www.jstatsoft.org/v076/i01 doi: 10.18637/jss.v076.i01

Côté, M.-P., Genest, C., & Abdallah, A. (2016, Dec 01). Rank-based methods for modeling dependence between loss triangles. *European Actuarial Journal*, *6*(2), 377–408. Retrieved from https://doi.org/10.1007/s13385-016-0134-y doi: 10.1007/s13385-016-0134-y

Diggle, P. J., Tawn, J. A., & Moyeed, R. A. (1998). Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *47*(3), 299-350. Retrieved from https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9876.00113 doi: 10.1111/1467-9876.00113

Flaxman, S., Gelman, A., Neill, D., Smola, A., Vehtari, A., & Wilson, A. G. (2016). *Fast hierarchical Gaussian processes.* (Unpublished manuscript)

Frees, E. W. (2009). *Regression Modeling with Actuarial and Financial Applications.* Cambridge University Press. doi: 10.1017/CBO9780511814372

Gelfand, A. E., & Schliep, E. M. (2016). Spatial statistics and gaussian processes: A beautiful marriage. *Spatial Statistics*, *18*, 86 - 104. Retrieved from http://www.sciencedirect.com/science/article/pii/S2211675316300033 (Spatial Statistics Avignon: Emerging Patterns) doi: https://doi.org/10.1016/j.spasta.2016.03.006

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*, 1593-1623. Retrieved from http://jmlr.org/papers/v15/hoffman14a.html

Kaas, R., Goovaerts, M., Dhaene, J., & Denuit, M. (2008). *Modern Actuarial Risk Theory: Using R.*

Kuang, D., Nielsen, B., & Perch Nielsen, J. (2011). Forecasting in an Extended Chain-Ladder-Type Model. *Journal of Risk and Insurance*, *78*(2), 345-359. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6975.2010.01395.x` doi: 10.1111/j.1539-6975.2010.01395.x

Lally, N., & Hartman, B. (2018). Estimating loss reserves using hierarchical Bayesian Gaussian process regression with input warping. *Insurance: Mathematics and Economics*, *82*, 124 - 140. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0167668717304523` doi: https://doi.org/10.1016/j.insmatheco.2018.06.008

Maciak, M., Okhrin, O., & Pešta, M. (2018, 01 08). *Dynamic and granular loss reserving with copulae.*

Mack, T. (1993, 11). Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *Astin Bulletin - ASTIN BULL*, *23*. doi: 10.2143/AST.23.2.2005092

Mack, T. (1999). The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor. *ASTIN Bulletin*, *29*(2), 361–366. doi: 10.2143/AST.29.2.504622

Martinez-Miranda, M. D., Nielsen, J., & Verrall, R. (2012, 05). Double Chain Ladder. *ASTIN Bulletin*, *42*. doi: 10.2143/AST.42.1.2160712

Meyers, G. G., & Shi, P. (2011). *Loss reserves.* `https://www.casact.org/research/index.cfm?fa=loss\_reserves\_data`. ([Online])

Polson, N. G., & Scott, J. G. (2012, 12). On the half-cauchy prior for a global scale parameter. *Bayesian Anal.*, *7*(4), 887–902. Retrieved from `https://doi.org/10.1214/12-BA730` doi: 10.1214/12-BA730

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Radtke, M., Schmidt, K., & Schnaus, A. (2016). *Handbook on loss reserving.* doi: 10.1007/978-3-319-30056-6

Rasmussen, C., & Williams, C. (2006). *Gaussian Processes for Machine Learning.* the MIT Press.

Shi, P. (2017). A multivariate analysis of intercompany loss triangles. *Journal of Risk and Insurance*, *84*(2), 717-737. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1111/jori.12102` doi: 10.1111/jori.12102

Shi, P., & Frees, E. W. (2011). Dependent loss reserving using copulas. *ASTIN Bulletin*, *41*(2), 449–486. doi: 10.2143/AST.41.2.2136985

Snoek, J., Swersky, K., S. Zemel, R., & P. Adams, R. (2014, 02). Input Warping for Bayesian Optimization of Non-stationary Functions. *31st International Conference on Machine Learning, ICML 2014*, *5*.

Wüthrich, M. V. (2018). Machine learning in individual claims reserving. *Scandinavian Actuarial Journal*, *2018*(6), 465-480. Retrieved from `https://doi.org/10.1080/03461238.2018.1428681` doi: 10.1080/03461238.2018.1428681

Wüthrich, M. V., & Merz, M. (2006). *Stochastic claims reserving methods in non-life insurance.* (Retrieved from `https://www.actuaries.ch/fr/downloads/aid!b4ae4834-66cd-464b-bd27-1497194efc96/id!98/Vorlesung\_SLR.pdf`)

Zhang, Y. (2010). A general multivariate chain ladder model. *Insurance: Mathematics and*

Economics, *46*(3), 588 - 599. Retrieved from http://www.sciencedirect.com/science/article/pii/S0167668710000302 doi: https://doi.org/10.1016/j.insmatheco.2010.03.002

A | Glossary of abbreviations

Below is a table which includes the abbreviations that are used in this thesis.

Abbreviation	Meaning
BE	Best Estimate
BF	Bornheutter-Ferguson
CDF	Cumulative Distribution Function
CA	Commercial Automobile
CL	Chain Ladder
FRM	Financial Risk Management
GP	Gaussian Process
IBN(E)R	Incurred but not (enough) reported
LoB	Line of Business
MAE	Mean Absolute Error
MCMC	Markov Chain Monte Carlo
MM	Medical Malpractice
NAIC	National Association of Insurance Commissioners
P&C	Property & Casualty
PA	Personal Automobile
PDF	Probability Density Function
PL	Product Liability
PSD	Positive Semidefinite
RBNS	Reported but not settled
RMSE	Root Mean Square Error
WC	Workers' Compensation

B | Run-off triangle: Example

Example B.1: Run-off triangle

An insurance company has been writing policies in 2016, 2017 and 2018.
It has received the following claims:

- Two claims in calendar year 2017:
  - One originating from a policy written in 2016 of €1.000
  - One originating from a policy written in 2017 of €500
- Two claims in calendar year 2018:
  - One originating from a policy written in 2016 of €250
  - One originating from a policy written in 2018 of €2.000

We can summarise this in the following incremental run-off triangle:

	Development lag		
Incurral year	1	2	3
2016	€0	€1.000	€250
2017	€500	€0	·
2018	€2.000	·	·

Fields that are marked with a · are unknown, as they refer to years that have not yet passed/finished. This can also be presented in a cumulative triangle, which is commonly done:

	Development lag		
Incurral year	1	2	3
2016	€0	€1.000	€1.250
2017	€500	€500	·
2018	€2.000	·	·

C | Run-off triangles of the data set

In this appendix, the triangles that are used for training and testing the model are presented. We have selected 5 different Lines of Business, and have selected 3 insurers (and thus: triangles) from each Line of Business. All triangles are extracted from the data set made available by the National Association of Insurance Commissioners on the website of the Casualty Actuarial Society(Meyers & Shi, 2011). All insurance companies are located in the United States of America, and the selected triangles are paid claims (as opposed to incurred claims).

Of each Line of Business, we have selected triangles with non-zero claims for all data points and have attempted to select different characteristics that test the accuracy of the model. The Lines of Business and the selected insurers are discussed in more detail in Chapter 5. Also, 3D plots of both the upper triangle and the observed losses are presented in that section.

Table C.1: Triangle 1.1: Medical Malpractice: *Physicians' Reciprocal Insurers*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	258	2,546	6,709	14,322	26,353	37,519	47,651	52,359	55,997	58,703	73,259
1989	548	4,645	11,259	26,348	36,464	43,677	51,381	55,971	60,119	*61,272*	75,519
1990	1,113	5,540	15,937	30,230	40,287	53,564	62,345	71,358	*74,387*	*76,947*	85,661
1991	976	7,609	19,104	37,548	58,428	67,809	78,260	*83,123*	*88,195*	*89,983*	86,797
1992	914	10,504	22,641	35,314	48,978	61,901	*73,525*	*82,150*	*85,302*	*89,784*	91,918
1993	1,242	8,605	18,760	32,304	50,990	*69,471*	*75,750*	*84,168*	*89,044*	*94,038*	93,210
1994	2,087	11,566	28,145	44,708	*61,536*	*77,778*	*87,094*	*91,609*	*98,291*	*102,083*	97,266
1995	2,596	15,296	33,315	*50,784*	*70,978*	*86,742*	*94,441*	*103,330*	*108,103*	*113,189*	106,992
1996	3,196	12,987	*24,692*	*46,068*	*63,115*	*75,547*	*84,909*	*90,259*	*94,560*	*95,935*	107,264
1997	4,039	*11,910*	*27,631*	*49,559*	*65,411*	*77,425*	*84,494*	*92,102*	*95,511*	*101,971*	110,832

Table C.2: Triangle 1.2: Medical Malpractice: *Promutual Group*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	292	4,697	13,938	21,175	24,834	22,360	24,779	25,792	25,891	25,942	53,178
1989	524	3,666	11,778	19,663	19,941	25,068	26,915	26,971	27,022	*27,039*	60,809
1990	829	2,943	9,327	14,676	18,355	19,650	19,820	20,199	*20,312*	*20,374*	51,806
1991	1,658	6,121	11,419	16,512	24,338	25,625	26,041	*28,735*	*28,772*	*28,805*	44,101
1992	518	5,166	9,775	16,649	20,587	23,465	*24,508*	*24,979*	*25,453*	*25,531*	29,242
1993	1,850	5,076	7,974	18,872	29,528	*33,034*	*35,005*	*37,073*	*37,144*	*37,326*	26,379
1994	1,064	6,647	13,035	21,891	*29,199*	*33,047*	*34,699*	*35,500*	*36,097*	*36,154*	25,343
1995	703	4,803	10,388	*17,420*	*21,629*	*26,032*	*28,087*	*28,467*	*29,132*	*29,147*	26,396
1996	796	3,143	*7,196*	*14,290*	*24,606*	*30,148*	*31,099*	*31,963*	*32,489*	*32,659*	27,300
1997	556	*3,283*	*6,505*	*12,970*	*18,850*	*23,259*	*26,116*	*28,810*	*29,193*	*29,292*	22,719

Table C.3: Triangle 1.3: Medical Malpractice: *Scpie Indemnity Company*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	2,716	24,576	43,990	59,722	71,019	76,354	76,792	77,207	77,588	77,656	135,318
1989	3,835	25,158	45,145	60,331	67,457	70,821	71,769	72,085	72,035	*72,171*	111,938
1990	4,838	27,965	50,873	66,400	71,875	74,755	75,176	75,250	*75,250*	*75,250*	99,293
1991	4,456	34,241	64,737	79,390	84,465	87,375	89,119	*89,825*	*90,333*	*90,343*	96,483
1992	5,970	36,080	68,268	81,783	86,076	87,167	*88,282*	*88,857*	*89,087*	*89,102*	98,608
1993	9,398	46,210	77,045	86,298	91,796	*93,827*	*95,871*	*95,943*	*95,987*	*96,008*	99,133
1994	6,181	39,204	70,006	82,385	*86,523*	*88,387*	*90,408*	*90,525*	*91,640*	*91,796*	97,097
1995	7,828	42,356	70,729	*79,340*	*81,142*	*81,891*	*81,905*	*81,897*	*81,781*	*81,782*	101,600
1996	8,854	51,400	*81,653*	*90,504*	*94,284*	*96,456*	*97,305*	*97,175*	*97,212*	*97,225*	101,537
1997	7,818	*47,098*	*84,142*	*93,724*	*97,401*	*97,726*	*98,250*	*98,355*	*98,627*	*98,655*	108,198

