Multilevel Non-Linear Random Effects Claims Reserving Models And Data Variability Structures

Graciela Vera

Abstract

Characteristic of many reserving methods designed to analyse claims data aggregated by contract or sets of contracts, is the assumption that features typifying historical data are representative of the underwritten risk and of future losses likely to affect the contracts. Kremer (1982), Bornheutter and Ferguson (1972), de Alba (2002), and many others, consider models with development patterns common to all underwriting years and known mean-variance relationships. Data amenable to such assumptions are indeed rare. More usual are large variations in settlement speeds, exposure and claim volumes. Also typifying many published models are Incurred But Not Reported (*IBNR*) predictions limited to periods with known claims, frequently adjusted with "tail factors" generated from market statistics. Of concern could be analytical approach inconsistencies behind reserves for delay periods before and after the last known claims, under reserving and unfair reserve allocation at underwriting year, array or contract levels.

As applications of Markov Chain Monte Carlo (MCMC) methods, the models proposed in this paper depart from the neat assumptions of quasi-likelihood and extended quasi-likelihood, and introduce random effects models. The primary focus is the close dependency of the *IBNR* on data variability structures and variance models, built with reference to the generic model derived in Vera (2003). The models have been implemented in BUGS (http://www.mrc-bsu.cam.ac.uk/bugs)

Keywords: Markov Chain Monte Carlo, Non-linear Random Effects and GLM, Reserving.

1. INTRODUCTION

Insurance data reflect and react to financial uncertainty associated with external events, quantifiable time varying factors such as inflation and interest rate fluctuations, and non-quantifiable factors such as variations in litigation practices and underwriting policy terms. In an interesting historical account of legislative changes introduced in Israel to deal with inflation, Kahane (1987) illustrates how external events can be given functional interpretation in a reserving model. Further examples can be found in Taylor

(2000). Data distortions due to external events could undermine all stochastic assumptions. Concerned with the analysis of claims data, from the simplest aggregation levels, such as class of business, to multiple-nested groups, this paper deals with the construction of claims reserving models capable of capturing variability structures in a claims portfolio.

Hierarchical or multi-level claims reserving models are potential source of wide-ranging contribution to claims portfolio analysis beyond reserving. Identification of the causes of data variability with reference to hierarchical model structures could provide a statistical framework for parametric analyses of claims across a number of underwriting years. This would enhance our ability to construct more discriminating models, set initial parameter values, review and update our assumptions on risk premium calculations, related management strategies for commutations, portfolio composition, analysis, etc.

1.1 Research Context

As one of the simplest claims reserving methods, the chain ladder has motivated an extensive body of work intended to establish statistical basis for the problem of reserving. Models that fall within the category of generalized linear models (GLM) (McCullagh and Nelder (1989)), such as Renshaw (1989), Renshaw and Verrall (1998), Verrall (1991), Wright (1990), Mack (1991) and many others, have extended the research beyond assumptions of lognormality and explored applications from exponential family distributions. Carroll (2003) remarks "there are many instances where understanding the structure of variability is just as central as understanding the mean structure". The *IBNR* definition given in this paper is integral to the definition of the model itself, and its value is highly sensitive to model specification. Hence, the emphasis of this research is in the identification of suitable representations for the mean and data variability structures beyond assumptions of known and specific mean-variance relationships.

Reserving model structures depend on the intended use of the predicted reserves and on the sector of interest in the claims portfolio, such as insurance class, contract, specific loss, etc. The data assessment should determine the selection of the analytical approach.

For instance, an insurance contract provides cover against the hazards listed in the contract. Premium calculations reflect policy management expenses, expected returns and risk premiums for all the perils covered by the contract. Risk premium analyses, in general, are carried out by peril, ignoring the fact that a particular event could simultaneously hit more than one kind of cover. When reserve analysis of all perils with a single model is viable, it could deliver, for example, relative cost measures capable of generating more competitive commercial premiums, hence allowing cover assessment on statistical basis, identification of cross-subsidies and unexplored niches, etc.

Within the context of hierarchical models, claims data can be differently interpreted depending on their levels of aggregation. For instance:

- Each underwriting year data set could be described as a set or cohort of longitudinal data.
- A claims array could be considered single-level longitudinal data for more than one subject.
- A book of business segmented by class, type of loss and underwriting year, could be treated as multilevel longitudinal data or as multiple nested groups of single level longitudinal data.

Davinian and Giltinan (1993 and 1996) provide an introduction to the theory of non-linear random effects models and an overview of various techniques for the analysis of non-linear models with repeated observations. More recently, Pinheiro and Bates (2000) reviews the theory and applications of linear and non-linear mixed effect models to the analysis of grouped data.

In this paper it is shown that the generic model in Vera (2003), briefly outlined below, is key to the extension of random effect models to the analysis of reserves. If the claims process for underwriting year w is reported at times $t_1, t_2, \ldots, t_e$, such that $0 < t_1 < t_2 < \ldots < t_e$, and t_e is the final settlement period, the generic model is given in terms of a percentage cash flow and a ultimate claim amount functions, denoted respectively by P_{w,t_j} and C_w. $P_{w,t_j} = \int_0^{t_j} \pi(w,z)dz$, where $\pi(w,t)$ is a probability density function taking values from positive real numbers. $S_{w,t_j} = 1 - P_{w,t_j} = \int_{t_j}^{t_e} \pi(w,z)dz$, $P_{w,t_j} \le 1$ for $j < e$ and $P_{w,t_j} = 1$ otherwise. Finally, h_{w,t_j} and H_{w,t_j} are the instant and cumulative hazard rate functions, defined for underwriting year w and payment year τ (τ = w + delay time - 1) by

$$h_{w,\tau-w+1} = -\left(\frac{\partial\left(\ln\left(1-P_{w,z}\right)\right)}{\partial z}\right)_{z=\tau-w+1} = \left(\frac{C_w}{IBNR_{(w,\tau-w+1)}}\right)\left(\frac{\partial P_{w,z}}{\partial z}\right)_{z=\tau-w+1} \tag{1.1}$$

$$H_{w,\tau-w+1} = -\ln\left(1-P_{w,\tau-w+1}\right)$$

Hence, the following are alternative representations of the claims process for cumulative data $Y_{w,\tau-w+1}$:

$$Y_{w,\tau-w+1} = C_w P_{w,\tau-w+1} \tag{1.2}$$

$$Y_{w,\tau-w+1} = C_w\left(1-\exp\left(-H_{w,\tau-w+1}\right)\right) \tag{1.3}$$

$$Y_{w,\tau-w+1} = C_w\left(1-S_{w,\tau-w+1}\right) \tag{1.4}$$

Equivalently, for incremental data $y_{w,\tau-w+1}$

$$y_{w,\tau-w+1} = C_w * \left(P_{w,\tau-w+1} - P_{w,\tau-w+1-1}\right) \tag{1.5}$$

$$y_{w,\tau-w+1} = C_w\left(\exp\left(-H_{w,\tau-w}\right) - \exp\left(-H_{w,\tau-w+1}\right)\right) \tag{1.6}$$

