Chain-ladder method: dynamic run-off uncertainty analysis

Mario V. Wüthrich*†

March 2, 2016

Abstract

We review the claims run-off uncertainty analysis derived for the chain-ladder reserving method. In a first step, we consider the total prediction uncertainty using the conditional mean square error of prediction. In a second step, we describe how this total prediction uncertainty is released dynamically over time. This provides a run-off of uncertainty pattern which allows to determine a market-value margin that can be used for market-consistent valuation and for risk-based solvency considerations.

Keywords. Chain-ladder method; claims reserving uncertainty; claims development result; Mack's formula; Merz-Wüthrich's formula; conditional mean square error of prediction; run-off.

1 Introduction

The aim of this contribution is to review the claims run-off uncertainty analysis derived for the chain-ladder (CL) reserving method. Originally, the CL method was introduced as an algorithm to set reserves for outstanding loss liabilities in insurance. This algorithm is simple and can be calculated in a spread sheet. The resulting CL reserves then serve as a point predictor for the outstanding loss liabilities. In order to quantify the prediction uncertainty in these CL reserves one needs a stochastic model foundation for the CL algorithm. Mack's [7] contribution in 1993 is considered to be a cornerstone in stochastic claims reserving modeling. It introduces the distribution-free CL model which gives a stochastic model foundation to the CL algorithm and it allows to quantify prediction uncertainty within these stochastic model assumptions. The main achievement of Mack was that he provided a conditional mean square error of prediction (MSEP) formula for the total prediction uncertainty in the CL reserves over the entire run-off of the outstanding loss liabilities. We call this the *static long-term view*.

During the financial crisis 2000-2001 (recession in European Union, dot-com bubble, 9/11 terrorist attack) insurance supervision has started new initiatives with the aim of improving the financial stability and risk management practice in the insurance industry. During these developments it became apparent that the static long-term view is not sufficient because claims

*ETH Zurich, RiskLab, Department of Mathematics, 8092 Zurich, Switzerland

†Swiss Finance Institute SFI Professor

reserves are updated periodically, for instance, whenever insurance companies close their balance sheets claims reserves are updated according to the latest available information. This viewpoint has generated a whole stream of new research, we refer to contributions [3, 4, 1, 8, 5, 9, 11]. These contributions study the change in claims prediction when additional information of a new accounting year becomes available. The updated prediction will (hopefully) fluctuate around the previous one and this new literature determines the potential size of these fluctuations. This viewpoint is called the *short-term view* or the *one-year view* and its consideration is motivated by the fact that it describes an important risk bearing position in the profit-and-loss statements of insurance companies. These days the 2008 paper [9] is a crucial module of solvency testing and in a 2014 industry discussion article [2] it is stated that almost all questions are solved for the CL method.
Is it really the case that almost all questions are solved for the CL method? Do we understand how the short-term view is related to the static long-term view? Two recent contributions [13, 10] show that this has not been the case and the aim of this review is to discuss the latest results for the CL method. The two papers [13, 10] show that claims prediction should be understood as a dynamic process. The sum over all innovations of this dynamic process exactly leads to the static long-term view, and the first innovation of this dynamic process to the short-term (one-year) view. Having this picture in mind, paper [10] shows how the total prediction uncertainty of the static long-term view is allocated to the different run-off periods. We call this the *dynamic view* and, in particular, it provides a run-off of uncertainty pattern which is a central component in market-consistent valuation of insurance products and should be contrasted to the run-off of expected reserves pattern.

Outline. In the next section we formally introduce the static long-term view, the short-term view and the dynamic view. Corollary 2.1 explains the connection between these three views. In Section 3 we introduce the gamma-gamma Bayesian chain-ladder (BCL) model. This model produces the CL reserves in the non-informative prior limit. For this model we provide the three uncertainty views and we compare them to Mack's formula (static-long term view) and to Merz-Wüthrich's (MW's) formula (short term view). Moreover, in Subsection 3.3 we complete this picture by the dynamic view. In Section 4 we provide a numerical example and we explain the R library ChainLadder [6]. Finally, in Section 5 we conclude.

2 Static long-term view, short-term view and dynamic view

We assume that the cumulative claim of accident year $i \in \{1, \dots, I\}$ and development year $j \in \{0, \dots, J\}$ is denoted by $C_{i,j}$. Throughout we assume that $I > J$ are fixed and that cumulative claims have finite first moments. Accident year i denotes the year of claims occurrence and development year j refers to the settlement delay of these cumulative claims. A general assumption is that there is a maximal settlement delay J within which all claims are settled, thus, the ultimate claim $C_{i,J}$ denotes the total claim amount of accident year i. At time $t = 0, \dots, I+J$

we have observed cumulative claims

$$\mathcal{D}_t = \{C_{i,j};\ i+j \le t,\ 1 \le i \le I,\ 0 \le j \le J\},$$

and we aim at predicting the future cumulative claims

$$\mathcal{D}_t^c = \{C_{i,j};\ i+j > t,\ 1 \le i \le I,\ 0 \le j \le J\}.$$

For detailed background information on this prediction problem and the meaning of cumulative claims we refer to [14].

The observed cumulative claims $\mathcal{D}_t$, $0 \le t \le I+J$, define a flow of information. Based on the latest available information $\mathcal{D}_t$ at time t we predict the ultimate claim $C_{i,J}$ of accident year $1 \le i \le I$ by the conditional expectation

$$\widehat{C}_{i,J}^{(t)} = \mathbb{E}\left[C_{i,J} \middle| \mathcal{D}_t \right]. \tag{2.1}$$

This predictor $\widehat{C}_{i,J}^{(t)}$ of the ultimate claim $C_{i,J}$ is optimal at time t in the sense that it minimizes the conditional MSEP among all $\sigma(\mathcal{D}_t)$-measurable predictors, and its conditional MSEP is given by

$$\mathrm{msep}_{C_{i,J}|\mathcal{D}_t}\left(\widehat{C}_{i,J}^{(t)}\right) = \mathbb{E}\left[\left(C_{i,J} - \widehat{C}_{i,J}^{(t)} \right)^2 \middle| \mathcal{D}_t \right]. \tag{2.2}$$

Here, we additionally assume existence of the right-hand side of (2.2). Conditional MSEP (2.2) gives a risk measure of the total prediction uncertainty at time t of $\widehat{C}_{i,J}^{(t)}$ over the *entire lifetime* of the claim $C_{i,J}$. This exactly corresponds to the *static long-term view* because it measures prediction uncertainty over the entire settlement period from t to $i+J$ (supposed that $t < i+J$, otherwise the ultimate claim $C_{i,J}$ is fully settled and observed at time t).