Table C.4: Triangle 2.1: Commercial Automobile: *Farmers' Automobile*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	2,491	4,426	5,484	6,275	6,544	6,828	7,103	7,103	7,111	7,115	9,147
1989	2,526	4,690	6,073	7,036	8,139	8,310	8,373	8,378	8,424	*8,425*	9,854
1990	2,760	4,442	5,529	6,965	7,533	8,192	8,919	8,983	*8,985*	*8,991*	11,142
1991	2,633	5,124	6,812	7,187	7,882	8,388	8,403	*8,524*	*8,538*	*8,549*	11,380
1992	2,502	4,780	5,790	6,705	7,193	7,473	*7,546*	*7,808*	*7,842*	*7,853*	11,808
1993	3,226	5,098	6,776	8,333	9,019	*9,391*	*9,514*	*9,665*	*9,943*	*9,980*	12,514
1994	3,567	5,796	7,553	8,322	*8,717*	*8,955*	*9,345*	*9,576*	*9,582*	*9,583*	13,460
1995	3,215	5,625	6,655	*7,785*	*8,570*	*8,644*	*8,721*	*8,797*	*8,802*	*9,038*	14,206
1996	3,516	6,025	*6,798*	*7,807*	*8,286*	*8,614*	*8,727*	*8,727*	*8,723*	*8,719*	14,231
1997	3,628	*5,907*	*8,950*	*10,581*	*11,453*	*12,364*	*12,431*	*12,432*	*12,433*	*12,434*	13,929

Table C.5: Triangle 2.2: Commercial Automobile: *State Farm*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	54,699	108,337	143,899	164,818	179,538	185,391	188,023	189,759	190,520	193,499	286,378
1989	60,091	119,366	151,151	174,665	185,469	192,213	196,152	198,013	199,997	*200,480*	308,908
1990	65,860	130,803	172,390	197,977	210,230	219,267	222,428	224,078	*224,855*	*225,430*	326,503
1991	61,946	121,108	158,880	182,689	195,247	201,854	204,911	*205,535*	*206,051*	*206,719*	332,616
1992	65,043	128,550	164,433	187,508	199,823	208,008	*210,925*	*213,006*	*214,097*	*214,524*	341,890
1993	72,295	144,579	185,446	208,388	219,345	*225,981*	*230,040*	*231,594*	*232,739*	*233,365*	355,840
1994	81,988	151,197	189,630	212,446	*229,511*	*237,406*	*240,779*	*242,463*	*242,860*	*244,280*	379,781
1995	83,207	152,470	190,974	*212,923*	*226,193*	*233,723*	*237,024*	*237,748*	*239,596*	*240,184*	398,755
1996	79,699	143,590	*184,346*	*206,780*	*224,327*	*233,298*	*237,216*	*236,725*	*236,825*	*237,368*	406,609
1997	75,827	*143,120*	*184,965*	*206,565*	*220,052*	*225,837*	*229,414*	*230,996*	*230,301*	*230,775*	406,516

Table C.6: Triangle 2.3: Commercial Automobile: *The Ins Co*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	152	288	276	429	442	802	802	791	791	791	1,300
1989	253	285	475	344	474	479	521	538	551	*553*	1,697
1990	217	404	567	587	594	479	479	479	*479*	*479*	1,553
1991	82	398	563	719	692	936	936	*936*	*936*	*936*	1,425
1992	602	1,495	1,703	1,855	1,903	1,919	*1,926*	*2,254*	*2,254*	*2,254*	3,252
1993	326	1,258	2,091	2,860	3,328	*3,363*	*4,333*	*4,352*	*4,452*	*4,458*	5,617
1994	731	2,209	3,218	4,243	*4,368*	*6,098*	*6,111*	*6,117*	*6,117*	*6,117*	9,536
1995	362	2,361	2,583	*2,861*	*3,973*	*4,409*	*4,450*	*4,463*	*4,472*	*4,480*	5,674
1996	1,293	2,705	*4,280*	*5,442*	*5,976*	*6,213*	*6,273*	*6,273*	*6,273*	*6,273*	7,389
1997	1,082	*2,360*	*3,726*	*4,687*	*4,880*	*5,165*	*5,322*	*5,322*	*5,322*	*5,322*	6,378

Table C.7: Triangle 3.1: Personal Automobile: *Farmers' Automobile*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	8,784	15,009	17,840	19,388	20,299	20,482	20,611	20,748	20,743	20,739	28,355
1989	10,262	18,849	24,027	26,024	27,483	27,907	28,054	27,979	28,013	*28,009*	31,312
1990	11,733	21,109	24,500	26,815	28,238	28,516	28,729	28,808	*28,811*	*28,826*	36,707
1991	11,915	21,078	25,506	28,481	30,054	30,645	31,065	*31,150*	*31,241*	*31,272*	40,265
1992	11,949	20,676	25,841	28,292	29,766	30,440	*30,571*	*30,746*	*30,781*	*30,780*	42,570
1993	13,270	23,902	28,318	31,328	33,197	*33,605*	*33,882*	*33,946*	*33,987*	*34,035*	42,879
1994	14,561	24,177	29,091	32,281	*34,287*	*34,962*	*35,027*	*35,094*	*35,226*	*35,226*	43,684
1995	14,545	23,908	29,665	*32,150*	*34,016*	*34,844*	*35,000*	*35,029*	*35,334*	*35,400*	44,500
1996	15,098	24,813	*29,524*	*31,727*	*33,144*	*33,221*	*33,274*	*33,317*	*33,315*	*33,316*	45,276
1997	14,846	*24,270*	*28,303*	*31,097*	*32,304*	*33,378*	*33,522*	*33,583*	*33,662*	*33,661*	47,331

Table C.8: Triangle 3.2: Personal Automobile: *Federal Insurance Company*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	13,440	35,680	48,703	56,319	61,018	61,119	63,049	63,556	63,744	63,835	83,473
1989	18,757	44,166	57,578	66,264	65,600	67,721	75,369	76,713	77,007	*77,029*	91,800
1990	19,834	42,225	56,347	63,194	67,112	69,459	74,267	79,208	*79,250*	*79,308*	95,877
1991	16,230	38,045	46,055	53,983	60,638	62,917	71,074	*71,885*	*72,560*	*72,579*	99,256
1992	14,629	22,427	33,873	43,339	53,168	60,413	*61,452*	*62,363*	*62,464*	*62,458*	96,170
1993	24,597	51,373	68,484	80,253	92,192	*94,939*	*97,226*	*97,536*	*97,654*	*97,787*	139,038
1994	31,723	59,733	77,398	94,395	*101,008*	*104,557*	*107,399*	*108,067*	*108,476*	*110,038*	152,174
1995	37,397	71,133	94,294	*103,996*	*107,948*	*109,478*	*110,401*	*111,051*	*111,108*	*111,598*	167,833
1996	53,670	98,628	*112,473*	*123,070*	*129,739*	*131,549*	*132,682*	*133,137*	*133,426*	*133,522*	180,523
1997	52,837	*77,758*	*95,357*	*104,789*	*109,025*	*111,835*	*112,467*	*113,000*	*113,086*	*113,371*	164,717

Table C.9: Triangle 3.3: Personal Automobile: *State Farm*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	2,439,272	4,722,902	5,705,646	6,238,289	6,519,491	6,677,426	6,750,431	6,787,444	6,808,809	6,815,646	7,809,394
1989	2,828,267	5,368,026	6,494,597	7,096,377	7,417,869	7,575,814	7,655,217	7,693,240	7,712,077	*7,721,911*	8,764,863
1990	3,186,948	5,913,490	7,140,613	7,774,615	8,096,374	8,251,086	8,325,184	8,364,955	*8,382,479*	*8,394,117*	9,796,463
1991	3,192,619	5,878,000	7,075,254	7,698,459	7,996,404	8,140,065	8,215,810	*8,256,338*	*8,279,401*	*8,288,143*	10,594,952
1992	3,561,950	6,474,291	7,763,969	8,404,713	8,716,926	8,876,813	*8,959,761*	*8,998,462*	*9,016,704*	*9,026,329*	11,457,922
1993	3,895,076	7,024,867	8,343,417	9,003,120	9,337,099	*9,509,203*	*9,599,735*	*9,642,522*	*9,660,532*	*9,673,610*	12,240,633
1994	4,323,103	7,590,944	8,924,062	9,640,098	*10,030,681*	*10,216,964*	*10,299,231*	*10,340,679*	*10,361,560*	*10,375,605*	13,277,675
1995	4,491,070	7,664,190	9,006,113	*9,736,570*	*10,120,120*	*10,312,088*	*10,412,421*	*10,461,753*	*10,497,027*	*10,512,108*	14,125,898
1996	4,444,088	7,486,113	*8,845,221*	*9,583,622*	*9,999,739*	*10,196,271*	*10,293,776*	*10,347,058*	*10,373,438*	*10,387,245*	14,664,665
1997	4,344,144	*7,305,064*	*8,614,474*	*9,379,418*	*9,792,901*	*9,988,209*	*10,076,219*	*10,123,792*	*10,148,983*	*10,165,481*	14,923,375

Table C.10: Triangle 4.1: Product Liability: *Allstate Insurance Company*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	1,501	3,916	8,834	17,450	22,495	28,687	31,311	32,039	36,357	36,358	48,622
1989	1,697	5,717	10,442	18,125	23,284	30,092	34,338	41,094	41,164	*41,189*	46,663
1990	1,373	4,002	10,829	16,695	21,788	25,332	34,875	34,893	*35,037*	*35,037*	39,473
1991	1,069	4,594	6,920	9,996	13,249	19,221	19,256	*19,230*	*19,232*	*19,233*	31,571
1992	1,134	3,068	5,412	8,210	19,164	19,187	*19,190*	*19,209*	*19,209*	*19,213*	22,317
1993	979	3,079	6,407	16,113	16,131	*16,148*	*16,153*	*16,201*	*16,222*	*16,247*	24,398
1994	1,397	2,990	25,688	26,030	*26,165*	*26,294*	*26,765*	*27,075*	*27,085*	*27,086*	34,271
1995	1,016	21,935	22,095	*22,388*	*22,694*	*22,882*	*23,253*	*23,969*	*24,861*	*25,216*	26,583
1996	9,852	10,071	*10,444*	*10,743*	*11,011*	*11,037*	*11,567*	*11,788*	*11,829*	*11,841*	23,440
1997	319	*964*	*3,410*	*3,179*	*3,460*	*4,006*	*4,333*	*4,650*	*4,812*	*5,049*	4,450