$$y_{w,\tau-w+1} = C_w * \left(S_{w,\tau-w} - S_{w,\tau-w+1}\right) \tag{1.7}$$

The underwriting year and array *IBNR* and reported *IBNR* projections are respectively

$$\begin{gathered} IBNR_{(w,\tau-w+1)} = C_w S_{w,\tau-w+1} \\ IBNR(\tau) = \sum_{w=1}^{u} IBNR_{(w,\tau-w+1)} \end{gathered} \tag{1.8}$$

$$\begin{gathered} RIBNR_{(w,\tau-w+1)} = IBNR_{(w,\tau-w+1)} + \left(C_w S_{w,\tau-w+1} - Y_{w,\tau-w+1}\right) \\ RIBNR(\tau) = \sum_{w=1}^{u} RIBNR_{(w,\tau-w+1)} \end{gathered} \tag{1.9}$$

where u is the number of underwriting years in the array. *RIBNR* links the reserving analysis to the accounting processes, by adjusting the *IBNR* by the difference between the total claim amount incurred to date and its estimate. Due to the additional noise induced by the adjustment, (1.9) is only applied in the final stages of the reserving analysis. In contrast to many published reserving methods, an important aspect of the models is the unrestricted *IBNR* projection periods, since the period before the last claim is generally unknown. The above equations could make explicit, and potentially highlight, the sources of data variability. Settlement speeds differences between underwriting years should be captured by $P_{w,r-w+1}$, $H_{w,r-w+1}$ or $S_{w,r-w+1}$. Although exposure levels are largely determined by underwriting volumes and contract terms, neither necessarily random, to accelerate convergence and formulate the final model variance function, random effects are introduced in C_w. When more than one claims array are analyzed, the additional aggregation level and source of variability is *array*, indexed by subscript *r*.

1.2 Objective

The examples' aim is to show that more than one model could fit historical data, but not all may reliably predict the reserves. The reliability of the *IBNR* and ultimate claim amount predictions depends on the models' capacity to extract from the data claims volume and settlement speeds measures. This is possible when the variability of both can be represented parametrically and formulated into the variance model. The scope of the models is made evident by their formulation and by the data. As the variability in settlement speeds and claims volumes increase the underlying assumptions of GLM are no longer sustainable, and more complex variance models and random effect parameters for the mean response become essential. To illustrate the process of constructing variance models two data sets are selected. One is a claims array simulated from a mixed portfolio, and the second consists of three arrays simulated from a marine hull, marine cargo and aviation hull portfolios. The second, selected to exacerbate the variability encountered in the first, in addition to large claims volume differences between underwriting years, contains also 20 negative incremental claims entries.

Since the concepts of population models (Zeger, Liang and Albert (1988)) are intended to average random variability between subjects, they are implemented around the percentage cash flow function. They can be used to obtain average (or array) *IBNR* predictions for a given ultimate loss. Other array or average results are the weighted average array or portfolio hazard rates. They provide thresholds, useful to quantify the impact on the claims portfolio of excluding from it underwriting contracts associated with particular underwriting years or arrays.

1.3 Outline

The paper structure is as follows. Section 2 introduces random effect models for one array with a general formulation of non-linear random effects models, and translated into a Bayesian framework in section 2.1.1. Noted in section 2.2 are amendments necessary to formulate multi-array models.

The models selected to analyze the two data sets are presented in sections 3 and 4 respectively. Denoted 1.0 and 2.0, in section 3.1 two preliminary models for one array are given, followed by numerical examples in section 3.3. The examples identify 2.0 as the basis for further analysis to construct the final models. In section 3.4.5 the results from two validation and two final models are discussed. Also in two stages, in section 4 multi-array models are constructed for two mean response functions denoted respectively 7.0 and 8.0. The preliminary models, used to establish data variability structures, are introduced in section 4.1, followed by numerical examples in section 4.2. For mean response functions 7.0 and 8.0, results for precision parameters σ^2, σ_r^2 and σ_{rw}^2 are obtained, identifying the three model versions by (a), (b) and (c). The final models, defined in section 4.3, are analyzed in section 4.5. They emphasise the contribution the generic model makes to the analysis of reserves, and to random effects models and variance models in general.

Section 4.4 extends the claims array average percentage cash flow definition given in section 3.2 to introduce portfolio model average for the percentage cash flow. As immediate by-products of the reserving analysis, hazard rates are discussed in section 4.6. The claims' hazard rate profile, essential for further portfolio analyses, can be used also as a portfolio management template. Discussion on the contribution made by the models proposed is given in section 5.

For the models in section 3, the results are fully reported in appendix A. Given the size of the data used in section 4, the reported results in this section are restricted to *IBNR* and ultimate claim amount projections for the selected preliminary and final models.

2. GENERAL FORMULATION OF NON-LINEAR RANDOM EFFECTS MODELS

In non-linear hierarchical models, inter and intra-underwriting year variations are analysed as a *two-stage process*. In the first, the intra-underwriting year variation is defined by a non-linear regression model for the underwriting year covariance structure. In the second stage, the inter-underwriting year variation is represented by both, systematic and random variability. The models can be constructed within a Bayesian hierarchical structure by noting that the intra-underwriting variation is associated with the sampling distribution, while the prior distribution is relevant to the inter-underwriting variation. Because the models' notation will depend on the number of aggregation levels, in sections 2.1 and 2.2 the array and multi-array analytical frameworks are respectively given.

2.1 Analytical Framework For a Claims Array

For the purpose of defining the general model, ignoring whether claims are cumulative or incremental, the observation at development time t of response vector for underwriting year w is simply denoted by $y_{w,t}$, and the model is defined as follows:

$$y_{w,t} = \mu_{w,t}(\phi_w) + \varepsilon_{w,t} \tag{2.1}$$

where $\mu_{w,t}$ is a non-linear function common to the entire array, while parameter vector ϕ_w is specific to underwriting year w. $t = t_1, \ldots, t_{n_w}$; with t_{n_w} representing the last period with known claims to date, $w = 1, \ldots, u$, and u is the number of cohorts or underwriting years in the claims array. Hence

$$y_w = \left[y_{w,t_1}, \ldots, y_{w,t_{n_w}} \right]^T$$

$$\mu_w = \left[\mu_{w,t_1}, \ldots, \mu_{w,t_{n_w}} \right]^T$$

$$\varepsilon_w = \left[\varepsilon_{w,t_1}, \ldots, \varepsilon_{w,t_{n_w}} \right]^T$$

and

$$\operatorname{cov}(\varepsilon_w) = \sigma^2 R_w \tag{2.2}$$

R_w is the intra-underwriting year covariance matrix for underwriting year w.