If we understand prediction as a dynamic process with periodic updates we obtain a sequence of predictors $(\widehat{C}_{i,J}^{(t)})_{0 \le t \le I+J}$. The sequence of predictors defined by (2.1) has the martingale property which implies that for all $0 \le t \le I+J-1$

$$\mathbb{E}\left[\widehat{C}_{i,J}^{(t+1)} \middle| \mathcal{D}_t \right] = \widehat{C}_{i,J}^{(t)}. \tag{2.3}$$

This martingale property (2.3) has the interpretation that in average tomorrow's prediction $\widehat{C}_{i,J}^{(t+1)}$ meets today's value $\widehat{C}_{i,J}^{(t)}$ (unbiasedness). The *short-term view* aims at studying the fluctuation in this one-period update. Therefore, we introduce the claims development result at time $t+1$ of accident year i defined by

$$\mathrm{CDR}_{i,t+1} = \widehat{C}_{i,J}^{(t)} - \widehat{C}_{i,J}^{(t+1)}.$$

This claims development result refers to a position in the profit-and-loss statements of insurance companies. They are facing a loss in this position in case of a negative claims development result and a gain for a positive claims development result. Martingale property (2.3) implies that

$$\mathbb{E}\left[\mathrm{CDR}_{i,t+1} \middle| \mathcal{D}_t \right] = 0.$$

Therefore, we predict the claims development result $\mathrm{CDR}_{i,t+1}$ at time t by 0 and its prediction uncertainty is analyzed by the following conditional MSEP (short-term view)

$$\mathrm{msep}_{\mathrm{CDR}_{i,t+1}|\mathcal{D}_t}(0) = \mathbb{E}\left[\left(\mathrm{CDR}_{i,t+1} - 0\right)^2 \Big| \mathcal{D}_t\right]. \tag{2.4}$$

This was state-of-the-art of risk-based solvency analysis at the end of 2013, and [7, 9] provide the relevant analytical formulas of the conditional MSEPs (2.2) and (2.4) for the CL method, see also [2]. Recent work [10] studies the relationship between (2.2) and (2.4). This leads to the *dynamic view* of prediction uncertainty. At time $t = i + J$ the ultimate claim $C_{i,J}$ is fully observed and we have $\widehat{C}_{i,J}^{(i+J)} = C_{i,J}$. This allows to rewrite the total prediction error at time $t < i + J$ as a telescoping sum

$$\widehat{C}_{i,J}^{(t)} - C_{i,J} = \widehat{C}_{i,J}^{(t)} - \widehat{C}_{i,J}^{(i+J)} = \sum_{s=t}^{i+J-1} \widehat{C}_{i,J}^{(s)} - \widehat{C}_{i,J}^{(s+1)} = \sum_{s=t}^{i+J-1} \mathrm{CDR}_{i,s+1}.$$

This identity expresses the total prediction error as a sum of innovations described by the claims development results $\mathrm{CDR}_{i,s+1}$, $t \le s \le i + J - 1$. Or in other words, it allocates the total prediction error at time $t < i + J$ to the different future accounting periods $t + 1 \le s + 1 \le i + J$. This gives a dynamic interpretation how prediction errors manifest over time. Moreover, martingale property (2.3) has the nice consequence that these innovations (claims development results) are uncorrelated, which immediately implies the following corollary.

Corollary 2.1. *Choose $t < i + J$. Assume that the second conditional moment of $C_{i,J}$, given $\mathcal{D}_t$, exists. We have for the ultimate claim predictors defined by* (2.1)

$$\mathrm{msep}_{C_{i,J}|\mathcal{D}_t}\left(\widehat{C}_{i,J}^{(t)}\right) = \sum_{s=t}^{i+J-1} \mathrm{Var}\left(\mathrm{CDR}_{i,s+1} | \mathcal{D}_t\right) = \sum_{s=t}^{i+J-1} \mathbb{E}\left[\mathrm{msep}_{\mathrm{CDR}_{i,s+1}|\mathcal{D}_s}(0) \Big| \mathcal{D}_t\right].$$

Corollary 2.1 combines the three views: on the left-hand side we have the total prediction uncertainty of the static long-term view; the right-hand side describes how it needs to be allocated to the different future accounting periods which gives the dynamic view; and the first term on the right-hand side exactly describes the prediction uncertainty in the short-term view. The aim here is to compute all single terms under the sum on the right-hand side. This then provides the run-off of uncertainty picture. The following formula is useful for this analysis: under the assumptions of Corollary 2.1 we have for $t \le s \le i + J - 1$

$$\mathbb{E}\left[\mathrm{msep}_{\mathrm{CDR}_{i,s+1}|\mathcal{D}_s}(0) \Big| \mathcal{D}_t\right] = \mathrm{Var}\left(\widehat{C}_{i,J}^{(s+1)} \Big| \mathcal{D}_t\right) - \mathrm{Var}\left(\widehat{C}_{i,J}^{(s)} \Big| \mathcal{D}_t\right).$$

For the proof we refer to formula (1.20) in [14]. In the next section we provide the relevant formulas for the CL method.

3 Gamma-gamma Bayesian chain-ladder model

In this section we analyze all terms of Corollary 2.1 for the CL method. As mentioned in the introduction, the CL method was originally introduced as an algorithm which was not based on a stochastic model. Only later actuaries defined stochastic models which produce the CL reserves as predictors. Not surprisingly, there are different stochastic models that lead to the CL reserves. The two most popular ones are the distribution-free CL model introduced in [7] and the over-dispersed Poisson (ODP) model, see [12]. We use a different stochastic model here, namely, the gamma-gamma Bayesian chain-ladder (BCL) model. The reason for this different model choice is that we would like to preserve martingale property (2.3) because this simplifies many considerations, in particular, Corollary 2.1 only holds true under (2.3). In general, this martingale property holds true in a Bayesian context but it fails in a frequentist's set-up as soon as one needs to estimate parameters (and parameter estimation error is involved). This is the case in the distribution-free CL model and also in the ODP model, but it is not the case in the gamma-gamma BCL model. Reassuring, often the numerical results of the gamma-gamma BCL model and of the distribution-free CL model are rather close; this will be stated and analyzed more precisely in the next subsections.

3.1 Model assumptions and claims prediction

Model Assumptions 3.1 (gamma-gamma BCL model)**.** *Assume $\sigma_j > 0$ are given fixed constants for $0 \le j \le J-1$.*

(a) Conditionally, given vector $\boldsymbol{\Theta} = (\Theta_0, \ldots, \Theta_{J-1})$, $(C_{i,j})_{0 \le j \le J}$ are independent (in accident year i) Markov processes (in development year j) with conditional distributions

$$C_{i,j+1}\,\big|_{\{C_{i,j},\boldsymbol{\Theta}\}} \sim \; \Gamma\left(C_{i,j}\sigma_j^{-2}, \Theta_j \sigma_j^{-2}\right),$$

for all $1 \le i \le I$ and $0 \le j \le J-1$.

(b) The components Θ_j of $\boldsymbol{\Theta}$ are independent and $\Gamma(\gamma_j, f_j(\gamma_j - 1))$-distributed with given prior parameters $f_j > 0$ and $\gamma_j > 1$ for $0 \le j \le J-1$.

(c) $\boldsymbol{\Theta}$ and $C_{1,0}, \ldots, C_{I,0}$ are independent and $C_{i,0} > 0$, $\mathbb{P}$-a.s., for all $1 \le i \le I$.

These model assumptions allow for an explicit calculation of the claims predictors given by the conditional expectations (2.1). This is stated in the next theorem.