Table C.11: Triangle 4.2: Product Liability: *Federal Insurance Company*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	1,249	4,092	8,893	15,516	23,806	34,070	30,699	31,130	31,716	32,430	102,414
1989	946	2,929	9,953	17,368	25,139	29,628	32,059	32,972	33,610	*34,292*	98,541
1990	1,765	4,743	9,854	20,471	26,756	33,654	43,282	42,588	*49,336*	*55,502*	100,215
1991	1,408	5,226	9,768	21,637	25,339	31,574	37,738	*50,255*	*53,313*	*53,586*	103,518
1992	1,647	7,628	12,848	18,323	21,943	23,139	*32,579*	*34,833*	*36,287*	*40,504*	104,926
1993	7,566	14,991	29,466	49,090	57,590	*66,854*	*69,663*	*72,885*	*84,050*	*93,531*	113,504
1994	2,299	6,856	23,954	61,496	*78,905*	*90,416*	*95,287*	*100,858*	*104,661*	*106,357*	123,984
1995	4,959	13,541	23,003	*44,313*	*73,566*	*83,381*	*89,134*	*93,387*	*97,683*	*101,142*	140,440
1996	6,063	9,707	*26,814*	*55,185*	*71,471*	*85,110*	*90,315*	*99,066*	*101,633*	*103,677*	150,180
1997	6,507	*29,036*	*43,380*	*66,827*	*82,385*	*107,656*	*116,351*	*120,486*	*124,430*	*129,300*	149,656

Table C.12: Triangle 4.3: Product Liability: *Federated Mutual Group*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	3,894	7,409	11,199	16,913	18,517	19,722	20,417	20,902	21,208	21,368	29,513
1989	3,883	6,869	10,489	14,410	16,605	18,485	20,112	20,654	20,835	*21,119*	27,835
1990	4,178	7,473	11,281	15,874	19,186	20,993	22,069	22,477	*22,742*	*23,279*	28,309
1991	3,872	8,137	14,577	18,071	22,125	23,770	24,559	*25,283*	*26,017*	*26,380*	28,091
1992	4,095	7,260	11,015	14,844	19,309	20,097	*20,927*	*21,138*	*21,424*	*21,658*	29,036
1993	2,223	4,048	6,651	8,229	9,141	*10,446*	*11,078*	*11,098*	*11,510*	*11,598*	23,916
1994	2,500	4,003	6,893	9,379	*11,092*	*11,386*	*11,602*	*11,684*	*12,304*	*12,353*	26,223
1995	2,442	4,657	7,303	*8,380*	*9,553*	*10,170*	*10,446*	*10,522*	*10,618*	*10,883*	28,379
1996	2,761	4,661	*7,009*	*10,018*	*11,108*	*12,333*	*13,023*	*13,131*	*14,071*	*14,100*	29,809
1997	2,642	*5,027*	*9,539*	*12,343*	*15,013*	*16,282*	*16,917*	*17,105*	*17,236*	*17,336*	31,953

Table C.13: Triangle 5.1: Workers' Compensation: *Allstate Insurance Company*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	70,571	155,905	220,744	251,595	274,156	287,676	298,499	304,873	321,808	325,322	394,742
1989	66,547	136,447	179,142	211,343	231,430	244,750	254,557	270,059	273,873	*277,574*	374,252
1990	52,233	133,370	178,444	204,442	222,193	232,940	253,337	256,788	*261,166*	*263,000*	280,320
1991	59,315	128,051	169,793	196,685	213,165	234,676	239,195	*245,499*	*247,131*	*248,319*	313,982
1992	39,991	89,873	114,117	133,003	154,362	159,496	*164,013*	*166,212*	*167,397*	*168,844*	252,698
1993	19,744	47,229	61,909	85,099	87,215	*88,602*	*89,444*	*89,899*	*90,446*	*90,686*	201,055
1994	20,379	46,773	88,636	91,077	*92,583*	*93,346*	*93,897*	*94,165*	*94,558*	*94,730*	174,381
1995	18,756	84,712	87,311	*89,200*	*90,001*	*90,247*	*90,687*	*91,068*	*91,001*	*91,161*	146,366
1996	42,609	44,916	*46,981*	*47,899*	*48,583*	*49,109*	*49,442*	*49,073*	*49,161*	*49,255*	93,294
1997	691	*2,085*	*2,795*	*2,866*	*2,905*	*2,909*	*2,908*	*2,909*	*2,909*	*2,909*	7,651

Table C.14: Triangle 5.2: Workers' Compensation: *Lumbermen's Underwriting Alliance*

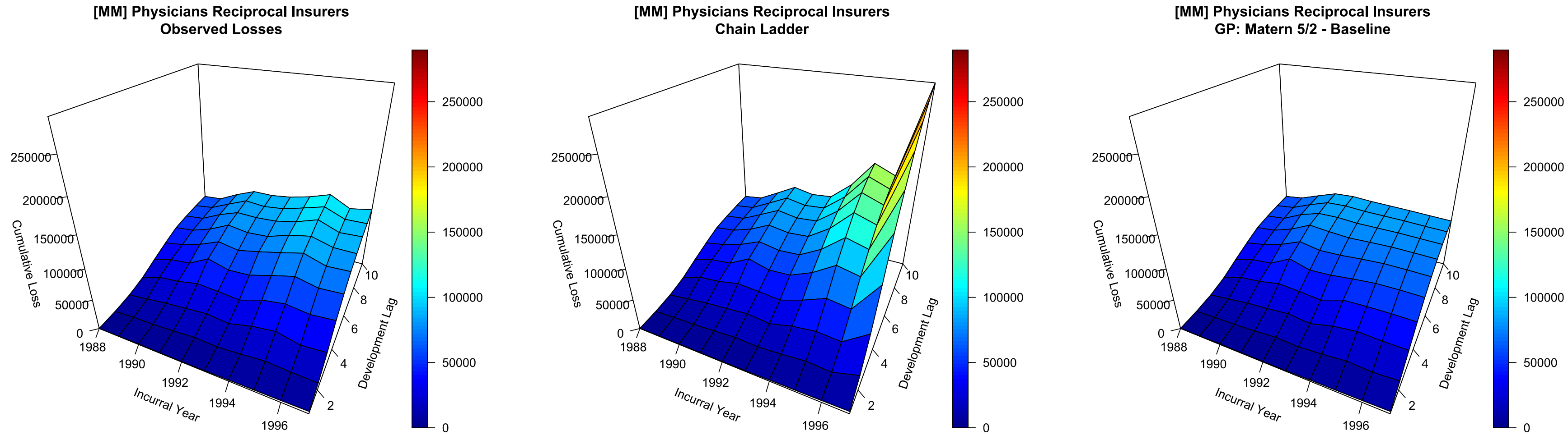
IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	19,016	44,632	59,804	66,052	70,115	72,219	73,565	74,273	75,112	75,655	94,389
1989	17,346	42,058	59,686	64,821	67,313	69,036	69,942	70,428	70,846	*71,001*	92,238
1990	12,212	28,087	42,719	46,564	48,016	49,030	49,700	49,994	*50,144*	*50,385*	68,123
1991	9,490	19,697	32,062	38,698	40,369	41,220	41,970	*42,359*	*42,720*	*42,887*	59,587
1992	7,605	14,874	21,105	29,016	35,208	36,884	*37,456*	*37,767*	*37,536*	*37,684*	62,295
1993	5,596	11,527	14,677	17,073	20,813	*24,037*	*26,021*	*27,557*	*27,593*	*27,752*	70,875
1994	4,885	10,118	13,103	14,570	*15,193*	*15,680*	*16,474*	*18,136*	*19,774*	*21,102*	62,394
1995	4,056	8,435	10,439	*11,424*	*13,000*	*13,611*	*13,847*	*14,252*	*14,526*	*14,678*	67,685
1996	4,213	8,768	*10,280*	*11,093*	*11,848*	*12,180*	*12,657*	*13,232*	*13,318*	*13,544*	50,118
1997	5,258	*10,661*	*13,346*	*15,717*	*17,620*	*18,884*	*19,382*	*19,914*	*20,523*	*20,734*	49,456

Table C.15: Triangle 5.3: Workers' Compensation: *State Farm*

IY	Development Lag										Premium
	1	2	3	4	5	6	7	8	9	10	
1988	22,190	60,834	85,104	100,151	108,802	114,967	118,790	121,558	123,492	125,049	177,104
1989	26,542	77,798	106,407	122,422	133,359	138,599	143,029	145,712	147,358	*149,252*	201,118
1990	32,977	100,494	134,886	157,758	168,991	178,065	182,787	187,760	*192,576*	*196,092*	246,010
1991	38,604	114,428	157,103	181,322	197,411	208,804	213,396	*221,596*	*226,756*	*229,642*	286,019
1992	42,466	125,820	164,776	189,045	204,377	213,904	*221,659*	*228,415*	*231,760*	*235,831*	340,183
1993	46,447	116,764	154,897	179,419	193,676	*206,073*	*212,193*	*216,710*	*220,900*	*224,868*	418,755
1994	41,368	100,344	132,021	151,081	*165,667*	*173,167*	*177,960*	*181,932*	*185,039*	*190,572*	366,031
1995	35,719	83,216	111,268	*127,221*	*136,597*	*142,371*	*146,359*	*148,538*	*151,476*	*153,299*	338,186
1996	28,746	66,033	*87,748*	*100,432*	*108,442*	*116,297*	*119,566*	*121,885*	*124,377*	*126,403*	286,631
1997	25,265	*61,872*	*81,038*	*92,519*	*99,006*	*103,738*	*106,810*	*108,329*	*110,509*	*111,592*	245,378

D | Results: 3D-plots of benchmark measurements

In this Appendix, we will give 3D-plots of all observed losses, the Chain Ladder prediction and the best performing Gaussian process model.