Inter-underwriting year variation accounted by ϕ_w is assumed to be random and, rather than simply regarding $\phi_i \neq \phi_w$ for $i \neq w$, the model represents

$$\phi_w = A_w \beta + B_w b_w$$

where β is a p-dimensional fixed parameter effects vector, and b_w is a q-dimensional underwriting year or random effects vector. Parameters b_w are independent and identically distributed with zero mean and variance covariance matrix Σ. Finally, A_w and B_w are $(n_w \times p)$ and $(n_w \times q)$ design matrices for the fixed and random effects respectively. While missing data from the earliest payment years and irregular reporting time intervals are allowed by the model formulation, the code and model specification for data given at regular intervals are simpler. The length of the response vector for the array is $M = \sum_{w=1}^{u} n_w$ and

$$\begin{array}{lll}
y = [y_1, \dots, y_u]^T & \phi = [\phi_1, \dots, \phi_u]^T & \Sigma = diag[\Sigma_1, \dots, \Sigma_u] \\
\mu = [\mu_1, \dots, \mu_u]^T & b = [b_1^T, \dots, b_u^T]^T & R = diag[R_1, \dots, R_u] \\
\varepsilon = [\varepsilon_1, \dots, \varepsilon_u]^T & & B = diag[B_1, \dots, B_u] \\
 & & A = [A_1^T, \dots, A_u^T]^T
\end{array}$$

Hence, the overall model becomes

$$\begin{aligned}
E(y) &= \mu(\phi) \\
\mathrm{var}(y) &= \sigma^2 R \\
\phi &= A\beta + Bb \\
b &\sim (0, \Sigma)
\end{aligned} \tag{2.3}$$

Corresponding to the two stages in the hierarchical models are two possible types of inferences or derived results: array and underwriting year cohort. Parameters common to all underwriting years relate to the array inferences, while underwriting year parameters measure underwriting year deviations from the claims array mean. Array inferences are generic when they represent insurance classes, and can help reassess or draft underwriting contracts, for instance. Alternatively, underwriting year parametric structures can set foundations for more discriminating premium rates reflecting systematic trends evident in the losses experienced. The latter can be viewed as a continuous calibration process.

Unless a book of business is closed, the number of observations in the most recent underwriting years could restrict the choice of viable variance and covariance models, particularly with non-linear model structures. Inferences on parameters of non-linear mixed effects models implemented in S-Plus (Pinheiro, Bates and Lindstrom (1994)) are based on the linear mixed effect model approximation of the log-likelihood function. This relies on the restricted maximum likelihood estimates derived from asymptotic results and on the approximate distribution for the maximum likelihood estimates. Since the maximum likelihood estimates in the linear mixed effect models are assumed to be asymptotically normal (Pinheiro and Bates (2000), Lindstrom and Bates (1990) and others), implementations with NLME library have to be approached with care to meet the criteria of the generic reserving model. Alternative assumptions are also considered. In non-parametric models the distribution of the random effects is left unspecified, hence completely unrestricted. Escobar and West (1992) propose a non-parametric approach, where ϕ_w are taken from distribution classes provided by the Dirichlet processes. Wakefield and Walker (1994) consider a non-parametric approach when random effect parameters are suspected to be neither normal nor Student t distributed, and allow for multimodality and skewness. Beal and Sleiner (1992) use a mixture of normal distributions and Wakefield (1996) a multivariate t-distribution for the random effect parameters and lognormal distribution for the response. The heavier tails in the t-distribution accommodate outlying cohorts. To define the parameters it is necessary to establish the curve's behaviour with parameter value changes, categorising the conditions, if any, for convergence, divergence, discontinuities etc (Ratkowski (1990)). The models' capacity to predict reserves depends on the stability of the projected curves, which in turn depends on the variance model structure. The most complex are more easily implemented within a Bayesian framework, as outlined below.

2.1.1 Three-Stage Models With Heterogeneous Intra-Underwriting Year Variation: A Bayesian Approach

Gibbs sampler application to Bayesian hierarchical models removes obstacles associated with non-linear multi-parameter structures integration. First to consider the problem of fully Bayesian non-linear regression is Wakefield et al. (1994). Bayesian random effects models can be represented by the following three-stage structure:

First Stage: Intra-underwriting year variation:

It accounts for variability within underwriting years, through scale parameter σ^2 and, in some cases, through functions $V_w(\phi_w, \Im(t), \vartheta)$ or $\Gamma_w(\rho)$, or both, such that ϕ_w, ϑ and ρ are parameters, $\Im(t)$ is some function of t and $\Gamma_w(\rho)$ is a correlation matrix. Hence, given

$$y_w = \mu_w(\phi_w) + \varepsilon_w$$

for the most general case

$$R_w(\phi_w, \vartheta^T, \rho) = V_w^{1/2}(\phi_w, \Im(t), \vartheta)\Gamma_w(\rho)V_w^{1/2}(\phi_w, \Im(t), \vartheta) \tag{2.4}$$

So ε_w are independently and identically distributed with zero mean and

$$Cov(\varepsilon_w \mid \phi_w, \sigma, \vartheta^T, \rho) = \sigma^2 R_w(\phi_w, \vartheta^T, \rho) \tag{2.5}$$

The functional form of $R_w(\phi_w, \vartheta^T, \rho)$ and covariance parameters $\zeta = \left[\sigma, \vartheta^T, \rho\right]^T$ are the same for all underwriting years. Implicit in $V_w(\phi_w, \Im(t), \vartheta)$ are functions of $\mu_w(\phi_w)$ or t, and of some or all parameters in ϕ_w. If probability distribution function is denoted by f then

$$(y_w \mid \phi_w, \zeta) \sim f_{y_w \mid \phi_w, \zeta}(y_w \mid \phi_w, \zeta)$$

Second Stage: Inter- underwriting year variation:

The inter-underwriting year variation in the values of ϕ_w is represented by

$$\phi_w = A_w \beta + B_w b_w \tag{2.6}$$

The degree of complexity of design matrices A_w and B_w will depend on the data and the percentage cash flow function. Random effect parameters are assumed to be independent and identically distributed:

$$b_w \sim f_{b_w \mid \Sigma}(b_w \mid \Sigma) \tag{2.7}$$

Zero mean assumption for b_w is not essential and, with software packages such as BUGS, may not be attainable. Non parametric and semiparametric model specifications for $\phi_{w\cdot}$ can be considered.

Third Stage: Hyperprior distribution:

Definition of parameters β,ζ and Σ completes the model formulation.

$$(\beta,\zeta,\Sigma)\sim f_{\beta,\zeta,\Sigma}(\beta,\zeta,\Sigma) \tag{2.8}$$

The joint posterior distribution of all parameters upon which the Bayesian inferences are based is

$$f_{\beta,\zeta,b,\Sigma|y}(\beta,\zeta,b,\Sigma\,|\,y)=\frac{f_{y|\beta,\zeta,b}(y\,|\,\beta,\zeta,b)f_{b|\Sigma}(b\,|\,\Sigma)f_{\beta,\zeta,\Sigma}(\beta,\zeta,\Sigma)}{f_y(y)} \tag{2.9}$$

The marginal posterior distributions of interest are $f_{\beta|y}(\beta\,|\,y)$, $f_{b|y}(b\,|\,y)$ and $f_{\Sigma|y}(\Sigma\,|\,y)$.

Implicit in the above are two simpler models:

- For uncorrelated intra-underwriting year observations $\Gamma_{w\cdot}(\rho)=I_{n_w\times n_w}$ and $\zeta=\left[\sigma,\vartheta^T\right]^T$.
- If the model is homoscedastic, then $\Gamma_{w\cdot}(\rho)=R_w(\phi_{w\cdot},\vartheta^T,\rho)=I_{n_w\times n_w}$ and $\zeta=\sigma$.