Theorem 3.2 (BCL predictor)**.** *Under Model Assumptions 3.1 the BCL predictor for $C_{i,n}$ with $t \ge I \ge i > t-n \ge t-J$ is given by*

$$\widehat{C}_{i,n}^{BCL(t)} = \mathbb{E}\left[\, C_{i,n} \middle|\, \mathcal{D}_t\right] = C_{i,t-i} \prod_{j=t-i}^{n-1} \widehat{f}_j^{BCL(t)},$$

with BCL factors $\widehat{f}_j^{BCL(t)}$ for $0 \le j \le J-1$ given by

$$\widehat{f}_j^{BCL(t)} = \mathbb{E}\left[\, \Theta_j^{-1} \middle|\, \mathcal{D}_t\right] = \omega_j^{(t)}\; \widehat{f}_j^{CL(t)} + (1 - \omega_j^{(t)}) f_j,$$

where we set

$$\widehat{f}_j^{CL(t)} = \frac{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j+1}}{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j}} \qquad \textit{and} \qquad \omega_j^{(t)} = \frac{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j}}{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j} + \sigma_j^2\;(\gamma_j - 1)}.$$

For the proof we refer to Theorem 2.10 in [14]; and we use notation $a \wedge b = \min\{a, b\}$.

Remarks 3.3.

- Theorem 3.2 states that we obtain a CL structure for the ultimate claim prediction of $C_{i,J}$, which for $t \geq I \geq i > t - J$ is given by

$$\widehat{C}_{i,J}^{BCL(t)} = C_{i,t-i} \prod_{j=t-i}^{J-1} \widehat{f}_j^{BCL(t)}.$$

  This ultimate claim predictor is obtained by multiplying the latest observation $C_{i,t-i}$ of accident year i at time t with the BCL factors $\widehat{f}_{t-i}^{BCL(t)}, \ldots, \widehat{f}_{J-1}^{BCL(t)}$, which project this observation $C_{i,t-i}$ to the ultimate claim. This is often called CL structure or link ratio structure.

- The BCL factor $\widehat{f}_j^{BCL(t)}$ is a credibility weighted average between the purely observation based estimate $\widehat{f}_j^{CL(t)}$ and the prior estimate f_j. The corresponding credibility weight is given by $\omega_j^{(t)} \in (0, 1)$. It crucially depends on $\gamma_j > 1$ which quantifies the degree of information contained in the prior distribution of Θ_j.

- The observation based estimate $\widehat{f}_j^{CL(t)}$ is the classical CL factor that is used in the original CL algorithm; and it is also used in the distribution-free CL model and in the ODP model.

- Define $\boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_{J-1})$. The limit $\boldsymbol{\gamma} \to 1$ means that every component of $\boldsymbol{\gamma}$ converges to 1. In the case of non-informative priors we let $\boldsymbol{\gamma} \to 1$ and obtain for all $0 \leq j \leq J - 1$

$$\lim_{\boldsymbol{\gamma} \to 1} \widehat{f}_j^{BCL(t)} = \widehat{f}_j^{CL(t)}.$$

  That is, in the non-informative prior limit the BCL factor $\widehat{f}_j^{BCL(t)}$ and the CL factor $\widehat{f}_j^{CL(t)}$ coincide. Therefore, we obtain in the non-informative prior limit the classical CL predictor for $t \geq I \geq i > t - J$ given by

$$\widehat{C}_{i,J}^{CL(t)} = \lim_{\boldsymbol{\gamma} \to 1} \widehat{C}_{i,J}^{BCL(t)}.$$

  It is exactly this analogy that allows us to use the non-informative prior gamma-gamma BCL model as a stochastic model that supports the CL algorithm, and we use this stochastic representation to analyze the prediction uncertainty in the CL method.

3.2 Prediction uncertainty: static long-term view

Under Model Assumptions 3.1 we can explicitly calculate the prediction uncertainty of the static long-term view given by (2.2). Here, we benefit of having a Bayesian model with conjugate priors. We define the second order terms for $0 \leq j \leq J - 1$ and $t > j + 1$, subject to existence (we will further comment on this below), by

$$\Psi_j^{(t)} = \frac{\sigma_j^2}{\sigma_j^2(\gamma_j - 2) + \sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j}},$$

and their non-informative prior limits by

$$\overline{\Psi}_j^{(t)} \stackrel{\text{def.}}{=} \lim_{\gamma \to 1} \Psi_j^{(t)} = \frac{\sigma_j^2}{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j} - \sigma_j^2}.$$

Note that these second orders term are observable at time $t-1$. We have the following theorem.

Theorem 3.4 (static long-term view)**.** *Under Model Assumptions 3.1 the BCL predictors satisfy in the non-informative prior limit $\gamma \to 1$ for $t \geq I \geq i > t - J$*

$$\begin{aligned}\text{msep}_{C_{i,J}|\mathcal{D}_t}\left(\widehat{C}_{i,J}^{CL(t)}\right) &= \widehat{C}_{i,J}^{CL(t)} \sum_{j=t-i}^{J-1} \sigma_j^2 \prod_{m=j}^{J-1} \left(\widehat{f}_m^{CL(t)} \left(1 + \overline{\Psi}_m^{(t)}\right)\right) \\ &\quad + \left(\widehat{C}_{i,J}^{CL(t)}\right)^2 \left(\prod_{j=t-i}^{J-1} \left(1 + \overline{\Psi}_j^{(t)}\right) - 1\right),\end{aligned}$$

under the assumption that $\sum_{\ell=1}^{t-j-1} C_{\ell,j}/\sigma_j^2 > 1$ for all $t - i \leq j \leq J - 1$; otherwise the second moment is infinite. For aggregated accident years the conditional MSEP is given by

$$\begin{aligned}\text{msep}_{\sum_{i=t-J+1}^{I} C_{i,J}|\mathcal{D}_t}\left(\sum_{i=t-J+1}^{I} \widehat{C}_{i,J}^{CL(t)}\right) &= \sum_{i=t-J+1}^{I} \text{msep}_{C_{i,J}|\mathcal{D}_t}\left(\widehat{C}_{i,J}^{CL(t)}\right) \\ &\quad + 2 \sum_{t-J+1 \leq i < n \leq I} \widehat{C}_{i,J}^{CL(t)}\, \widehat{C}_{n,J}^{CL(t)} \left(\prod_{j=t-i}^{J-1} \left(1 + \overline{\Psi}_j^{(t)}\right) - 1\right).\end{aligned}$$

Proof of Theorem 3.4. The proof follows from Theorem 3.8 in [10] by letting $\gamma \to 1$. □

Remarks 3.5.

- The second order terms $\Psi_j^{(t)}$ and $\overline{\Psi}_j^{(t)}$ were only defined subject to existence. In fact, we require that their denominators are strictly positive, which is exactly the necessary and sufficient condition for having finite conditional MSEPs.

- Theorem 3.4 provides the prediction uncertainty of the static long-term view in the non-informative prior limit $\gamma \to 1$. The results for informative priors are completely similar, one only needs to replace all variables $\widehat{C}_{i,J}^{CL(t)}$, $\widehat{f}_m^{CL(t)}$ and $\overline{\Psi}_j^{(t)}$ by $\widehat{C}_{i,J}^{BCL(t)}$, $\widehat{f}_m^{BCL(t)}$ and $\Psi_j^{(t)}$, respectively.