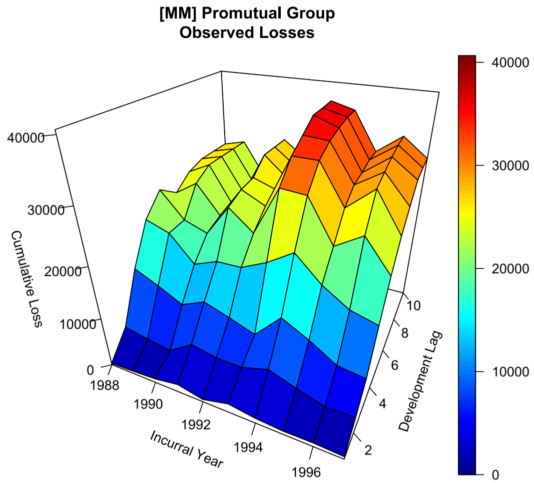


(a) Observed Losses

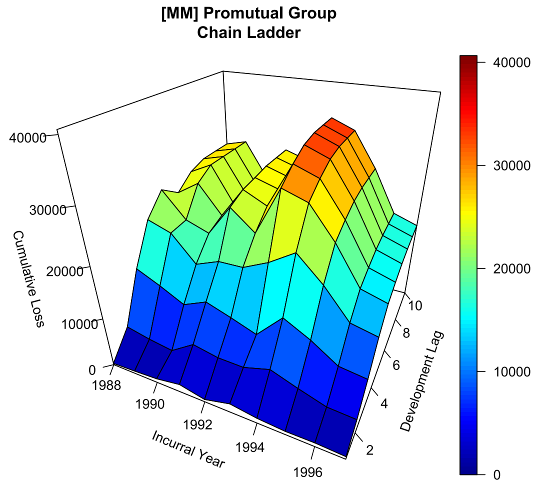
(b) Chain Ladder

(c) Matérn 5/2

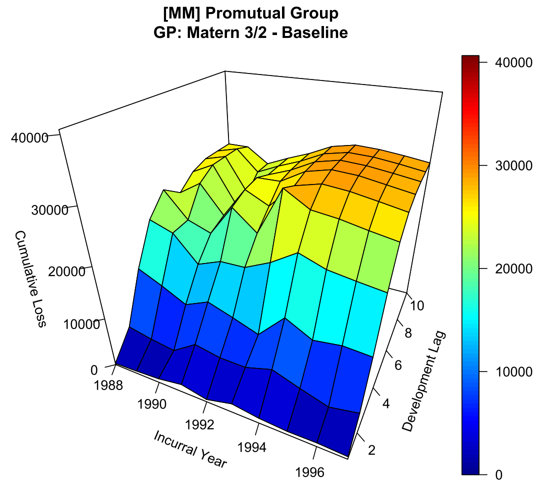
Figure D.1: Benchmark Results of Triangle 1.1



(a) Observed Losses

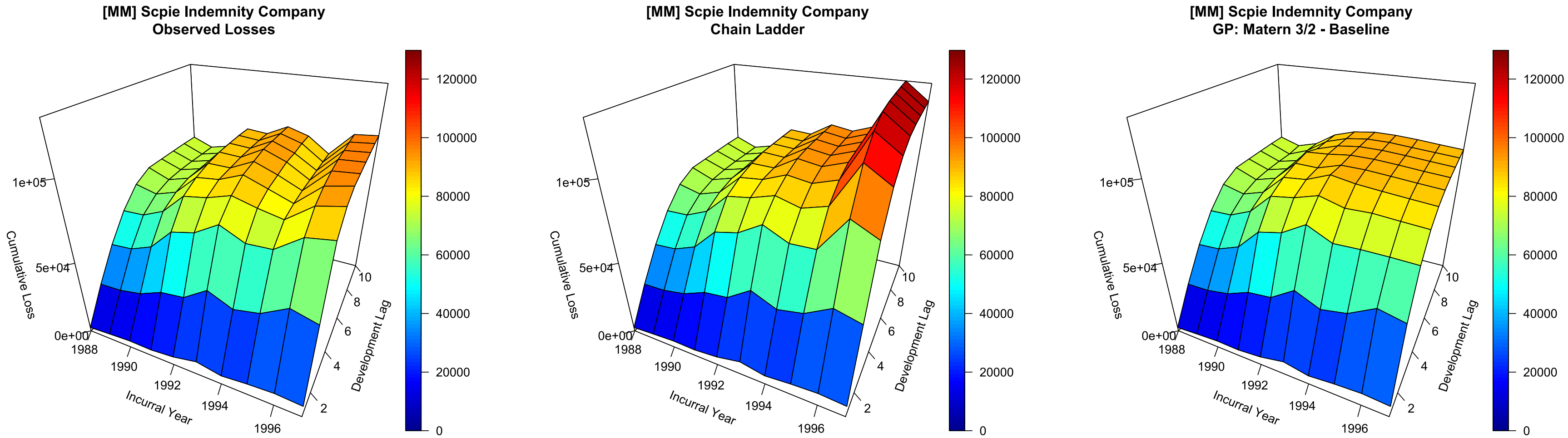


(b) Chain Ladder



(c) Matérn 3/2

Figure D.2: Benchmark Results of Triangle 1.2

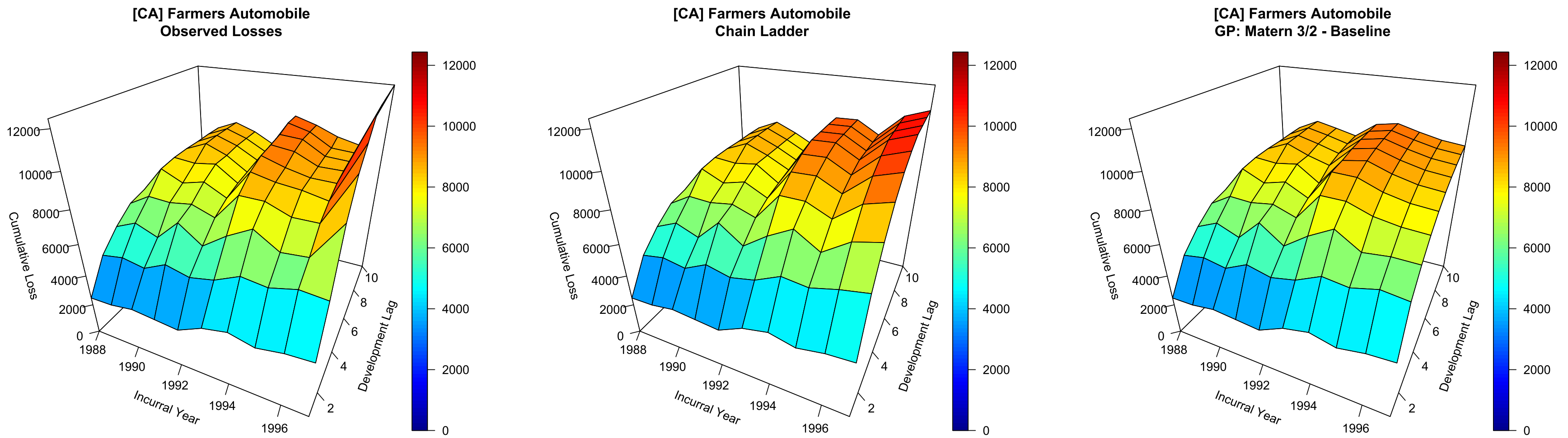


(a) Observed Losses

(b) Chain Ladder

(c) Matérn 3/2

Figure D.3: Benchmark Results of Triangle 1.3

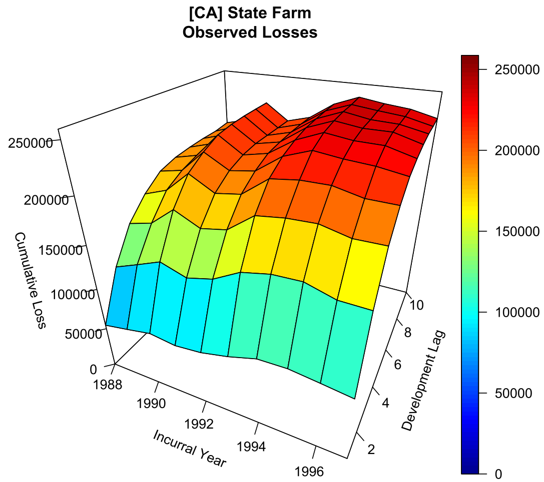


(a) Observed Losses

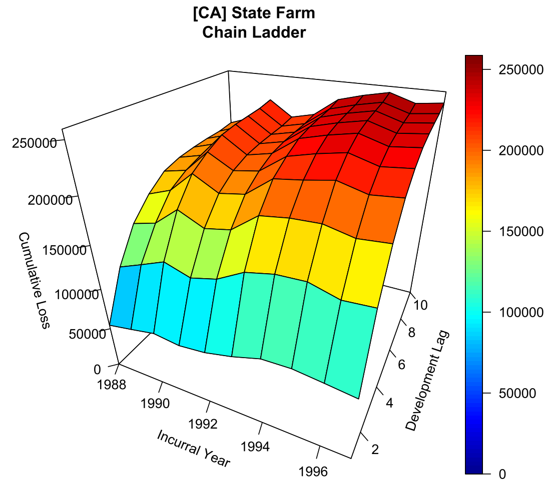
(b) Chain Ladder

(c) Matérn 3/2

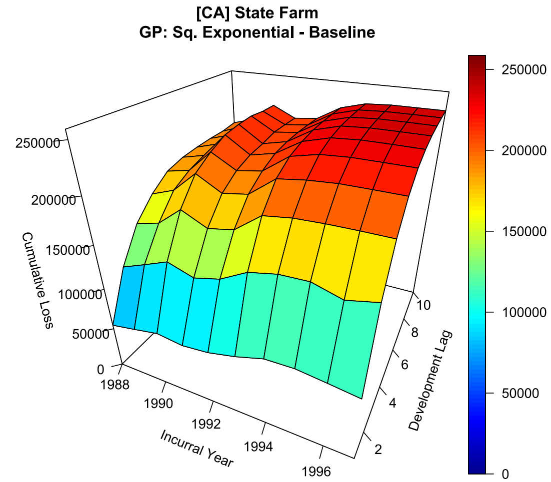
Figure D.4: Benchmark Results of Triangle 2.1