As a simple example, consider

$$b_w\big|\Sigma\sim N(0,\Sigma) \tag{2.10}$$

and

$$\begin{gathered}\beta\big|\beta^*,\Sigma_0\sim N\left(\beta^*,\Sigma_0\right)\\ \Sigma^{-1}\big|\Sigma^*,\nu^*\sim Wi\left(\left(\nu^*\Sigma^*\right)^{-1},\nu^*\right)\\ \frac{1}{\sigma^2}\bigg|\upsilon,\nu\sim Ga\left(\frac{\upsilon}{2},\frac{\upsilon\nu}{2}\right)\end{gathered} \tag{2.11}$$

where parameters $\beta^*,\Sigma_0,\Sigma^*,\nu^*,\upsilon,\nu$ are known. When a linearization method is used Σ_0 could be replaced by $\sigma^2\left(\hat{X}^T\hat{X}\right)^{-1}$, such that $\hat{X}=\left.\frac{\partial\mu}{\partial\beta}\right|_{\hat{\beta},\hat{b}}$. If $\bar{\beta}=\frac{1}{u}\sum_{w=1}^{u}\beta_{w\cdot}$, the parameters' conditional distributions for correlated observations are:

$$(\beta \mid y,\sigma,\Sigma,\phi_w, w=1,\ldots,u) \sim N\left(\begin{array}{l}\left(u\Sigma^{-1}+\left(\Sigma_0^{-1}+\Sigma^{-1}\right)^{-1}\right)^{-1}\left(u\Sigma^{-1}\bar{\beta}+\left(\Sigma_0^{-1}+\Sigma^{-1}\right)^{-1}\beta^{*}\right),\\ \left(u\Sigma^{-1}+\left(\Sigma_0^{-1}+\Sigma^{-1}\right)^{-1}\right)^{-1}\end{array}\right)$$

$$\left(\Sigma^{-1} \mid y,\sigma,\beta,\phi_w, w=1,\ldots,u\right) \sim Wi\left(\left[\sum_{w=1}^{u}(\phi_w-\beta)(\phi_w-\beta)^T+v^{*}\Sigma^{*}\right]^{-1}, u+v^{*}\right) \tag{2.12}$$

$$\left(\sigma^{-2} \mid y,\beta,\Sigma,\phi_w, w=1,\ldots,u\right) \sim$$

$$Ga\left(\left(\frac{\upsilon+M}{2}\right),\frac{1}{2}\left(\upsilon v+\sum_{w=1}^{u}\left(y_w-\mu_w(\phi_w)\right)^T R_w^{-1}(\phi_w,\vartheta^T,\rho)\left(y_w-\mu_w(\phi_w)\right)\right)\right)$$

Then, the conditional distributions of ϕ_w and ϑ are

$$f_{\phi_w \mid y,\beta,\Sigma,\sigma,\phi_i,w\neq i}(\phi_w \mid y,\beta,\Sigma,\sigma,\phi_i, w\neq i) \propto$$

$$\exp\left(-\frac{1}{2\sigma^2}\left(y_w-\mu_w(\phi_w)\right)^T R_w^{-1}(\phi_w,\vartheta^T,\rho)\left(y_w-\mu_w(\phi_w)\right)-\frac{1}{2}(\phi_w-\beta)\Sigma^{-1}(\phi_w-\beta)^T\right)\left|R_w(\phi_w,\vartheta^T,\rho)\right|^{-1/2}\sigma$$

$$f_{\vartheta \mid y,\beta,\Sigma,\sigma,\phi_w,w=1,\ldots,u}(\vartheta \mid y,\beta,\Sigma,\sigma,\phi_w, w=1,\ldots,u) \propto \prod_{w=1}^{u}\left(\frac{\exp\left(-0.5\sigma^{-2}\left(y_w-f_{wt}(\phi_w)\right)^T R_w^{-1}(\phi_w,\vartheta^T,\rho)\left(y_w-f_{wt}(\phi_w)\right)\right)}{\sigma^{-1}\left|R_w(\phi_w,\vartheta^T,\rho)\right|^{1/2}}\right)$$

Variations on the above general model, with $\Gamma_w(\rho)=I_{n_w \times n_w}$ can be found in Wakefield (1996). In relation to the purpose of this paper, in the first stage, where variability structures are established, the predicted values that contribute to the $IBNR_{(w,t_j)}$ (equation (1.8)) are simply defined as $C_w \int_{t_j}^{t_r} \pi(w,z)dz$. While for the final models, y_w^{*}, or predicted losses for underwriting year *w*, are sampled from the distribution $f_{y_w^{*}\mid y,\beta,\Sigma,\sigma,\phi_w}\left(y_w^{*} \mid y,\beta,\Sigma,\sigma,\phi_w\right)$ and applied to equation (1.9).

2.2 General Formulation Of Multi-Array Bayesian Models

Extending the general model for a single array, the three-stage multi-array hierarchical model requires the following notation. For underwriting year *w* in claims array *r*, where $r=1,\ldots,r_1$ and $w=1,\ldots,u_r$, let the response vector be

$$y_{r,w}=\left[y_{r,w,t_1},\ldots,y_{r,w,t_{n_{rw}}}\right]^T$$

Hence, for the entire data set

$$y = \left[y_{1,1}, \ldots, y_{1,u_1}, \ldots, y_{r_1,1} \ldots, y_{r_1,u_{r_1}} \right]^T$$

$t = t_1, \ldots, t_{n_{rw}}$ are the reporting times, such that $t_{n_{rw}}$ denotes the last period with known claims for underwriting year w. $y_{r,w}$ will be replaced by $Y_{r,w}$ when the data analysed is cumulative. The length of the response vector is M, such that $n_r = \sum_{w=1}^{u_r} n_{rw}$ and $M = \sum_{r=1}^{r_1} n_r$. We write

$$y_{r,w} = \mu_{r,w}(\phi_{rw}) + \varepsilon_{r,w}$$

where

$$\phi_{rw} = A_{rw}\beta + B_{r,w}b_r + B_{rw}b_{rw} \tag{2.13}$$

β is a p-dimensional fixed effects parameter vector, b_r is a q_1-dimensional first level random effects vector and b_{rw} a q_2-dimensional second level random effects vector. b_r and b_{rw} could be defined to have zero mean and variance/covariance matrices Σ_1 and Σ_2 respectively. Through design matrices A_{rw}, $B_{r,w}$ and B_{rw} information specific to each underwriting year data set can be brought into the analysis. By replacing (2.6) by (2.13) the three-stage models accounts also for array variation.

The models in section 3, and those in section 4, show that more flexible covariance structures could provide insight into the data variability structures by exploring alternative definitions for $\zeta = \left[\sigma^2, \vartheta^T, \rho\right]^T$. However, to avoid degrading inferences on first moment components, the final model should assume common parameters ζ for all underwriting years and arrays. Hence, the problem consists of identifying any relationship evident between $\Im(t)$, ϕ_{rw}, $\mu_{r,w}$ or any other function of ϕ_{rw}, and the patterns of variability revealed by parameters $\zeta = \left[\sigma^2, \vartheta^T, \rho\right]^T$. Outliers could lead to incorrect inferences, possibly indicate that the claims distribution is in fact multimodal and the data should be segmented for analytical purposes. Although the models proposed do not include specific functions to capture payment year effects of the kind of systematic inflation, they can be easily amended to do so.