- Obviously, our conditional MSEP formula differs from Mack's formula [7]. Our uncertainty formula is *exact* in the non-informative prior gamma-gamma BCL model and Mack's formula is an *estimate* in the distribution-free CL model. In the next step we are going to show, how the two uncertainty formulas are related to each other.

In the sequel we make the following assumption (second relationship "$\gg$"): for $t \geq I$ and all $0 \leq j \leq J-1$

$$\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j} \geq \sum_{\ell=1}^{I-j-1} C_{\ell,j} \gg \sigma_j^2. \tag{3.1}$$

Note that the first inequality always holds true for non-negative cumulative claims. This assumption is sufficient for having finite conditional MSEPs in Theorem 3.4 for all time points $t \geq I$. Moreover, (3.1) implies for $t \geq I$ relationship

$$0 < \frac{\sigma_j^2}{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j}} < \overline{\Psi}_j^{(t)} \leq \frac{\sigma_j^2}{\sum_{\ell=1}^{I-j-1} C_{\ell,j} - \sigma_j^2} \ll 1. \tag{3.2}$$

Applying this to the first term in Theorem 3.4 gives approximation

$$\widehat{C}_{i,J}^{CL(t)} \sum_{j=t-i}^{J-1} \sigma_j^2 \prod_{m=j}^{J-1} \left(\widehat{f}_m^{CL(t)} \left(1 + \overline{\Psi}_m^{(t)}\right)\right) \approx \widehat{C}_{i,J}^{CL(t)} \sum_{j=t-i}^{J-1} \sigma_j^2 \prod_{m=j}^{J-1} \widehat{f}_m^{CL(t)} = \left(\widehat{C}_{i,J}^{CL(t)}\right)^2 \sum_{j=t-i}^{J-1} \frac{\sigma_j^2}{\widehat{C}_{i,j}^{CL(t)}}.$$

In fact, the right-hand side is a *lower bound* for the left-hand side. For the second term in Theorem 3.4 we have under (3.1) the following approximation

$$\left(\widehat{C}_{i,J}^{CL(t)}\right)^2 \left(\prod_{j=t-i}^{J-1} \left(1 + \overline{\Psi}_j^{(t)}\right) - 1 \right) \approx \left(\widehat{C}_{i,J}^{CL(t)}\right)^2 \sum_{j=t-i}^{J-1} \overline{\Psi}_j^{(t)} \approx \left(\widehat{C}_{i,J}^{CL(t)}\right)^2 \sum_{j=t-i}^{J-1} \frac{\sigma_j^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}}.$$

The right-hand side is again a *lower bound* for the left-hand side. In [7] one uses a slightly different parametrization for the variance parameters σ_j^2, namely, one considers $s_j^2/(\widehat{f}_j^{CL(t)})^2$ instead. Therefore, we identify for $0 \leq j \leq J-1$

$$\sigma_j^2 = s_j^2/(\widehat{f}_j^{CL(t)})^2. \tag{3.3}$$

Making this change of variables we exactly obtain Mack's formula [7] which provides approximation under (3.1) and lower bound (in any case)

$$\begin{aligned} \text{msep}_{C_{i,J}|\mathcal{D}_t}\left(\widehat{C}_{i,J}^{CL(t)}\right) &\approx \text{msep}^{\text{Mack}}_{C_{i,J}|\mathcal{D}_t}\left(\widehat{C}_{i,J}^{CL(t)}\right) \\ &= \left(\widehat{C}_{i,J}^{CL(t)}\right)^2 \sum_{j=t-i}^{J-1} \left[\frac{s_j^2/(\widehat{f}_j^{CL(t)})^2}{\widehat{C}_{i,j}^{CL(t)}} + \frac{s_j^2/(\widehat{f}_j^{CL(t)})^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}} \right]. \end{aligned} \tag{3.4}$$

Exactly the same arguments apply for aggregated accident years and one finds approximation under (3.1) and lower bound (in any case)

$$\begin{aligned} \text{msep}_{\sum_{i=t-J+1}^{I} C_{i,J}|\mathcal{D}_t}\left(\sum_{i=t-J+1}^{I} \widehat{C}_{i,J}^{CL(t)}\right) &\approx \text{msep}^{\text{Mack}}_{\sum_{i=t-J+1}^{I} C_{i,J}|\mathcal{D}_t}\left(\sum_{i=t-J+1}^{I} \widehat{C}_{i,J}^{CL(t)}\right) \\ &= \sum_{i=t-J+1}^{I} \text{msep}^{\text{Mack}}_{C_{i,J}|\mathcal{D}_t}\left(\widehat{C}_{i,J}^{CL(t)}\right) \\ &\quad + 2 \sum_{t-J+1 \leq i < n \leq I} \widehat{C}_{i,J}^{CL(t)} \widehat{C}_{n,J}^{CL(t)} \sum_{j=t-i}^{J-1} \frac{s_j^2/(\widehat{f}_j^{CL(t)})^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}}. \end{aligned} \tag{3.5}$$

Let us briefly conclude on these findings. As mentioned in Remarks 3.5 the two stochastic models differ and also the resulting conditional MSEP formulas differ. But (3.4) and (3.5) show that the resulting values for the prediction uncertainties are rather similar under assumption (3.1). This assumption is often fulfilled for non-life insurance data. Below we will explicitly analyze this difference numerically for an example.

The distribution-free CL model looks at claims prediction from a frequentist's viewpoint. This viewpoint allows to separate the so-called process uncertainty (blue terms in (3.4) and (3.5)) from the parameter estimation uncertainty (red terms in (3.4) and (3.5)). This separation will be important to understand the short-term view. For a more detailed explanation of these two uncertainty terms we refer to Section 1.5 in [14].

3.3 Prediction uncertainty: short-term view

Under Model Assumptions 3.1 we can also explicitly calculate the prediction uncertainty of the short-term view given by (2.4). For this we again benefit of having a Bayesian model with conjugate priors. We define the credibility weights for $0 \le j \le J-1$ and $I \ge t-j \ge 1$ by

$$\alpha_j^{(t)} = \frac{C_{t-j,j}}{\sigma_j^2\ (\gamma_j - 1) + \sum_{\ell=1}^{t-j} C_{\ell,j}},$$

and their non-informative prior limits by

$$\overline{\alpha}_j^{(t)} = \lim_{\gamma \to 1} \alpha_j^{(t)} = \frac{C_{t-j,j}}{\sum_{\ell=1}^{t-j} C_{\ell,j}} \ \in\ (0,1].$$

We have the following theorem.