(a) Observed Losses

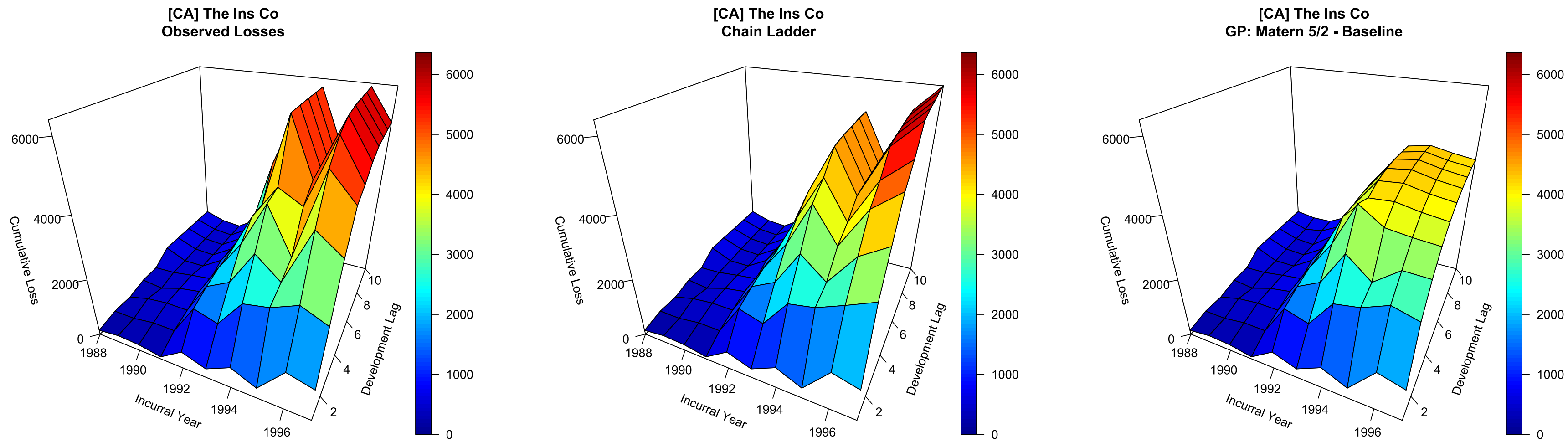


(b) Chain Ladder



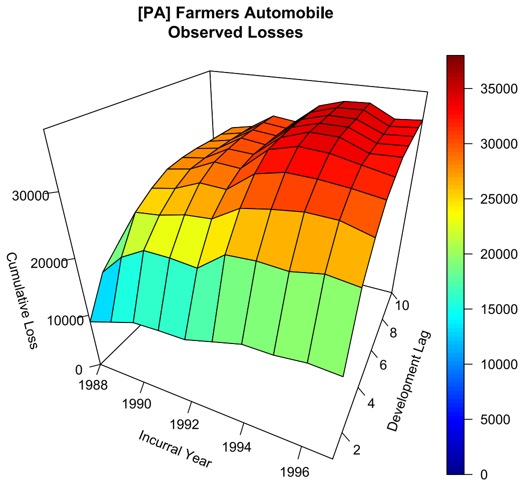
(c) Squared Exponential

Figure D.5: Benchmark Results of Triangle 2.2

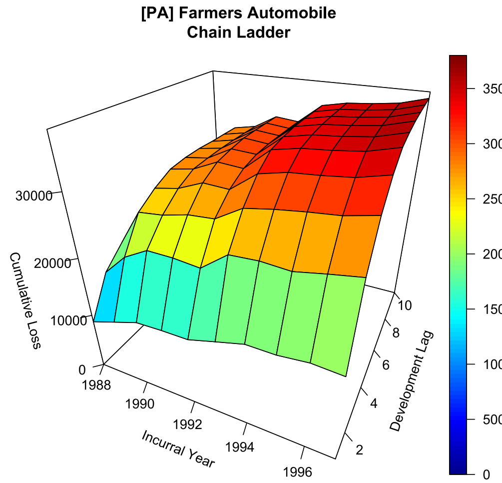


(a) Observed Losses (b) Chain Ladder (c) Matérn 5/2

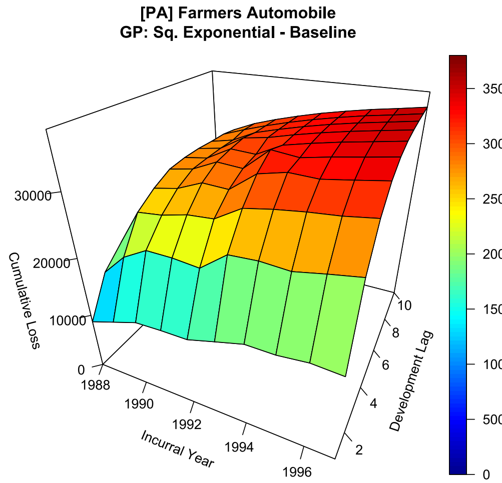
Figure D.6: Benchmark Results of Triangle 2.3