3. MODELS FOR ONE ARRAY

3.1 Examples Of Preliminary Models For One Array

Two preliminary models, denoted 1.0 and 2.0 respectively, are given below. Both have a power variance function. However, to assess variability assumptions and construct the final models, the power in model 2.0 is allowed to change with underwriting year. With the variance formulation of model 2.0 the standard variance parameter definition is disregarded, by using instead $\zeta_{w} = \left[\sigma, \vartheta_{w}^{T}, \rho\right]^{T}$, thereby weakening the inferential capability of the model. Hence, even if the *IBNR* and ultimate claim amount predictions for model 2.0 were satisfactory, model 2.0 should be treated as preliminary and used exclusively for exploratory purposes.

3.1.1 Model 1.0

The first heteroscedastic model is defined as follows:

$$Y_{w,t} = \mu_{w,t}(\phi_{w}) + \varepsilon_{w,t}$$

with

$$\left(Y_{w,t} \middle| \phi_{w}, \zeta\right) \sim N\left(\mu_{w,t}(\phi_{w}), \sigma^{2} \mu_{w,t}(\phi_{w})^{\exp(\vartheta)}\right)$$

such that, $\zeta = [\sigma, \vartheta]^{T}$

$$\mu_{w,t}(\phi_{w}) = \frac{\exp(L + l_{w})}{\left\{1 + \exp\left(D + d_{w} - \exp(Kc + kc_{w}) \ln\left(t^{*}\right) - \exp(Kd + kd_{w}) * t^{*}\right)\right\}} \quad (3.1)$$

and

$$\varepsilon_{w} \mid \phi_{w}, \zeta \sim N\left(0, \sigma^{2}\left(\mu_{w}(\phi_{w})^{\exp(\vartheta)}\right)^{T} I_{n_{w} \times n_{w}}\right)$$

where $t^{*} = \left(t + \frac{\exp(Ks_{1})}{t^{\exp(Ks_{2})}}\right)$ and

$$\phi_{w} = A_{w}\beta + B_{w} b_{w}$$

$$A_{w} = I_{7\times 7}$$

$$(B_{w})^{T} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$b_{w} = [l_{w}, d_{w}, kc_{w}, kd_{w}]^{T}, \qquad b_{w} \middle| b_{w}^{*}, \Sigma_{w} \sim MVN\left(b_{w}^{*}, \Sigma_{w}\right)$$

$$\beta = [Ks_{1}, Ks_{2}, \vartheta, L, D, Kc, Kd]^{T}, \quad \beta \middle| \beta^{*}, \Sigma_{0} \sim MVN\left(\beta^{*}, \Sigma_{0}\right)$$

The hyperprior distributions are $\frac{1}{\sigma^2} \sim Ga(0.001, 0.001)$, and

$$\beta^{\bullet}|\beta^{\bullet\bullet},\Sigma_0^{\bullet\bullet} \sim MVN(\beta^{\bullet\bullet},\Sigma_0^{\bullet\bullet}) \qquad (\Sigma_0)^{-1}|\Sigma_0^{\bullet} \sim Wi\left((7\Sigma_0^{\bullet})^{-1},7\right)$$
$$b_{w}^{\bullet}|b^{\bullet\bullet},\Sigma_{w}^{\bullet\bullet} \sim MVN(b^{\bullet\bullet},\Sigma_{w}^{\bullet\bullet}) \qquad (\Sigma_{w})^{-1}|\Sigma_{w}^{\bullet} \sim Wi\left((4\Sigma_{w}^{\bullet})^{-1},4\right)$$

for given parameters $\beta^{\bullet\bullet},b^{\bullet\bullet},\Sigma_0^{\bullet},\Sigma_0^{\bullet\bullet},\Sigma_w^{\bullet},\Sigma_w^{\bullet\bullet}$. Functions C_{w}, $P_{w,t}$ and $h_{w,t}$ describing the underlying claims process for model 2.0 are:

$$C_{w} = \exp(L + l_{w})$$

$$P_{w,t} = \left\{1+\exp\left(D+d_{w} - \exp(Kc+kc_{w})\ln(t^{\bullet}) - \exp(Kd+kd_{w})*t^{\bullet}\right)\right\}^{-1} \qquad (3.2)$$

$$S_{w,t} = 1 - P_{w,t} \qquad (3.3)$$

$$h_{w,t} = \left(\frac{\exp(Kc+kc_{w})}{t^{\bullet}} + \exp(Kd+kd_{w})\right)\left(1 - \frac{\exp(Ks_2+Ks_1)}{t^{\exp(Ks_2)+1}}\right)P_{w,t} \qquad (3.4)$$

3.1.2 Model 2.0

In model 2.0 $\mu_{w,t}(\phi_{w})$ is given by (3.1), but

$$\left(Y_{w,t}|\phi_{w},\zeta_{w}\right) \sim N\left(\mu_{w,t}(\phi_{w}),\sigma^2\mu_{w,t}(\phi_{w})^{\exp(\vartheta+\vartheta_{w})}\right)$$

where $\zeta_{w} = [\sigma,\vartheta,\vartheta_{w}]^T$. Having included random effect parameters in the variance function, the following further amendments to model 1.0 are needed:

$$A_{w} = B_{w} = I_{5\times5}$$
$$\varepsilon_{w} \mid \phi_{w},\zeta_{w} \sim N\left(0,\sigma^2\left(\mu_{w}(\phi_{w})^{\exp(\vartheta+\vartheta_{w})}\right)^T I_{n_w\times n_w}\right)$$
$$b_{w} = [\vartheta_{w},l_{w},d_{w},kc_{w},kd_{w}]^T, \quad b_{w}|b_{w}^{\bullet},\Sigma_{w} \sim MVN(b_{w}^{\bullet},\Sigma_{w})$$

such that

$$b_{w}^{\bullet}|b^{\bullet\bullet},\Sigma_{w}^{\bullet\bullet} \sim MVN(b^{\bullet\bullet},\Sigma_{w}^{\bullet\bullet}) \qquad (\Sigma_{w})^{-1}|\Sigma_{w}^{\bullet} \sim Wi\left((5\Sigma_{w}^{\bullet})^{-1},5\right)$$

The power parameters in models 1.0 and 2.0 are formulated as multivariate normal, together with the mean response parameters. The reason becomes evident in section 3.4, where the relationship between the parameters is analysed to construct the final models.

3.2 Claims Array Average Percentage Cash Flow Model

Non-linear mixed effect or population models (Zeger, Liang and Albert (1988)) are intended to deliver population parameter distributions to derive population inferences. The inter-subject variability allowed by the models assumes that the subject-specific parameters are identically and independently distributed. The generic claims reserving model describes the data as the product of functions for the percentage cash flow and the ultimate claim amount. In the best scenario the ultimate claim amount function would account for differences in claim and exposure volumes. Since both could be largely determined by underwriting contract terms, for array inferences to be representative of the type of peril the contract covers, they are better based on the percentage cash flow functions alone.