Theorem 3.6 (short-term view)**.** *Under Model Assumptions 3.1 the BCL predictors satisfy in the non-informative prior limit* $\boldsymbol{\gamma} \to 1$ *for* $t \ge I \ge i > t-J$

$$\mathrm{msep}_{\mathrm{CDR}_{i,t+1}|\mathcal{D}_t}(0) = (\widehat{C}_{i,J}^{CL(t)})^2\ \left[\left(1 + \frac{\sigma_{t-i}^2}{C_{i,t-i}}\right)\left(1 + \overline{\Psi}_{t-i}^{(t)}\right)\prod_{j=t-i+1}^{J-1}\left(1 + \overline{\alpha}_j^{(t)}\overline{\Psi}_j^{(t)}\right) - 1\right],$$

where we assume that $\sum_{\ell=1}^{t-j-1} C_{\ell,j}/\sigma_j^2 > 1$ *for all* $t-i \le j \le J-1$*, otherwise the corresponding conditional MSEP is infinite. For aggregated accident years the conditional MSEP is given by*

$$\begin{aligned}\mathrm{msep}_{\sum_{i=t-J+1}^{I}\mathrm{CDR}_{i,t+1}|\mathcal{D}_t}(0) &= \sum_{i=t-J+1}^{I}\mathrm{msep}_{\mathrm{CDR}_{i,t+1}|\mathcal{D}_t}(0)\\ &\quad + 2\sum_{t-J+1\le i<n\le I}\widehat{C}_{i,J}^{CL(t)}\widehat{C}_{n,J}^{CL(t)}\ \left[\left(1 + \overline{\Psi}_{t-i}^{(t)}\right)\prod_{j=t-i+1}^{J-1}\left(1 + \overline{\alpha}_j^{(t)}\overline{\Psi}_j^{(t)}\right) - 1\right].\end{aligned}$$

Proof of Theorem 3.6. The proof follows from Theorem 3.10 and Corollary A.4 in [10] by letting $\boldsymbol{\gamma} \to 1$. □

Remarks 3.5 also apply to the short-term view. We compare the result of Theorem 3.6 to Merz-Wüthrich's (MW's) formula which was derived in the distribution-free CL model, see [9]. We do this under assumption (3.1) and we need a second assumption, namely, for $0 \le t-i \le J-1$

$$C_{i,t-i} \gg \sigma_{t-i}^2. \tag{3.6}$$

Assumptions (3.1) and (3.6) imply

$$0 < \left(1 + \frac{\sigma_{t-i}^2}{C_{i,t-i}}\right)\left(1 + \overline{\Psi}_{t-i}^{(t)}\right) - 1\ \approx\ \frac{\sigma_{t-i}^2}{C_{i,t-i}} + \overline{\Psi}_{t-i}^{(t)} \ll 1,$$

where again the approximation on the right-hand side provides also a lower bound on the left-hand side. Moreover, relation (3.2) also holds true if we multiply all terms with credibility weights $\overline{\alpha}_j^{(t)} \in (0,1)$. This implies that the first statement of Theorem 3.6 can be approximated and bounded below by (we use identification (3.3))

$$\begin{aligned} \mathrm{msep}_{\mathrm{CDR}_{i,t+1}|\mathcal{D}_t}(0) \;&\approx\; \varrho_{i,t+1}^{(t)} \stackrel{\text{def.}}{=} \mathrm{msep}^{\mathrm{MW}}_{\mathrm{CDR}_{i,t+1}|\mathcal{D}_t}(0) \\ &= (\widehat{C}_{i,J}^{CL(t)})^2 \left[\frac{s_{t-i}^2/(\widehat{f}_{t-i}^{CL(t)})^2}{C_{i,t-i}} + \frac{s_{t-i}^2/(\widehat{f}_{t-i}^{CL(t)})^2}{\sum_{\ell=1}^{i-1} C_{\ell,t-i}} + \sum_{j=t-i+1}^{J-1} \overline{\alpha}_j^{(t)} \frac{s_j^2/(\widehat{f}_j^{CL(t)})^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}} \right]. \end{aligned} \tag{3.7}$$

This is exactly MW's formula [9] of the short-term view prediction uncertainty.

Remark 3.7. If we compare the static long-term view (3.4) to the short-term view (3.7) we observe that from the (blue) process uncertainty terms exactly the first term with index $j = t - i$ of the static long-term view appears in the short term view. For the (red) parameter estimation uncertainty terms the picture is slightly different, the first term with index $j = t - i$ appears in both views and the remaining terms with indexes $t - i + 1 \le j \le J - 1$ are scaled in the short-term view with credibility weights $\overline{\alpha}_j^{(t)} \in (0,1]$ compared to the static long-term view. These scalings reflect the reduction in parameter estimation uncertainty by the arrival of the new observations in $\mathcal{D}_{t+1} \setminus \mathcal{D}_t$.

For aggregated accident years we obtain completely analogously approximation and lower bound

$$\begin{aligned} \mathrm{msep}_{\sum_{i=t-J+1}^{I} \mathrm{CDR}_{i,t+1}|\mathcal{D}_t}(0) \;&\approx\; \varrho_{t+1}^{(t)} \stackrel{\text{def.}}{=} \mathrm{msep}^{\mathrm{MW}}_{\sum_{i=t-J+1}^{I} \mathrm{CDR}_{i,t+1}|\mathcal{D}_t}(0) \\ &= \sum_{i=t-J+1}^{I} \mathrm{msep}^{\mathrm{MW}}_{\mathrm{CDR}_{i,t+1}|\mathcal{D}_t}(0) \\ &\quad + 2 \sum_{t-J+1 \le i < n \le I} \widehat{C}_{i,J}^{CL(t)} \widehat{C}_{nJ}^{CL(t)} \left[\frac{s_{t-i}^2/(\widehat{f}_{t-i}^{CL(t)})^2}{\sum_{\ell=1}^{i-1} C_{\ell,t-i}} + \sum_{j=t-i+1}^{J-1} \overline{\alpha}_j^{(t)} \frac{s_j^2/(\widehat{f}_j^{CL(t)})^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}} \right]. \end{aligned} \tag{3.8}$$

3.4 Prediction uncertainty: dynamic view

Although the relationship between the non-informative prior gamma-gamma BCL model and the distribution-free CL model has not been worked out so clearly in the actuarial literature, the previous results of Subsections 3.2 and 3.3 are generally known and well-established in the actuarial community; this is also what the industry discussion article [2] was referring to. The interpretation in Remark 3.7 explains how the short-term view is related to the static long-term view. But from this interpretation we can extract much more which leads to the right intuition how these two views can be completed by the dynamic view! This exactly motivates the results derived in [10]. Before we give these new results we would like to explain the intuition that leads to the dynamic view.