(a) Observed Losses

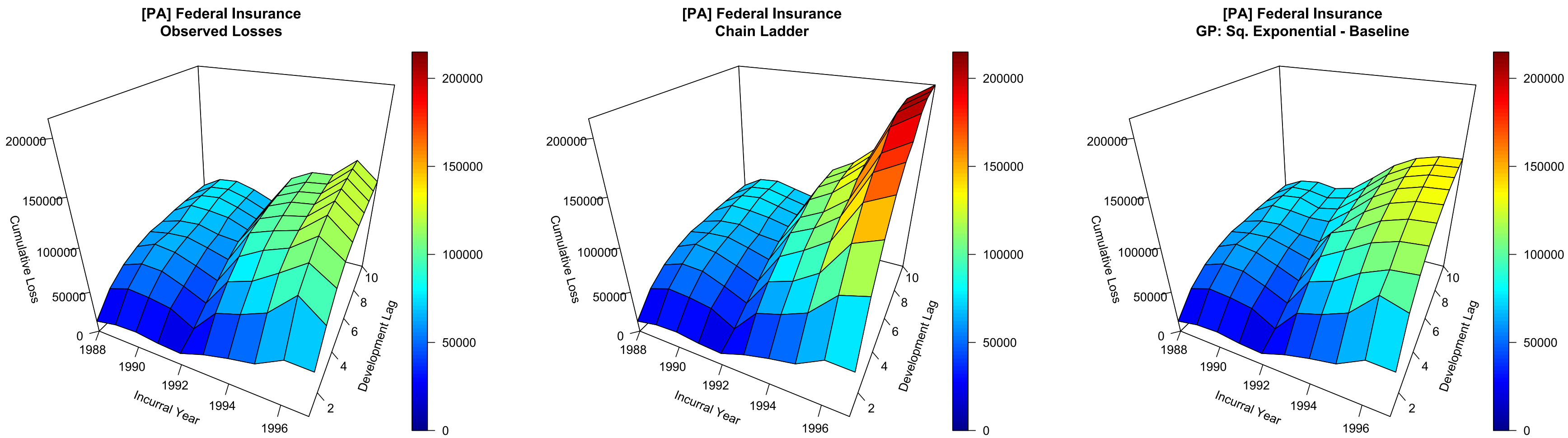


(b) Chain Ladder



(c) Squared Exponential

Figure D.7: Benchmark Results of Triangle 3.1

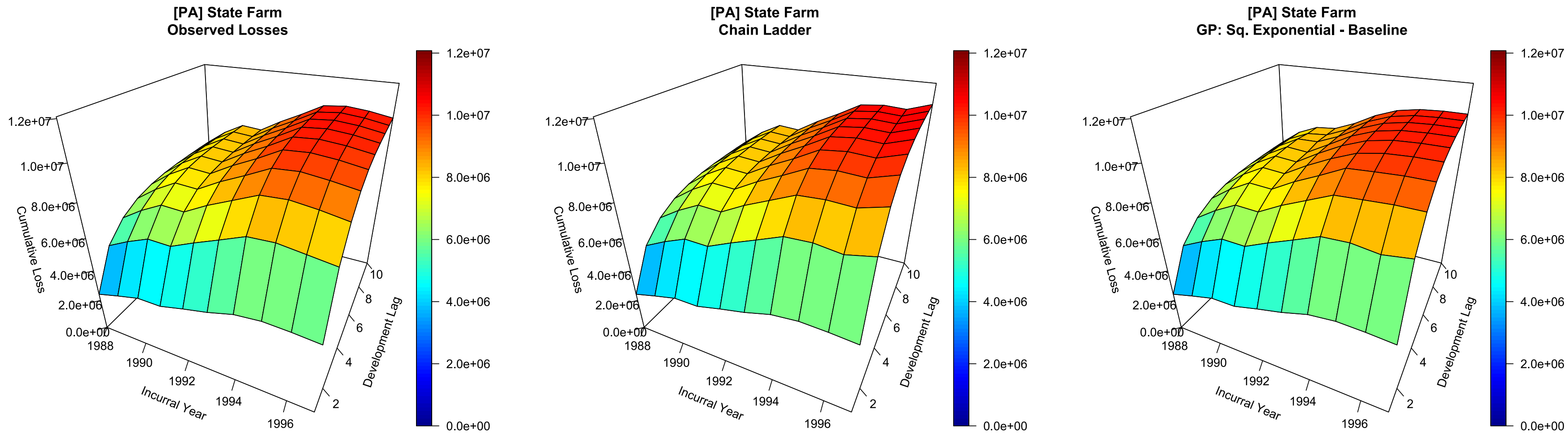


(a) Observed Losses

(b) Chain Ladder

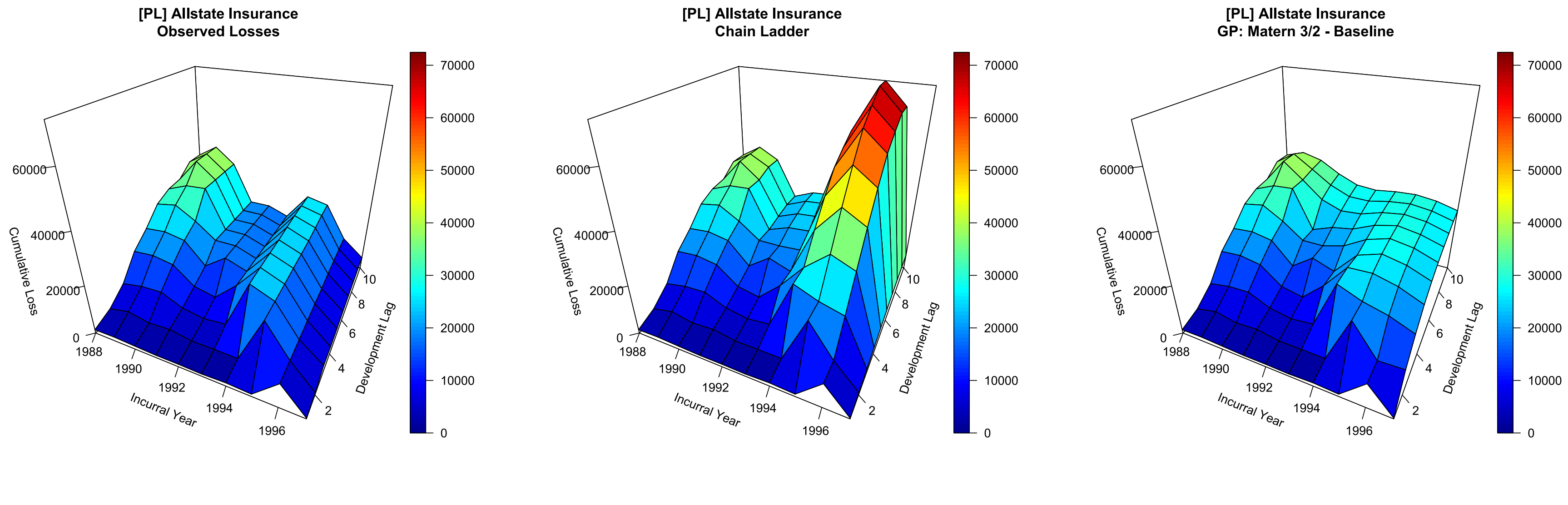
(c) Squared Exponential

Figure D.8: Benchmark Results of Triangle 3.2



(a) Observed Losses (b) Chain Ladder (c) Squared Exponential

Figure D.9: Benchmark Results of Triangle 3.3

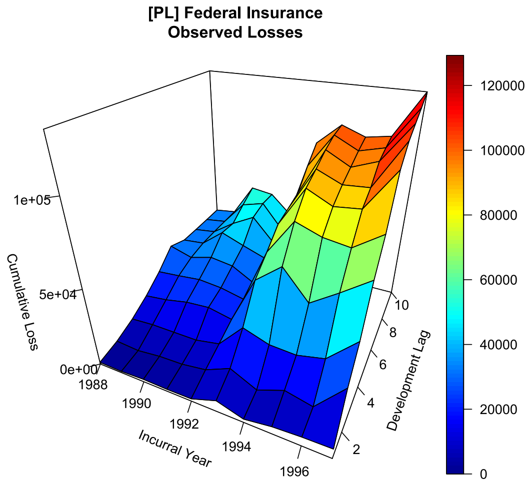


(a) Observed Losses

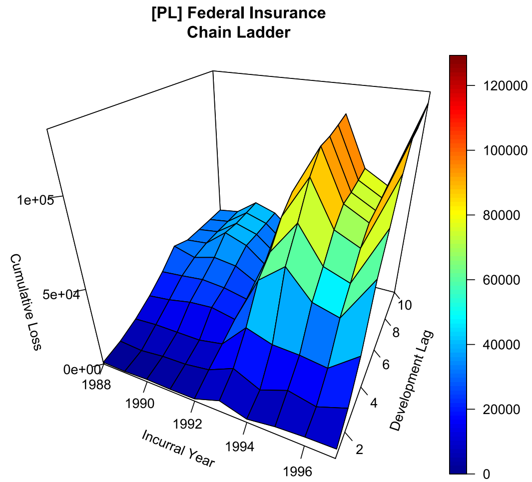
(b) Chain Ladder

(c) Matérn 3/2

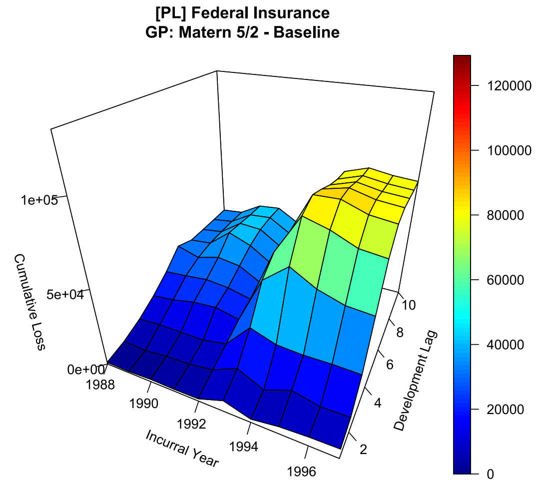
Figure D.10: Benchmark Results of Triangle 4.1



(a) Observed Losses

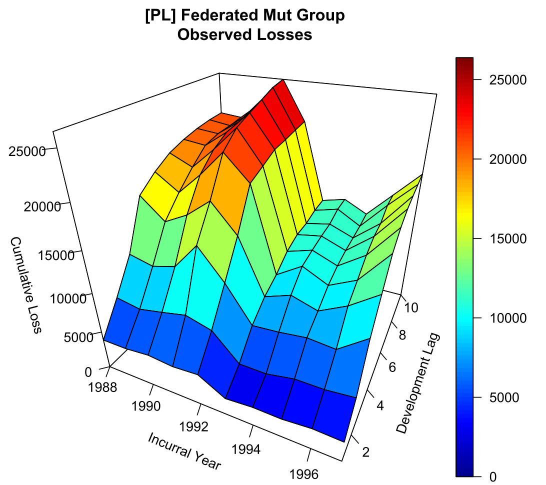


(b) Chain Ladder

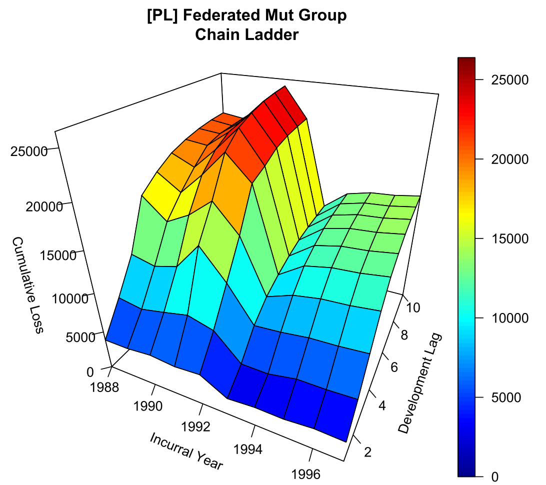


(c) Matérn 5/2

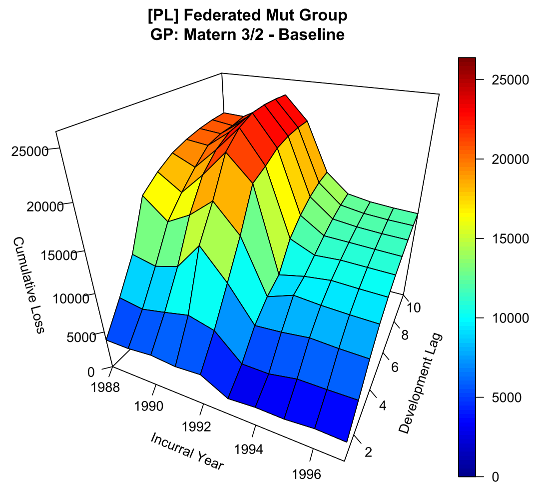
Figure D.11: Benchmark Results of Triangle 4.2



(a) Observed Losses

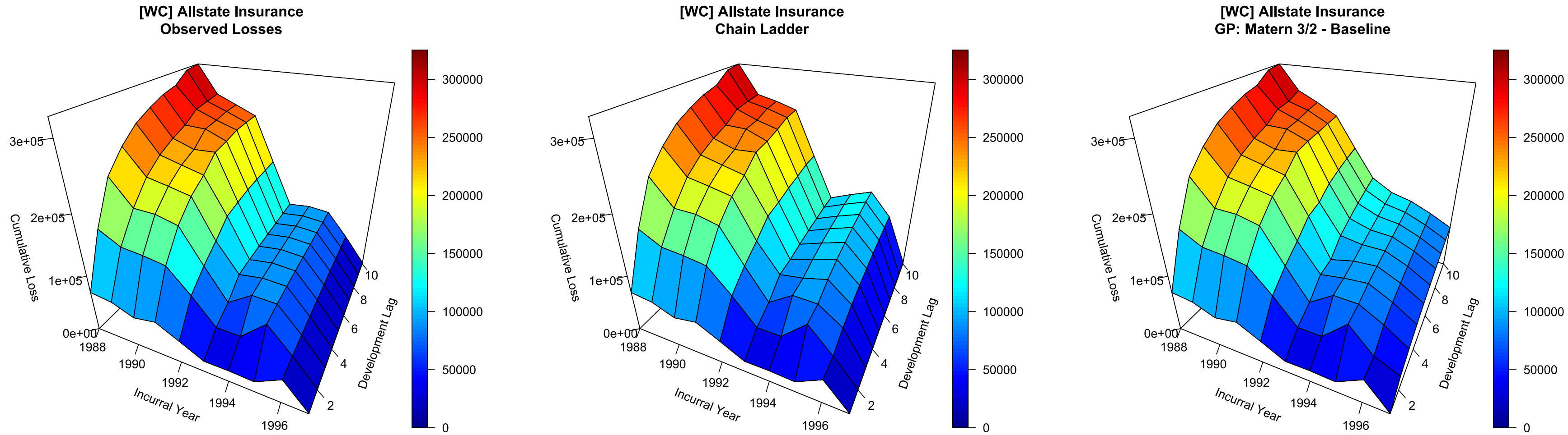


(b) Chain Ladder



(c) Matérn 3/2

Figure D.12: Benchmark Results of Triangle 4.3

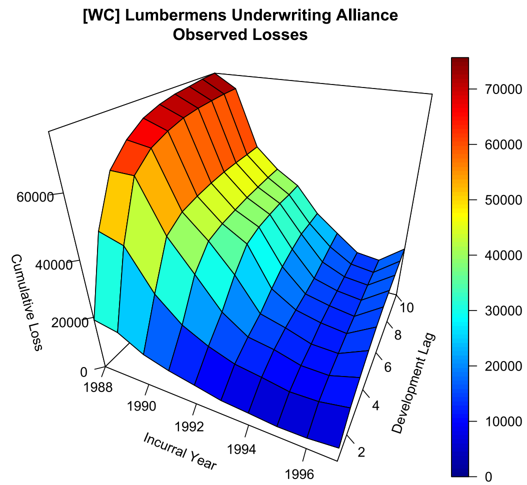


(a) Observed Losses

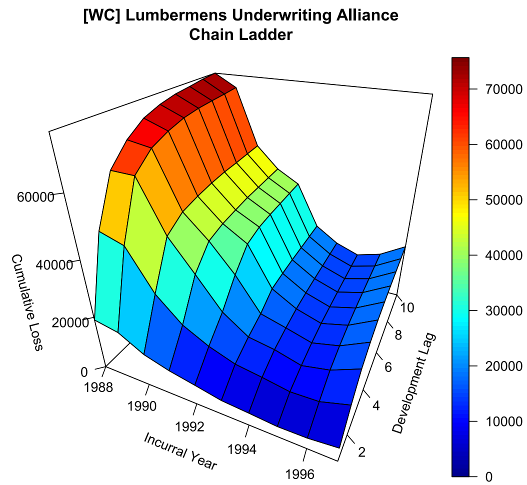
(b) Chain Ladder

(c) Matérn 3/2

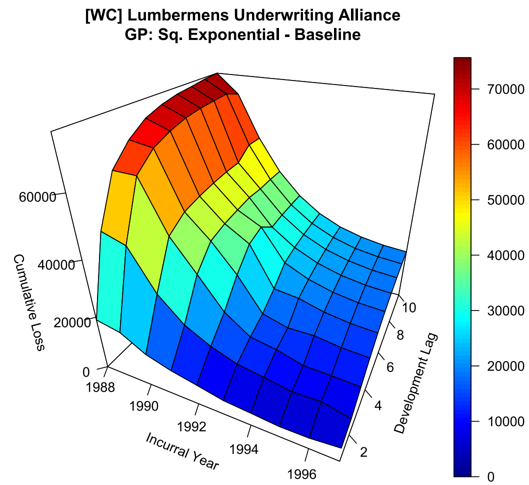
Figure D.13: Benchmark Results of Triangle 5.1



(a) Observed Losses

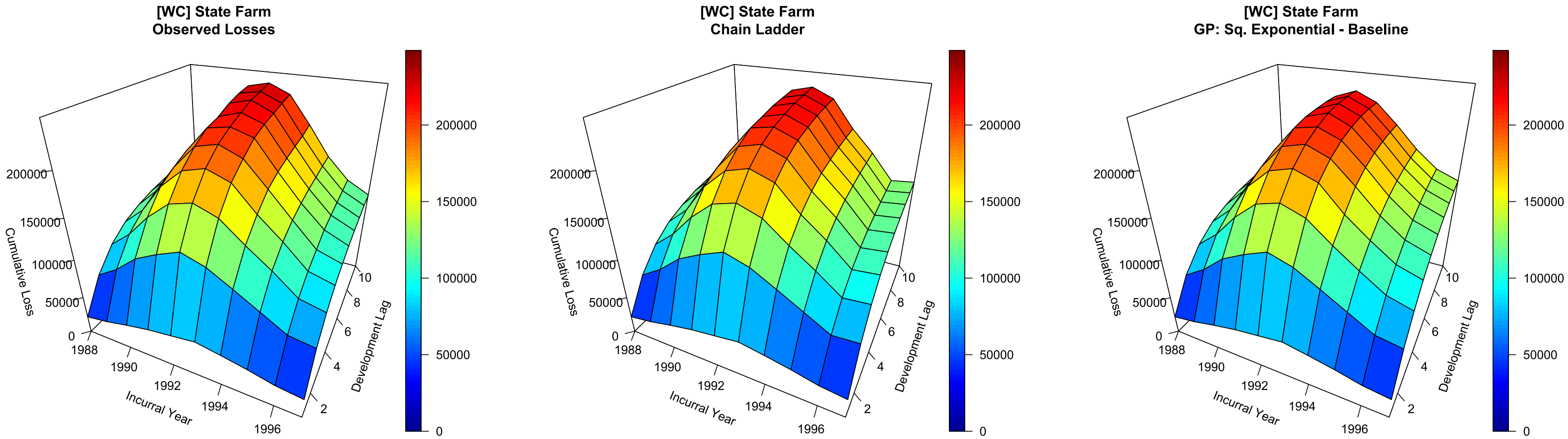


(b) Chain Ladder



(c) Squared Exponential

Figure D.14: Benchmark Results of Triangle 5.2



(a) Observed Losses

(b) Chain Ladder

(c) Squared Exponential

Figure D.15: Benchmark Results of Triangle 5.3

E | Results: Inference of prior parameters

In this Appendix, we will present the inferred parameters by varying the prior distribution of that parameter.