In the example that follows the general formulation of the random effects model, the random effects parameters are set to be $b_w|\Sigma \sim N(0,\Sigma)$, while observing that alternative definitions are feasible. In some applications or models it may not be possible to assume a zero mean for the random effects parameters, particularly when they are defined to belong to multivariate distributions. BUGS, for instance, cannot handle multivariate range restrictions, but can accommodate some simpler univariate centering forms.

Replacing design matrices A_w and B_w in models 1.0 and 2.0 by A and B respectively, the parameters for the claims array average percentage cash flow model have to be extracted from the parameter vector given by

$$\phi_A = A\beta + B\left(\frac{1}{u}\sum_{w=1}^{u} b_w\right)$$

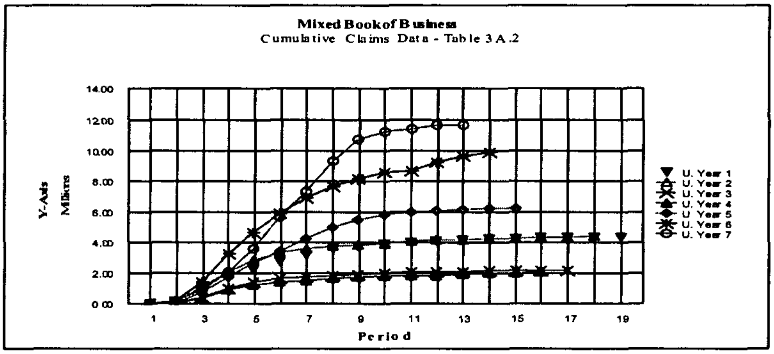
Hence, for model 2.0, the claims array average percentage cash flow curve is:

$$P_{A,t} = \left\{1+\exp\left(D+\left(\frac{1}{u}\sum_{w=1}^{u} d_w\right)-\exp\left(Kc+\left(\frac{1}{u}\sum_{w=1}^{u} kc_w\right)\right)\ln\left(t^*\right)-\exp\left(Kd+\left(\frac{1}{u}\sum_{w=1}^{u} kd_w\right)\right)*t^*\right)\right\}^{-1} \qquad (3.5)$$

with t^* defined as before. To ascertain if (3.5) is representative of the array, the curve is compared to the plots for the percentage cash flow for all underwriting years in the array.

3.3 Numerical Examples Of Preliminary Models 1.0 Aand 2.0

Extracted from a book of business containing more than one type of claim, the data selected for the examples display significant differences in the development patterns and exposure volumes across underwriting years, particularly evident in the last three underwriting years. (See graph 3.3.1 and tables A.1 and A.2). Another characteristic is the zero claims in the first reporting period. To ensure they are not interpreted as missing data, they have been set to one. This artifice is often necessary with non-linear models for the mean response or when the mean response is formulated into the variance.

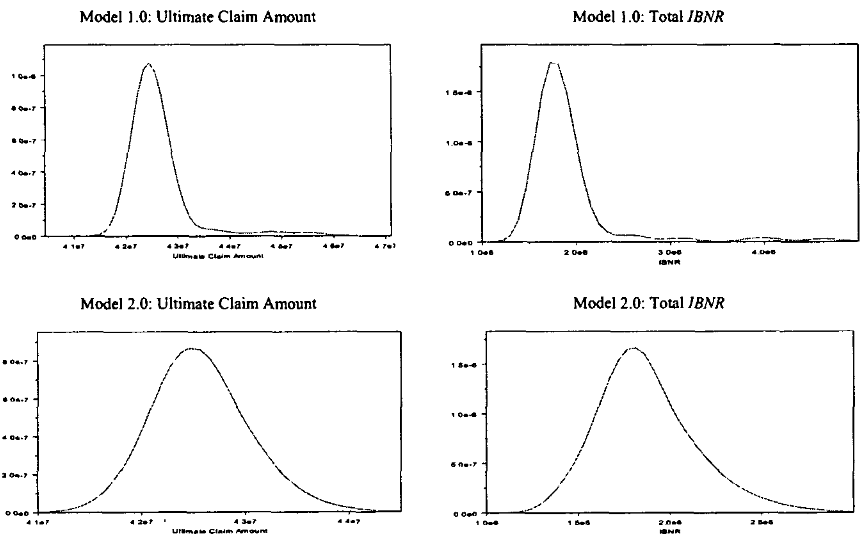


Graph 3.3.1 Cumulative paid claims data aggregated on annual basis.

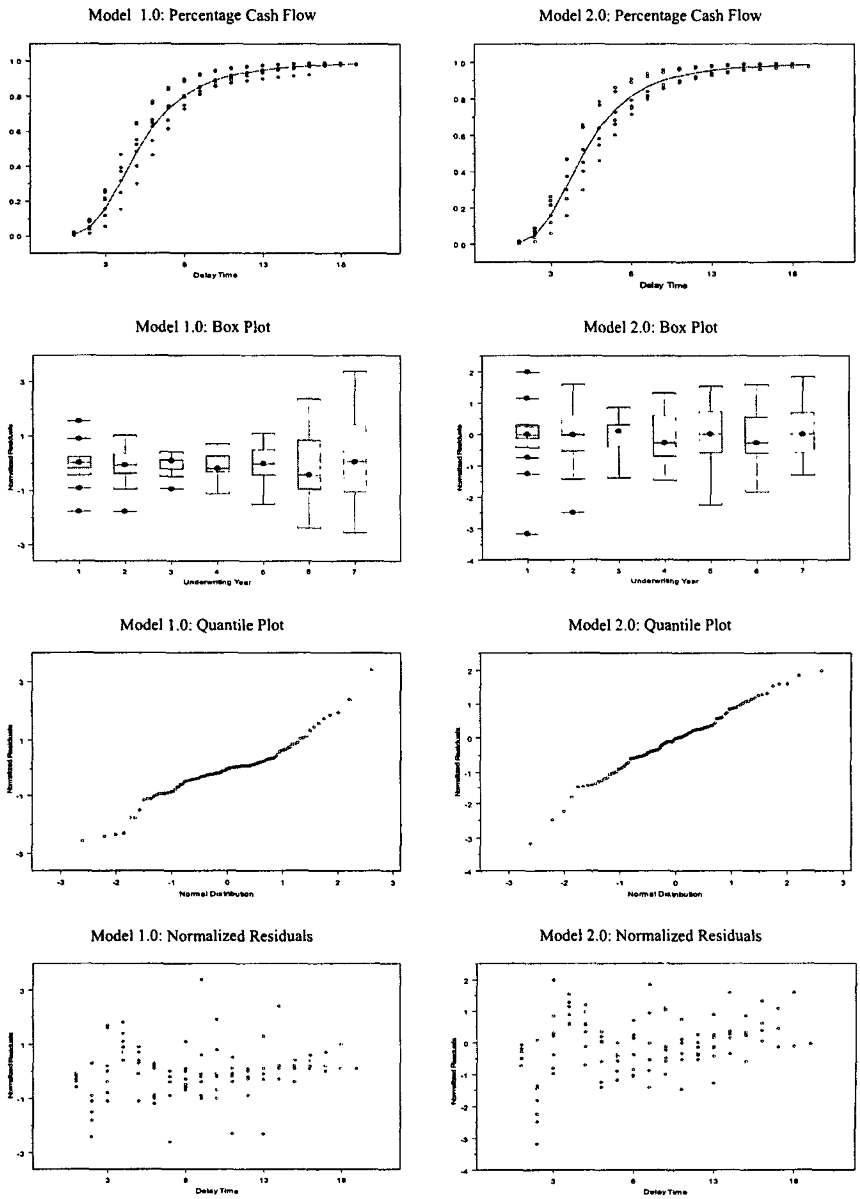
Of interest in the examples are the repercussions of hierarchical variance models. To facilitate the analysis of the preliminary models, the *IBNR* predictions do not include the accounting adjustment in (1.9) Graphs for observed claims and fitted values for the preliminary and final models would show that the fitted curves are almost indistinguishable and very close to the data. However, from table 3.3.1 and graph 3.3.3 observe that the *IBNR* predictions at underwriting year level for model 1.0 cannot be reliably used. The plot for the percentage cash flow for underwriting year 4 is unlikely to converge to 1. The model compensates by producing a higher *IBNR*. As graphical representations of spread, location and skewness for error distributions, the box plots show that, in contrast with model 1.0, with the introduction of parameter ϑ_w in the variance function, model 2.0 deals effectively with scale variability and with some of the outliers evident in the quantile plots.