The aim is to derive the second next term $\mathbb{E}[\mathrm{msep}_{\mathrm{CDR}_{i,t+2}|\mathcal{D}_{t+1}}(0)|\mathcal{D}_t]$ of the decomposition provided by Corollary 2.1. If we compare (3.4) and (3.7) the natural guess is that it should involve the next (blue) processes uncertainty term of the static long-term view (3.4) which is given by

$$\frac{s^2_{t-i+1}/(\widehat{f}^{CL(t)}_{t-i+1})^2}{\widehat{C}^{CL(t)}_{i,t-i+1}}.$$

From the (red) parameter estimation uncertainty terms it should involve the next term with index $j = t - i + 1$, but only the part that has not been treated in the previous period $t + 1$, that is,

$$\left(1 - \overline{\alpha}^{(t)}_{t-i+1}\right) \frac{s^2_{t-i+1}/(\widehat{f}^{CL(t)}_{t-i+1})^2}{\sum_{\ell=1}^{i-2} C_{\ell,t-i+1}},$$

and all remaining (red) parameter estimation uncertainty terms are reduced by the new incoming information in $\mathcal{D}_{t+2} \setminus \mathcal{D}_{t+1}$ which provides additional terms

$$\sum_{j=t-i+2}^{J-1} \overline{\alpha}^{(t)}_{j-1} \left(1 - \overline{\alpha}^{(t)}_{j}\right) \frac{s^2_j/(\widehat{f}^{CL(t)}_j)^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}}.$$

This is what has been shown in [10] and, moreover, this idea can be iterated to all future periods which exactly provides the dynamic (run-off) view.
We define for $0 \le k \le J - 1$ and $t \ge I$ and $I \ge i > t - J + k$

$$\begin{aligned} \varrho^{(t)}_{i,t+1+k} \;&=\; \left(\widehat{C}^{CL(t)}_{i,J}\right)^2 \left[\frac{\sigma^2_{t-i+k}}{\widehat{C}^{CL(t)}_{i,t-i+k}} + \prod_{m=1}^{k} \left(1 - \overline{\alpha}^{(t)}_{t-i+m}\right) \frac{\sigma^2_{t-i+k}}{\sum_{\ell=1}^{i-1-k} C_{\ell,t-i+k}} \right] \\ &\quad + \left(\widehat{C}^{CL(t)}_{i,J}\right)^2 \sum_{j=t-i+1+k}^{J-1} \left[\overline{\alpha}^{(t)}_{j-k} \prod_{m=0}^{k-1} \left(1 - \overline{\alpha}^{(t)}_{j-m}\right) \frac{\sigma^2_j}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}} \right], \end{aligned} \tag{3.9}$$

and for aggregated accident years, where summations $i < n$ run over $t - J + 1 + k \le i < n \le I$,

$$\begin{aligned} \varrho^{(t)}_{t+1+k} \;&=\; \sum_{i=t-J+1+k}^{I} \varrho^{(t)}_{i,t+1+k} + 2 \sum_{i<n} \widehat{C}^{CL(t)}_{i,J} \widehat{C}^{CL(t)}_{n,J} \prod_{m=1}^{k} \left(1 - \overline{\alpha}^{(t)}_{t-i+m}\right) \frac{\sigma^2_{t-i+k}}{\sum_{\ell=1}^{i-1-k} C_{\ell,t-i+k}} \\ &\quad + 2 \sum_{i<n} \widehat{C}^{CL(t)}_{i,J} \widehat{C}^{CL(t)}_{n,J} \sum_{j=t-i+1+k}^{J-1} \left[\overline{\alpha}^{(t)}_{j-k} \prod_{m=0}^{k-1} \left(1 - \overline{\alpha}^{(t)}_{j-m}\right) \frac{\sigma^2_j}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}} \right]. \end{aligned} \tag{3.10}$$

In Theorem 6.4 of [10] there is an *exact* statement for the conditional expectation $\mathbb{E}[\varrho^{(t+1)}_{i,t+1+k} | \mathcal{D}_t]$. We will not give the full statement here because its formulation is a bit cumbersome, but we give the approximation and lower bound which is formulated as the next property, for the derivation we refer to Section 6.4 in [10] and Section 2.4.2 in [14].

Property 3.8 (dynamic view)**.** *Assume that Model Assumptions 3.1 hold and that the BCL predictors are considered in the non-informative prior limit* $\gamma \to 1$. *We have approximation under assumptions* (3.1) *and* (3.6) *and lower bound (in any case) for* $1 \le k \le J - 1$, $t \ge I$ *and* $I \ge i > t - J + k$

$$\mathbb{E}\left[\varrho^{(t+1)}_{i,t+1+k} \middle| \mathcal{D}_t \right] \;\approx\; \varrho^{(t)}_{i,t+1+k},$$

and for aggregated accident years

$$\mathbb{E}\left[\varrho^{(t+1)}_{t+1+k} \middle| \mathcal{D}_t \right] \;\approx\; \varrho^{(t)}_{t+1+k}.$$

Property 3.8 implies that in the non-informative prior gamma-gamma BCL model under assumptions (3.1) and (3.6) we can approximate any term in the sum of Corollary 2.1, and we obtain for $t \geq I$ and $0 \leq k < I + J - t$

$$\mathbb{E}\left[\mathrm{msep}_{\sum_{i=t-J+1+k}^{I} \mathrm{CDR}_{i,t+1+k} | \mathcal{D}_{t+k}} (0) \middle| \mathcal{D}_t \right] \;\approx\; \varrho_{t+1+k}^{(t)}, \tag{3.11}$$

where in fact the approximation is a *lower bound* for each index k. The same statements apply for single accident years $I \geq i > t - J$. As a result, the sequence

$$\varrho_{t+1}^{(t)}, \ldots, \varrho_{I+J}^{(t)}$$

provides the *run-off of uncertainty profile* at time t (called dynamic view) in the non-informative prior gamma-gamma BCL model. There remains the question about the approximation error in (3.11). Fortunately, we know that $\varrho_{t+1+k}^{(t)}$ are lower bounds for all $k \geq 0$ which allows to uniformly control the approximation error.

Corollary 3.9. *Assume that Model Assumptions 3.1 hold and that the BCL predictors are considered in the non-informative prior limit $\gamma \to 1$ and under parametrization* (3.3). *For $I \leq t < I + J$ and $0 \leq k < I + J - t$ we have*

$$\mathbb{E}\left[\mathrm{msep}_{\sum_{i=t-J+1+k}^{I} \mathrm{CDR}_{i,t+1+k} | \mathcal{D}_{t+k}} (0) \middle| \mathcal{D}_t \right] \geq \varrho_{t+1+k}^{(t)},$$

with aggregation property

$$\mathrm{msep}^{\mathrm{Mack}}_{\sum_{i=t-J+1}^{I} C_{i,J} | \mathcal{D}_t} \left(\sum_{i=t-J+1}^{I} \widehat{C}_{i,J}^{CL(t)} \right) = \sum_{k=0}^{I+J-t-1} \varrho_{t+k+1}^{(t)}.$$

The total approximation error is given by the difference

$$\begin{aligned} &\sum_{k=0}^{I+J-t-1} \mathbb{E}\left[\mathrm{msep}_{\sum_{i=t-J+1+k}^{I} \mathrm{CDR}_{i,t+k+1} | \mathcal{D}_{t+k}} (0) \middle| \mathcal{D}_t \right] - \sum_{k=0}^{I+J-t-1} \varrho_{t+k+1}^{(t)} \\ &= \mathrm{msep}_{\sum_{i=t-J+1}^{I} C_{i,J} | \mathcal{D}_t} \left(\sum_{i=t-J+1}^{I} \widehat{C}_{i,J}^{CL(t)} \right) - \mathrm{msep}^{\mathrm{Mack}}_{\sum_{i=t-J+1}^{I} C_{i,J} | \mathcal{D}_t} \left(\sum_{i=t-J+1}^{I} \widehat{C}_{i,J}^{CL(t)} \right) \;\geq\; 0. \end{aligned}$$

Property 3.8 and Corollary 3.9 conclude the dynamic view. They tell us how to split the static long-term view across the future development periods $t + 1, \ldots, I + J$, and the resulting approximations $\varrho_{t+k+1}^{(t)}$ can be calculated explicitly. Moreover, the approximation error can be determined explicitly from Corollary 3.9, formula (3.5) and Theorem 3.4 applied to the non-informative prior limit $\gamma \to 1$. In many applied applications, as for instance the example below, this approximation error is negligible, which means that in Property 3.8 we almost have the martingale property in the upper index.