E.1 | Bandwidth parameter (ψ)

Table E.1: Inferred values of ψ_1 (Development Lag), by variations on the prior distribution

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
	LoB	Cauchy	Gamma	Cauchy	Gamma	Cauchy	Gamma
1.1	MM	0.9011	0.759	1.5666	1.011	2.0112	1.3025
1.2	MM	1.8657	1.0696	3.6252	1.46	3.9048	1.4869
1.3	MM	1.7194	1.057	2.3654	1.2587	2.3059	1.3705
2.1	CA	0.9572	0.5899	1.6179	1.1411	2.0127	1.1697
2.2	CA	0.7161	0.6716	0.9816	0.7754	1.2269	0.9325
2.3	CA	1.4056	0.592	2.377	1.0622	2.9131	1.2317
3.1	PA	0.8552	0.7606	1.2409	0.918	1.3248	1.0163
3.2	PA	0.6965	0.3983	1.2279	0.4822	1.7508	0.9917
3.3	PA	0.7707	0.741	1.0265	0.799	1.1367	0.895
4.1	PL	2.9466	0.9792	3.4757	1.2253	2.7313	1.3839
4.2	PL	2.3923	0.9886	4.0955	1.3872	4.7952	1.8635
4.3	PL	0.6576	0.5898	1.1864	0.8247	1.4953	1.0867
5.1	WC	0.3939	0.3992	0.6899	0.5168	0.8435	0.7688
5.2	WC	0.8526	0.7151	1.3228	0.9011	1.3575	0.9721
5.3	WC	0.5408	0.594	0.8642	0.771	1.0241	0.8764

Considering several remarkably high values of the bandwidth parameter, we investigate the Standard Deviation of the predicted Loss Reserve, given in Table E.3. For some triangles, we see that the high inferred values result in a lower standard deviation.

Table E.2: Inferred values of ψ_2 (Incurral year), by variations on the prior distribution

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
	LoB	Cauchy	Gamma	Cauchy	Gamma	Cauchy	Gamma
1.1	MM	1.8771	1.2233	3.9375	1.7262	5.7244	1.5311
1.2	MM	4.2627	1.2793	4.3455	0.8887	0.7579	0.8109
1.3	MM	0.842	0.7447	1.4059	1.0019	1.6367	1.1599
2.1	CA	5.1424	2.1202	11.3557	1.0481	18.7123	0.842
2.2	CA	1.5652	1.1588	3.3262	1.7291	5.1914	2.1369
2.3	CA	16.7376	2.7361	34.3191	2.132	45.0727	2.1477
3.1	PA	1.209	0.9369	1.6958	1.0855	1.6692	0.9507
3.2	PA	7.1923	2.4592	14.9481	3.785	25.471	2.9871
3.3	PA	1.1165	0.9635	2.2756	1.476	3.479	1.9834
4.1	PL	11.5624	1.6699	14.2031	1.8285	11.0983	2.0517
4.2	PL	8.5283	2.2414	14.2702	3.1457	18.1432	4.777
4.3	PL	3.4516	1.6806	6.4467	2.2157	8.0759	2.6933
5.1	WC	3.6954	1.8756	7.7373	2.7808	12.3416	3.3932
5.2	WC	1.7569	1.1723	3.4214	1.7134	4.9405	1.4946
5.3	WC	1.2261	1.0546	1.9667	1.3962	2.3216	1.5795

Table E.3: Standard Deviation of Loss Reserve predictions, by variations on the prior distribution of ψ

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
	LoB	Cauchy	Gamma	Cauchy	Gamma	Cauchy	Gamma
1.1	MM	89,726	**86,625**	98,296	**90,804**	112,678	**83,825**
1.2	MM	34,405	**33,058**	35,024	**27,092**	29,277	**26,187**
1.3	MM	43,673	**41,430**	46,733	**36,476**	49,412	**37,098**
2.1	CA	**5,175**	5,458	5,737	**4,712**	6,094	**4,128**
2.2	CA	73,820	**73,495**	80,097	**74,844**	104,107	**65,842**
2.3	CA	**2,821**	3,931	**3,080**	4061	**3,314**	4,335
3.1	PA	8,474	**8,039**	9,854	**8,852**	10,948	**8,223**
3.2	PA	**80,404**	92,948	**89,888**	118,434	**92,693**	98,847
3.3	PA	2,550,905	**2,400,639**	2,654,248	**2,616,817**	3,135,120	**2,541,395**
4.1	PL	**51,172**	57,195	**52,128**	56,209	58,583	**58,213**
4.2	PL	**88,689**	95,199	**94,270**	128,196	**94,220**	116,028
4.3	PL	**17,664**	21,155	**20,138**	21,202	20,169	**18,783**
5.1	WC	**179,668**	229,973	**188,070**	226,372	**189,414**	202,139
5.2	WC	**36,374**	37,775	38,995	**37,127**	37,352	**26,968**
5.3	WC	**103,864**	107,898	**111,896**	120,175	109,435	**104,266**

E.2 | Noise parameter (σ^2)

Table E.4: Inferred values of σ^2, by variations on the prior distribution

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
	LoB	Log-normal	Student-T	Log-normal	Student-T	Log-normal	Student-T
1.1	MM	0.0117	0.0020	0.0142	0.0027	0.0197	0.0086
1.2	MM	0.0439	0.0228	0.0529	0.0425	0.0582	0.0484
1.3	MM	0.0083	0.0014	0.0083	0.0016	0.0090	0.0019
2.1	CA	0.0354	0.0039	0.0435	0.0348	0.0474	0.0405
2.2	CA	0.0067	0.0009	0.0068	0.0008	0.0158	0.0007
2.3	CA	0.0388	0.0075	0.0618	0.0495	0.0669	0.0584
3.1	PA	0.0085	0.0015	0.0091	0.0020	0.0098	0.0027
3.2	PA	0.0154	0.0024	0.0205	0.0029	0.0768	0.0634
3.3	PA	0.0061	0.0008	0.0065	0.0008	0.0090	0.0008
4.1	PL	0.0767	0.0641	0.0872	0.0763	0.0911	0.0794
4.2	PL	0.0192	0.0042	0.0203	0.0063	0.0465	0.0133
4.3	PL	0.0093	0.0016	0.0091	0.0017	0.0096	0.0017
5.1	WC	0.0097	0.0014	0.0113	0.0017	0.0272	0.0033
5.2	WC	0.0063	0.0009	0.0074	0.0010	0.0088	0.0019
5.3	WC	0.0050	0.0007	0.0044	0.0006	0.0041	0.0006

E.3 | Signal parameter (η^2)

Table E.5: Inferred values of η^2 by variations on the prior distribution

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
$\hat{\eta}$	LoB	Log-normal	Student-T	Log-normal	Student-T	Log-normal	Student-T
1.1	MM	2.6722	2.2072	4.271	3.0145	2.9359	2.149
1.2	MM	2.923	2.1061	2.4737	1.8748	1.9976	1.5398
1.3	MM	1.6074	1.4984	2.3112	1.8552	2.0993	1.735
2.1	CA	4.3613	3.341	3.0393	2.0477	2.1474	1.7373
2.2	CA	2.2507	1.9389	3.62	2.7136	3.3295	2.4536
2.3	CA	9.1009	6.6204	7.0691	4.3056	4.5172	2.9945
3.1	PA	2.1043	1.8129	2.9691	2.3423	3.0276	2.2183
3.2	PA	10.1939	7.5789	23.0048	17.765	6.7857	4.5751
3.3	PA	2.1512	1.8537	3.5401	2.6887	3.7304	2.7291
4.1	PL	3.5845	2.6519	4.401	3.1203	3.8026	2.5663
4.2	PL	8.2508	6.1809	14.8427	11.096	8.9884	6.1213
4.3	PL	2.3928	2.0015	3.4916	2.5203	2.4177	1.8492
5.1	WC	2.8707	2.3103	5.2233	3.5938	3.5168	2.5707
5.2	WC	1.3889	1.3048	2.1633	1.7856	1.7648	1.5268
5.3	WC	2.0471	1.7975	3.2412	2.5005	2.9885	2.3228

E.4 | Warping parameters ($\alpha_1, \beta_1, \alpha_2, \beta_2$)

Table E.6: Inferred values of α_1, by variations on the prior distribution

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
$\hat{\alpha_1}$	LoB	Snoek	Lally	Snoek	Lally	Snoek	Lally
1.1	MM	1.0901	1.1121	1.1111	1.1537	1.0836	1.094
1.2	MM	0.939	1.1883	1.0106	1.1894	1.0244	1.1904
1.3	MM	0.5156	0.7005	0.6352	0.8245	0.7074	0.8696
2.1	CA	0.5234	0.6033	0.573	0.6486	0.5818	0.6373
2.2	CA	0.4517	0.5804	0.6018	0.6764	0.6011	0.6958
2.3	CA	0.504	0.5846	0.5986	0.6671	0.6647	0.7209
3.1	PA	0.4309	0.5523	0.538	0.6323	0.5879	0.6589
3.2	PA	0.4782	0.5285	0.5566	0.5729	0.5146	0.574
3.3	PA	0.3812	0.4945	0.497	0.572	0.5473	0.6036
4.1	PL	0.7656	0.8249	0.8229	0.864	0.8646	0.8866
4.2	PL	1.5888	1.7867	1.7253	1.9277	1.3074	2.0347
4.3	PL	0.9409	1.0043	1.0361	1.0879	1.0595	1.0886
5.1	WC	0.4791	0.5509	0.571	0.6123	0.5691	0.5969
5.2	WC	0.4709	0.6107	0.5807	0.7163	0.6032	0.7411
5.3	WC	0.4593	0.5459	0.5308	0.5876	0.5629	0.597