	Ultimate Claim Amount				*IBNR* (1.8)			
	Mean	Standard Mean Sq. Pred. Error	Predictive Interval 2.50%	Predictive Interval 97.50%	Mean	Standard Mean Sq. Pred. Error	Predictive Interval 2.50%	Predictive Interval 97.50%
Model 1.0								
Und. Year 1	4,442,000	80,410	4,302,000	4,611,000	95,490	28,990	53,920	161,100
2	4,342,000	73,480	4,219,000	4,487,000	46,260	18,980	19,770	87,900
3	2,180,000	65,470	2,059,000	2,311,000	22,850	14,550	3,489	57,970
4	2,179,000	539,200	1,889,000	4,383,000	173,100	467,000	4,487	2,116,000
5	6,642,000	115,100	6,413,000	6,863,000	290,800	51,360	197,400	397,800
6	10,170,000	153,100	9,892,000	10,500,000	607,200	85,100	457,700	791,100
7	12,650,000	131,400	12,380,000	12,880,000	676,100	54,630	556,600	774,600
Total	42,600,000	607,100	41,940,000	44,830,000	1,912,000	499,400	1,546,000	3,915,000
Model 2.0								
Und. Year 1	4,481,000	58,480	4,373,000	4,604,000	113,500	23,960	74,340	168,000
2	4,327,000	48,600	4,233,000	4,424,000	40,920	11,260	21,990	66,410
3	2,165,000	39,650	2,093,000	2,249,000	16,400	8,756	5,335	38,440
4	2,007,000	54,090	1,909,000	2,128,000	31,570	17,560	8,824	79,130
5	6,644,000	88,190	6,474,000	6,822,000	293,500	42,590	217,800	382,600
6	10,160,000	257,300	9,668,000	10,700,000	606,400	144,000	348,000	933,100
7	12,790,000	348,100	12,170,000	13,580,000	771,200	200,400	458,100	1,259,000
Total	42,570,000	471,100	41,720,000	43,590,000	1,873,000	268,600	1,421,000	2,502,000

Table 3.3.1 Ultimate losses and *IBNR* predictive distributions for models 1.0 and 2.0



Graph 3.3.2 Kernel densities for ultimate losses and *IBNR* totals for preliminary models 1.0 and 2.0.



Graph 3.3.3 Percentage cash flow plots and normalized residuals for models 1.0 and 2.0.

Hence, model 2.0 in general, and ϑ_w in particular, should be analysed to formulate the final variance model. The Kernel densities for ultimate claim amount and *IBNR* projections in graph 3.3.2 suggest possible bi-modality, particularly for model 1.0.

Note that a variance derived directly from model 2.0 may not deal completely with the pattern evident in the plots for the normalized residuals (graph 3.3.3). Portfolio transfers or account consolidations often produce data sets where the settlement speeds of the new and old data differ significantly. The quantile and scatter plots point to the second observation in underwriting years 1, 2, 5 and 6 as possible outliers. These give an indication that the correction needed in the variance model may involve a function dependent on delay period t. In the next section the variance function for model 5.0 is derived form the output of model 2.0. With the variance function for model 6.0 it is aimed to deal with remaining outliers.

3.4. Final One-Array Models

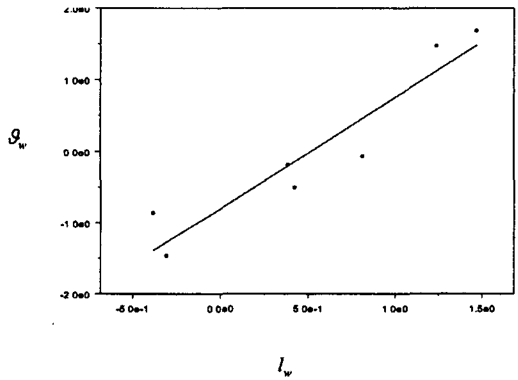
The generic model conveniently separates the percentage cash flow and the ultimate claim amount functions and, through the percentage cash flow function, can extract from the data settlement speed characteristics. Deviations induced by large differences in underwriting volumes between underwriting years may not be captured by random effect models, and the introduction of cluster structures may be necessary. The criteria needed to establish them remains to be determined.

	Model 1.0		Model 2.0	
	Fixed parameters		Fixed parameters	
	L	ϑ	L	ϑ
	15.7400	-6.2470	14.8900	-3.4710
Book Year	Random parameters		Random parameters	
	l_w	ϑ_w	l_w	ϑ_w
1	-0.4364		0.4214	-0.5061
2	-0.4592		0.3865	-0.1910
3	-1.1480		-0.3060	-1.4670
4	-1.1690		-0.3818	-0.8611
5	-0.0342		0.8153	-0.0693
6	0.3919		1.2400	1.4680
7	0.6101		1.4700	1.6780

Table 3.4.1 Parameter estimates for variance model and C_w function.

From tables 3.3.1 and 3.4.1 note the approximate correspondence between the order of magnitudes of C_w and ϑ_w for model 2.0. Hence consider the following regression line:

$$\vartheta_w = \delta_1 + \delta_2 l_w \tag{3.6}$$



Graph 3.4.1 Line $\vartheta_w = \delta_1 + \delta_2 l_w$ and scatter plot of l_w vs ϑ_w.