4 Example

In this section we present a numerical example and we show how the R library ChainLadder [6] can be used to obtain the dynamic view. In order to analyze an example we still need to calibrate the variance parameters $\sigma_0^2, \ldots, \sigma_{J-1}^2$ to the data. This could be done in a (full) Bayesian way by also choosing a prior distribution on these variance parameters; in this outline we prefer an empirical Bayesian approach and estimate these variance parameters directly from the data. This has the advantage of preserving analytical tractability. Under Model Assumptions 3.1 we obtain for the first two conditional moments

$$\begin{aligned}
\mathbb{E}\left[C_{i,j+1} \middle| C_{i,0}, \ldots, C_{i,j}, \boldsymbol{\Theta}\right] &= \Theta_j^{-1}\, C_{i,j}, \\
\operatorname{Var}\left(C_{i,j+1} \middle| C_{i,0}, \ldots, C_{i,j}, \boldsymbol{\Theta}\right) &= \Theta_j^{-2}\sigma_j^2\, C_{i,j},
\end{aligned}$$

for all $1 \le i \le I$ and $0 \le j \le J-1$. This implies that for known model parameters $\boldsymbol{\Theta}$ the gamma-gamma BCL model exactly fulfills the assumptions of the distribution-free CL model as defined in [7]: Θ_j^{-1} plays the role of the CL factor and $\Theta_j^{-2}\sigma_j^2$ plays the role of the variance parameter in the distribution-free CL model. This is also the motivation behind parametrization (3.3). For this reason we use the classical CL estimates and consider at time $t \ge I$ for $j < t-2$

$$\widehat{s}_j^2 = \frac{1}{((t-j-1)\wedge I) - 1} \sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j} \left(\frac{C_{\ell,j+1}}{C_{\ell,j}} - \widehat{f}_j^{CL(t)} \right)^2, \tag{4.1}$$

and if $j = J-1 = I-2 = t-2$ we set

$$\widehat{s}_{J-1}^2 = \min\left\{ \widehat{s}_{J-3}^2, \widehat{s}_{J-2}^2, \widehat{s}_{J-2}^4/\widehat{s}_{J-3}^2 \right\} = \min\left\{ \widehat{s}_{J-3}^2, \widehat{s}_{J-2}^4/\widehat{s}_{J-3}^2 \right\}.$$

Using re-parametrization (3.3) then motivates estimates at time $t \ge I$ for $0 \le j \le J-1$

$$\widehat{\sigma}_j^2 = \widehat{s}_j^2 / (\widehat{f}_j^{CL(t)})^2.$$

In the sequel we use these sample estimates $\widehat{s}_j^2$ and $\widehat{\sigma}_j^2$ for s_j^2 and σ_j^2, respectively, in all terms appearing in Corollary 3.9.

We revisit Example 2.4 of [14]. The data is given in Table 4 in Appendix A, and the parameter estimates for this data are provided in Table 1. With these parameter estimates we can now

$j=$	0	1	2	3	4	5	6	7	8	9
$\widehat{f}_j^{CL(I)}$	1.4925	1.0778	1.0229	1.0148	1.0070	1.0051	1.0011	1.0010	1.0014	
$\widehat{s}_j$	135.25	33.80	15.76	19.85	9.34	2.00	0.82	0.22	0.06	
$\widehat{\sigma}_j$	90.62	31.36	15.41	19.56	9.27	1.99	0.82	0.22	0.06	

Table 1: Estimated CL factors $\widehat{f}_j^{CL(I)}$ and standard deviation parameters $\widehat{s}_j$ and $\widehat{\sigma}_j$ at time $t = I = J+1$.

calculate all quantities of interest for the CL method. We start by defining the CL reserves. If we assume that at time $t \ge I$ the cumulative claims $C_{i,(t-i)\wedge J}$ are already settled, we can define the claims reserves of accident years $I \ge i > t-J$ for the outstanding loss liabilities by

$$\mathcal{R}_i^{CL(t)} = \widehat{C}_{i,J}^{CL(t)} - C_{i,t-i},$$

accident year i	CL reserves $\mathcal{R}_i^{CL(I)}$	BCL msep$^{1/2}$	Mack's msep$^{1/2}$	in % reserves
1	0	0	0	–
2	15'126	267	267	1.8%
3	26'257	914	914	3.5%
4	34'538	3'058	3'058	8.9%
5	85'302	7'628	7'628	8.9%
6	156'494	33'341	33'341	21.3%
7	286'121	73'467	73'467	25.7%
8	449'167	85'399	85'398	19.0%
9	1'043'242	134'338	134'337	12.9%
10	3'950'815	410'850	410'817	10.4%
total	6'047'061	462'990	462'960	7.7%

Table 2: CL reserves and prediction uncertainty in the static long-term view: rooted conditional MSEPs in the non-informative prior gamma-gamma BCL model (Theorem 3.4) and Mack's formula (3.4) and (3.5) at time $t = I = 10$.

and aggregated over all accident years by

$$\mathcal{R}^{CL(t)} = \sum_{i=t-J+1}^{I} \mathcal{R}_i^{CL(t)}.$$

In Table 2 we provide the CL reserves and the rooted conditional MSEPs of the static long-term view, that is, the rooted MSEP formula in the non-informative prior gamma-gamma BCL model (Theorem 3.4) and the rooted Mack's formula (3.4) and (3.5). As stated in Corollary 3.9 we see that Mack's formula gives a lower bound to the conditional MSEP formula in the non-informative prior gamma-gamma BCL model. The approximation error $462'960$ versus $462'990$ is fairly small and therefore we may consider approximations $\varrho_{I+1+k}^{(I)}$ for $k \geq 0$ being appropriate. Using (3.9) and (3.10) we calculate the dynamic run-off view. This describes how Mack's static long-term prediction uncertainty is expected to be released over time. We therefore consider two different quantities, the first one being the rooted expected claims development result uncertainties $\sqrt{\varrho_{I+1+k}^{(I)}}$, for $k = 0, \ldots, J-1$ at time $t = I$, and the second one being the rooted expected remaining prediction uncertainty at time $s = I, \ldots, I+J-1$ (viewed from time $t = I$) defined by

$$\chi_s^{(I)} = \sqrt{\sum_{k=s-I}^{J-1} \varrho_{I+k+1}^{(I)}}. \tag{4.2}$$

This rooted expected run-off of prediction uncertainty is compared to the expected run-off of the CL reserves given by

$$\mathbb{E}\left[\mathcal{R}^{CL(s)} \middle| \mathcal{D}_I\right] = \sum_{i=s-J+1}^{I} \widehat{C}_{i,J}^{CL(I)} - \widehat{C}_{i,s-i}^{CL(I)}. \tag{4.3}$$