Table E.7: Inferred values of β_1, by variations on the prior distribution

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
$\hat{\beta_1}$	LoB	Snoek	Lally	Snoek	Lally	Snoek	Lally
1.1	MM	1.2497	1.4873	1.3075	1.5958	1.2576	1.5253
1.2	MM	1.3111	2.4547	1.4146	2.4519	1.5303	2.5075
1.3	MM	1.4577	3.0128	1.5287	3.0144	1.6209	2.8011
2.1	CA	1.2195	1.7028	1.2357	1.7282	1.2285	1.68
2.2	CA	1.4334	2.5556	1.4194	2.5284	1.3726	2.3668
2.3	CA	1.1486	1.6145	1.1536	1.6114	1.1815	1.551
3.1	PA	1.351	2.8381	1.3567	2.6662	1.4068	2.4101
3.2	PA	1.1099	1.2914	1.105	1.2574	1.1518	1.4711
3.3	PA	1.3734	2.7678	1.3456	2.6783	1.3542	2.4626
4.1	PL	1.1298	1.317	1.1555	1.3937	1.1737	1.4711
4.2	PL	1.6556	2.5565	1.7762	2.6642	1.5819	2.97
4.3	PL	1.2731	1.8569	1.3309	1.9894	1.4691	2.1408
5.1	WC	1.1242	1.2762	1.1545	1.3778	1.1609	1.388
5.2	WC	1.3288	2.3214	1.3646	2.3217	1.4443	2.503
5.3	WC	1.2342	1.7929	1.2393	1.7239	1.2686	1.7567

Table E.8: Inferred values of α_2, by variations on the prior distribution

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
$\hat{\alpha_2}$	LoB	Snoek	Lally	Snoek	Lally	Snoek	Lally
1.1	MM	1.3093	1.4051	1.4991	1.68	1.1532	1.7014
1.2	MM	9.4698	1.1952	8.5882	1.2394	7.7088	1.3909
1.3	MM	0.9557	0.9604	0.8559	0.9691	0.7891	1.0013
2.1	CA	1.1078	1.2688	0.5313	0.6313	0.4469	0.5368
2.2	CA	1.133	1.5127	1.3153	2.0453	1.2318	2.8785
2.3	CA	2.0925	2.1649	2.5022	2.3705	2.3933	2.2898
3.1	PA	0.442	0.5661	0.3999	0.4888	0.3369	0.4025
3.2	PA	1.3893	1.425	1.4888	1.5419	2.1892	1.4886
3.3	PA	0.9152	0.9981	0.972	1.1999	1.0037	1.5519
4.1	PL	1.7774	1.52	1.8358	1.707	2.155	1.9708
4.2	PL	2.2556	2.6427	2.4534	2.8741	2.7115	3.4028
4.3	PL	1.8373	2.3685	2.2602	3.3787	3.2987	4.455
5.1	WC	1.0366	1.0188	1.0929	1.1495	1.3147	1.6144
5.2	WC	0.6744	0.9501	0.5813	0.9617	0.4088	0.5692
5.3	WC	1.1269	1.1422	1.2254	1.2634	1.3193	1.3636

Table E.9: Inferred values of β_2, by variations on the prior distribution

		Matérn 3/2		Matérn 5/2		Sq. Exponential	
$\hat{\beta_2}$	LoB	Snoek	Lally	Snoek	Lally	Snoek	Lally
1.1	MM	1.3712	1.8887	1.5294	2.17	1.402	2.4755
1.2	MM	1.0138	1.9851	1.0105	1.4009	1.0189	1.2237
1.3	MM	1.2142	1.7445	1.3139	2.2691	1.4356	2.719
2.1	CA	1.4132	1.8947	1.0736	1.2992	1.0176	1.1735
2.2	CA	1.5341	2.8664	1.8901	4.0755	1.673	6.1015
2.3	CA	1.3925	1.5891	1.618	2.069	1.7408	2.2767
3.1	PA	1.0718	1.399	1.0341	1.2732	0.9418	1.0574
3.2	PA	1.294	1.4211	1.4159	1.5798	1.5237	2.1588
3.3	PA	1.1664	1.6642	1.2629	2.2294	1.3768	3.3471
4.1	PL	1.1734	1.3539	1.294	1.7375	1.4336	2.1014
4.2	PL	1.8487	2.6756	2.1151	3.0004	1.9258	3.7116
4.3	PL	1.702	2.7363	2.1257	3.7784	3.0205	5.0685
5.1	WC	1.0364	1.0632	1.1985	1.483	1.5441	3.0899
5.2	WC	1.3766	2.2062	1.3976	2.39	1.3532	1.9528
5.3	WC	1.2248	1.3679	1.3589	1.5461	1.5172	1.7022

F | Results: Bornheutter-Ferguson Estimators

In this Appendix, we will give the estimated development and estimated loss as used in generating the results of Section 6.3.3

Table F.1: Estimators of Development and Cumulative Loss of Triangle 1.1

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	58,703	61,272	76,947	89,983	89,784	94,038	102,083	113,189	95,935	101,971
Est. Development	1.00	0.95	0.89	0.80	0.68	0.54	0.36	0.20	0.09	0.01
Est. Loss	58,703	62,677	84,511	96,977	105,529	91,147	153,263	82,430	103,813	115,788

Table F.2: Estimators of Development and Cumulative Loss of Triangle 1.2

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	25,942	27,039	20,374	28,805	25,531	37,326	36,154	29,147	32,659	29,292
Est. Development	1.00	1.00	1.00	0.98	0.93	0.86	0.67	0.40	0.18	0.03
Est. Loss	25,942	27,109	20,498	26,685	24,758	33,857	23,262	32,455	32,439	16,558

Table F.3: Estimators of Development and Cumulative Loss of Triangle 1.3

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	77,656	72,171	75,250	90,343	89,102	96,008	91,796	81,782	97,225	98,655
Est. Development	1.00	1.00	1.00	0.99	0.98	0.94	0.87	0.70	0.40	0.07
Est. Loss	77,656	72,091	75,524	89,621	88,581	97,352	96,014	104,691	112,039	53,280

Table F.4: Estimators of Development and Cumulative Loss of Triangle 2.1

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	7,115	8,425	8,991	8,549	7,853	9,980	9,583	9,038	8,719	12,434
Est. Development	1.00	1.00	1.00	0.99	0.96	0.91	0.84	0.73	0.57	0.33
Est. Loss	7,115	8,428	9,018	8,460	7,850	9,928	10,028	9,972	10,726	10,594

Table F.5: Estimators of Development and Cumulative Loss of Triangle 2.2

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	193,499	200,480	225,430	206,719	214,524	233,365	244,280	240,184	237,368	230,775
Est. Development	1.00	0.98	0.98	0.97	0.95	0.92	0.86	0.76	0.59	0.31
Est. Loss	193,499	203,117	228,458	212,130	218,991	238,821	247,230	255,548	258,646	249,556

Table F.6: Estimators of Development and Cumulative Loss of Triangle 2.3

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	791	553	479	936	2,254	4,458	6,117	4,480	6,273	5,322
Est. Development	1.00	1.00	0.99	0.99	0.97	0.86	0.79	0.64	0.48	0.17
Est. Loss	791	551	488	950	1,970	3,910	5,926	4,266	5,884	3,768

Table F.7: Estimators of Development and Cumulative Loss of Triangle 3.1

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	20,739	28,009	28,826	31,272	30,780	34,035	35,226	35,400	33,316	33,661
Est. Development	1.00	1.00	1.00	1.00	0.99	0.97	0.92	0.84	0.69	0.40
Est. Loss	20,739	28,008	28,820	31,138	30,794	34,115	34,890	35,158	36,247	36,442

Table F.8: Estimators of Development and Cumulative Loss of Triangle 3.2

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	63,835	77,029	79,308	72,579	62,458	97,787	110,038	111,598	133,522	113,371
Est. Development	1.00	1.00	1.00	0.96	0.89	0.85	0.77	0.65	0.49	0.25
Est. Loss	63,835	77,154	79,579	73,898	69,471	108,949	124,464	150,117	179,219	166,331

Table F.9: Estimators of Development and Cumulative Loss of Triangle 3.3

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	6,815,646	7,721,911	8,394,117	8,288,143	90,26,329	9,673,610	10,375,605	10,512,108	10,387,245	10,165,481
Est. Development	1.00	1.00	1.00	0.99	0.98	0.96	0.92	0.85	0.71	0.40
Est. Loss	6,815,646	7,719,135	8,396,402	8,290,064	9,048,159	9,729,040	10,414,583	10,589,695	10,731,391	12,077,363

Table F.10: Estimators of Development and Cumulative Loss of Triangle 4.1

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	36,358	41,189	35,037	19,233	19,213	16,247	27,086	25,216	11,841	5,049
Est. Development	1.00	1.00	0.94	0.88	0.76	0.62	0.46	0.31	0.16	0.05
Est. Loss	36,358	41,167	33,728	21,585	30,142	19,845	59,670	34,562	42,012	10,328

Table F.11: Estimators of Development and Cumulative Loss of Triangle 4.2

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	32,430	34,292	55,502	53,586	40,504	93,531	106,357	101,142	103,677	129,300
Est. Development	1	0.98	0.96	0.95	0.86	0.69	0.55	0.28	0.13	0.05
Est. Loss	32,430	34,476	43,444	37,628	26,231	85,281	98,243	78,314	13,191	104,953

Table F.12: Estimators of Development and Cumulative Loss of Triangle 4.3

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	21,368	21,119	23,279	26,380	21,658	11,598	12,353	10,883	14,100	17,336
Est. Development	1.00	0.99	0.98	0.96	0.91	0.85	0.71	0.53	0.33	0.18
Est. Loss	21,368	20,988	22,815	25,083	21,089	11,921	16,445	15,925	17,353	23,114

Table F.13: Estimators of Development and Cumulative Loss of Triangle 5.1

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	325,322	277,574	263,000	248,319	168,844	90,686	94,730	91,161	49,255	2,909
Est. Development	1.00	0.99	0.95	0.93	0.89	0.84	0.77	0.66	0.49	0.22
Est. Loss	325,322	277,104	265,993	254,406	183,880	114,889	129,013	123,220	57,261	6,347

Table F.14: Estimators of Development and Cumulative Loss of Triangle 5.2

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	75,655	71,001	50,385	42,887	37,684	27,752	21,102	14,678	13,544	20,734
Est. Development	1.00	0.99	0.98	0.98	0.96	0.94	0.87	0.76	0.54	0.24
Est. Loss	75,655	71,192	50,817	42,739	39,511	23,007	18,781	17,490	17,604	25,048

Table F.15: Estimators of Development and Cumulative Loss of Triangle 5.3

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Ult. Loss	125,049	149,252	196,092	229,642	235,831	224,868	190,572	153,299	126,403	111,592
Est. Development	1.00	0.99	0.97	0.95	0.93	0.88	0.81	0.70	0.53	0.20
Est. Loss	125,049	149,157	192,076	222,433	229,174	221,338	194,013	152,589	136,506	207,938