Formula (4.3) can be considered as the expected run-off of the liabilities that corresponds to a first moment, and formula (4.2) can be considered as the expected run-off of the corresponding second moment. The results of the dynamic view are provided in Table 3.

calendar years s	exp. run-off CL reserves $\mathbb{E}[\mathcal{R}^{CL(s)}\|\mathcal{D}_I]$	run-off rooted MSEP $\chi_s^{(I)}$	in % reserves	rooted expected CDR MSEP $\sqrt{\varrho_{s+1}^{(I)}}$
10	6'047'061	462'960	8%	420'220
11	2'173'856	194'285	9%	150'544
12	1'048'144	122'813	12%	93'390
13	570'584	79'758	14%	72'882
14	293'063	32'397	11%	31'459
15	148'951	7'739	5%	7'172
16	67'824	2'906	4%	2'803
17	36'036	769	2%	744
18	13'655	191	1%	191
19	0	0		0

Table 3: Dynamic view: expected run-off of CL reserves (4.3), the corresponding rooted remaining prediction uncertainty $\chi_s^{(I)}$ defined in (4.2) and the rooted expected claims development result (CDR) uncertainties $\sqrt{\varrho_{s+1}^{(I)}}$ for $I \leq s \leq I+J-1$ at time $t = I = 10$.

The first line of Table 3 for $s = 10$ corresponds to the last line of Table 2, providing the CL reserves, Mack's static long-term formula (3.5), as well as MW's short-term formula (3.8). This first line is complemented by the full expected run-off picture (dynamic view) on lines $s = 11, \ldots, 19$. The last column $\sqrt{\varrho_{s+1}^{(I)}}$ provides the rooted expected claims development result uncertainties for all future calendar years s, and the second and third columns describe the expected run-off of the CL reserves and of the corresponding prediction uncertainties. Interestingly, the relative uncertainty in the fourth column is increasing until calendar year $s = 13$ (from 8% to 14%) and then decreasing. This is not a typical picture, often in non-life insurance data one finds that this column is monotonically increasing, which says (colloquially speaking) that the release of reserves is faster than the release of uncertainty.

The results of Table 3 can be obtained from the R library ChainLadder [6]. Let us briefly describe the corresponding commands. Assume that the cumulative data is stored in `data.cumulative`, then we use the following commands:

```
# bringing data in appropriate triangular form
> tri <- as.triangle(as.matrix(data.cumulative))
> dimnames(tri)=list(origin=1:nrow(tri),dev=1:ncol(tri))

# calculation of CL method using (4.1)
> M <- MackChainLadder(tri,est.sigma="Mack")
> M
```

Columns 'IBNR' and 'Mack.S.E' of the R output provide the CL reserves and the rooted conditional MSEP per accident year $1 \leq i \leq I$, see (3.4), and the summary table gives the corresponding statistics for aggregated accident years, see (3.5), in the static long-term view.

dynamic view (3.10)

```
> CDR(M,dev="all")
```

Columns 'IBNR', 'CDR(1)S.E.', ..., 'CDR($J+1$)S.E.' and 'Mack.S.E.' of the R output provide the CL reserves and the rooted expected conditional MSEPs of the claims development results (short-term view) and of the total run-off uncertainty (static long-term view). The last line of these columns of the R output denoted 'Total' exactly displays $\sqrt{\varrho_{s+1}^{(I)}}$, $I \leq s \leq I+J-1$, which gives the dynamic view.

5 Conclusions

We have revisited the CL algorithm. For analyzing prediction uncertainty of the CL algorithm we have introduced the non-informative prior gamma-gamma BCL model which provides the classical CL predictors. In this model we have studied the full dynamic run-off of uncertainty picture. The resulting uncertainty formulas were compared to Mack's formula (static long-term view) and to MW's formula (short-term view). The dynamic view allows to analyze the release of prediction uncertainty over time, which was demonstrated in an example.

References

[1] Böhm, H.; Glaab, H. Modellierung des Kalenderjahr-Risikos im additiven und multiplikativen Schadenreservierungsmodell. Presented at the German ASTIN Colloquium 2006.

[2] Dal Moro, E.; Lo, J. An industry question: the ultimate and one-year reserving uncertainty for different non-life reserving methodologies. *ASTIN Bulletin* **2014**, *44*, 495-499.

[3] De Felice, M.; Moriconi, F. Risk based capital in P&C loss reserving or stressing the triangle. *Research Group on Insurance Companies and Pension Funds* **2003**, Working Paper n. 1.

[4] De Felice, M.; Moriconi, F. Process error and estimation error of year-end reserve estimation in the distribution free chain-ladder model. *Alef Working Paper* **2006**.

[5] England, P.D. Private communication **2008**.

[6] Gesmann, M.; Murphy, D.; Zhang, W.; Carrato, A.; Crupi G.; Wüthrich, M.V.; Concina, F. ChainLadder: statistical methods and models for the calculation of outstanding claims reserves in general insurance. R package. *CRAN* **2015**, version 0.2.1.

[7] Mack, T. Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin* **1993**, *23*, 213-225.

[8] Merz, M.; Wüthrich, M.V. Prediction error of the expected claims development result in the chain ladder method. *Bulletin of the Swiss Association of Actuaries* **2007**, 117-137.

[9] Merz, M.; Wüthrich, M.V. Modelling the claims development result for solvency purposes. *CAS E-Forum* **Fall 2008**, 542-568.

[10] Merz, M.; Wüthrich, M.V. Claims run-off uncertainty: the full picture. *SSRN manuscript* **2014**, ID 2524352.

[11] Merz, M.; Wüthrich, M.V.; Lysenko, N. Uncertainty in the claims development result in the chain ladder method. *Scandinavian Actuarial Journal* **2009**, *109*, 63-84.

[12] Renshaw, A.E.; Verrall, R.J. A stochastic model underlying the chain-ladder technique. *British Actuarial Journal* **1998**, *4*, 903-923.

[13] Röhr, A. Chain-ladder and error propagation. Preprint **2014**.

[14] Wüthrich, M.V.; Merz, M. Stochastic claims reserving manual: advances in dynamic modeling. *SSRN manuscript* **2015**, ID 2649057.

A Data

acc.	development year j									
year i	0	1	2	3	4	5	6	7	8	9
1	5'946'975	9'668'212	10'563'929	10'771'690	10'978'394	11'040'518	11'106'331	11'121'181	11'132'310	11'148'124
2	6'346'756	9'593'162	10'316'383	10'468'180	10'536'004	10'572'608	10'625'360	10'636'546	10'648'192	
3	6'269'090	9'245'313	10'092'366	10'355'134	10'507'837	10'573'282	10'626'827	10'635'751		
4	5'863'015	8'546'239	9'268'771	9'459'424	9'592'399	9'680'740	9'724'068			
5	5'778'885	8'524'114	9'178'009	9'451'404	9'681'692	9'786'916				
6	6'184'793	9'013'132	9'585'897	9'830'796	9'935'753					
7	5'600'184	8'493'391	9'056'505	9'282'022						
8	5'288'066	7'728'169	8'256'211							
9	5'290'793	7'648'729								
10	5'675'568									

Table 4: Observed cumulative claims $C_{i,j}$ for $i + j \leq 10$ and $I = J + 1 = 10$.