(3.6) gives $[\delta_1, \delta_2] = [-0.79914, 1.54866]$. From table A.6, there is no evident relationship, similar to (3.6), between ϑ_w and any of d_w, kc_w or kd_w, directly or through a suitable transformation. Graph 3.4.1 displays the scatter plot of ϑ_w versus l_w and regression line $\vartheta_w = \delta_1 + \delta_2 l_w$. $\exp(\vartheta + \vartheta_w)$ lies between 0.007 and 0.166, such that the minimum and maximum values correspond to $w = 3$ and $w = 7$ respectively. Had model 1.0 provided a better fit, the magnitudes of $\exp(\vartheta + \vartheta_w)$ could have influenced a decision to select a homoscedastic model. However, from table 3.3.1 note that $C_3 = 2.17$ million and $C_7 = 12.79$ million. Hence, it is justifiable to integrate $\exp(\vartheta + \vartheta_w) \cong \exp\left((\vartheta + \delta_1) + l_w * \delta_2\right)$ in the variance function definition as follows:

$$\mathrm{var}(Y_{w,t}) = \sigma^2 \mu_{w,t} (\phi_w)^{\exp\left(\vartheta_1^* + l_w \vartheta_2^*\right)} \tag{3.7}$$

(3.7) satisfies covariance definition (2.5) because parameter vector $\zeta = \left[\sigma, \left[\vartheta_1^*, \vartheta_2^*\right]^T\right]^T$ is invariant with underwriting year. It is intuitively obvious that if the variance parameters for preliminary model 2.0 would have been defined as $\zeta_w = \left[\sigma_w, \vartheta^T\right]^T$ instead of $\zeta_w = \left[\sigma, \vartheta_w^T\right]^T$, a similar relationship to (3.6) would be evident between σ_w and l_w. This is further explored in section 4. Although the variance for final model 5.0 will be (3.7), to validate the model and explore alternative analytical approaches, the final model is preceded by other two. The first gives an appreciation of the *IBNR* reserve values that an analysis of the data segmented into *K* subsets would deliver, where subset membership criteria is determined by the values of c_w or ϑ_w. Hence, the variance function considered is

$$\mathrm{var}(Y_{w,t}) = \sigma_k^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta)}$$

The second preliminary model assumes an autoregressive error structure. Variance function (3.7) may not successfully explain the variability evident in the normalized residual plot pattern of model 2.0 (graph 3.3.3) and function $\mu_{w,t}(\phi_w)^{\exp(\vartheta_1^* + l_w\vartheta_2^*)}$ may need to be adjusted. Hence, the function proposed for model 6.0 is

$$\mathrm{var}(Y_{w,t}) = \sigma^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta_1^* + l_w\vartheta_2^*)} \exp\left(\vartheta_3^* \left(t^*\right)^{\vartheta_4^*}\right)$$

3.4.1 Model 3.0 – Validation Model

Model 3.0 is equal to model 1.0 in all respects, except that subset membership for each underwriting year is taken into account only at the point of calculating the variance, and for subset *k* σ_k^2 is estimated independently from the rest of the data. For underwriting year *w*, member of subset *k*

$$Y_{w,t} = \mu_{w,t}(\phi_w) + \varepsilon_{w,t}$$

with

$$\left(Y_{w,t} \middle| \phi_w, \zeta_k\right) \sim N\left(\mu_{w,t}(\phi_w), \sigma_k^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta)}\right)$$

such that $\zeta_k = [\sigma_k, \vartheta]^T$, $\frac{1}{\sigma_k^2} \sim Ga(0.001, 0.001)$ and

$$\varepsilon_w \mid \phi_w, \zeta_k \sim N\left(0, \sigma_k^2 \left(\mu_w(\phi_w)^{\exp(\vartheta)}\right)^T I_{n_w \times n_w}\right)$$

3.4.2 Model 4.0 – Validation Model

In model 4.0 the option of using an autoregressive error structure is explored, to ascertain if this can effectively deal with scale variability between underwriting years:

$$Y_{w,t} = \mu_{w,t}(\phi_w) + W_{w,t}$$

such that

$$W_{w,t} = \rho W_{w,t-1} + \varepsilon_{w,t}$$

$$VAR(Y_{w,t}) = VAR(W_{w,t}) = \sigma^2$$

$$\varepsilon_w \mid \phi_w, \zeta_w \sim N\left(0, \sigma^2\left(1-\rho^2\right)\right)$$

and $\zeta = [\sigma, \rho]^T$.

3.4.3 Model 5.0 – Validation Model

Final model 5.0 integrates regression (3.6) into the variance model:

$$Y_{w,t} = \mu_{w,t}(\phi_w) + \varepsilon_{w,t}$$

with

$$\left(Y_{w,t} \middle| \phi_w, \zeta\right) \sim N\left(\mu_{w,t}(\phi_w), \sigma^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta_1^* + I_w \vartheta_2^*)}\right)$$

where $\vartheta^T = \left[\vartheta_1^*, \vartheta_2^*\right]$, $\varepsilon_w \mid \phi_w, \zeta \sim N\left(0, \sigma^2\left(\mu_w(\phi_w)^{\exp(\vartheta_1^* + I_w \vartheta_2^*)}\right)^T I_{n_w \times n_w}\right)$, $\frac{1}{\sigma^2} \sim Ga(0.001, 0.001)$ and $\zeta = \left[\sigma, \vartheta^T\right]^T$. Since $\mu_{w,t}$ is given by (3.1), in addition to the obvious changes in the design matrices, the other necessary amendments to model 2.0 are:

$$\beta = \left[Ks_1, Ks_2, \vartheta_1^*, \vartheta_2^*, L, D, Kc, Kd\right]^T, \quad \beta \middle| \beta^*, \Sigma_0 \sim MVN\left(\beta^*, \Sigma_0\right)$$

$$\beta^* \middle| \beta^{**}, \Sigma_0^{**} \sim MVN\left(\beta^{**}, \Sigma_0^{**}\right) \qquad \left(\Sigma_0\right)^{-1} \middle| \Sigma_0^* \sim Wi\left(\left(8\Sigma_0^*\right)^{-1}, 8\right)$$

3.4.4 Model 6.0 – Final Model

Final model 6.0 extends the variance model (3.7) as follows:

$$\left(Y_{w,t} \middle| \phi_w, \zeta\right) \sim N\left(\mu_{w,t}(\phi_w), \sigma^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta_1^* + I_w \vartheta_2^*)} \exp\left(\vartheta_3^*\left(t + \frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)^{\vartheta_4^*}\right)\right)$$

where $\zeta = \left[\sigma, \vartheta^T\right]^T$ and $\vartheta^T = \left[\vartheta_1^*, \vartheta_2^*, \vartheta_3^*, \vartheta_4^*\right]$. Hence the fixed effect parameter vector and related distributions are:

$$\beta = \left[Ks_1, Ks_2, \vartheta_1^*, \vartheta_2^*, \vartheta_3^*, \vartheta_4^*, L, D, Kc, Kd\right]^T, \quad \beta \middle| \beta^*, \Sigma_0 \sim MVN\left(\beta^*, \Sigma_0\right)$$

$$\beta^* \middle| \beta^{**}, \Sigma_0^{**} \sim MVN\left(\beta^{**}, \Sigma_0^{**}\right) \qquad \left(\Sigma_0\right)^{-1} \middle| \Sigma_0^* \sim Wi\left(\left(10\Sigma_0^*\right)^{-1}, 10\right)$$

3.4.5 Numerical Examples And Discussion For Validation Models 3.0 And 4.0 And Final Models 5.0 And 6.0

Data segmentation criteria described by the last column in table 3.4.5.1 and applied to model 3.0 is given by the values of l_w from model 2.0.

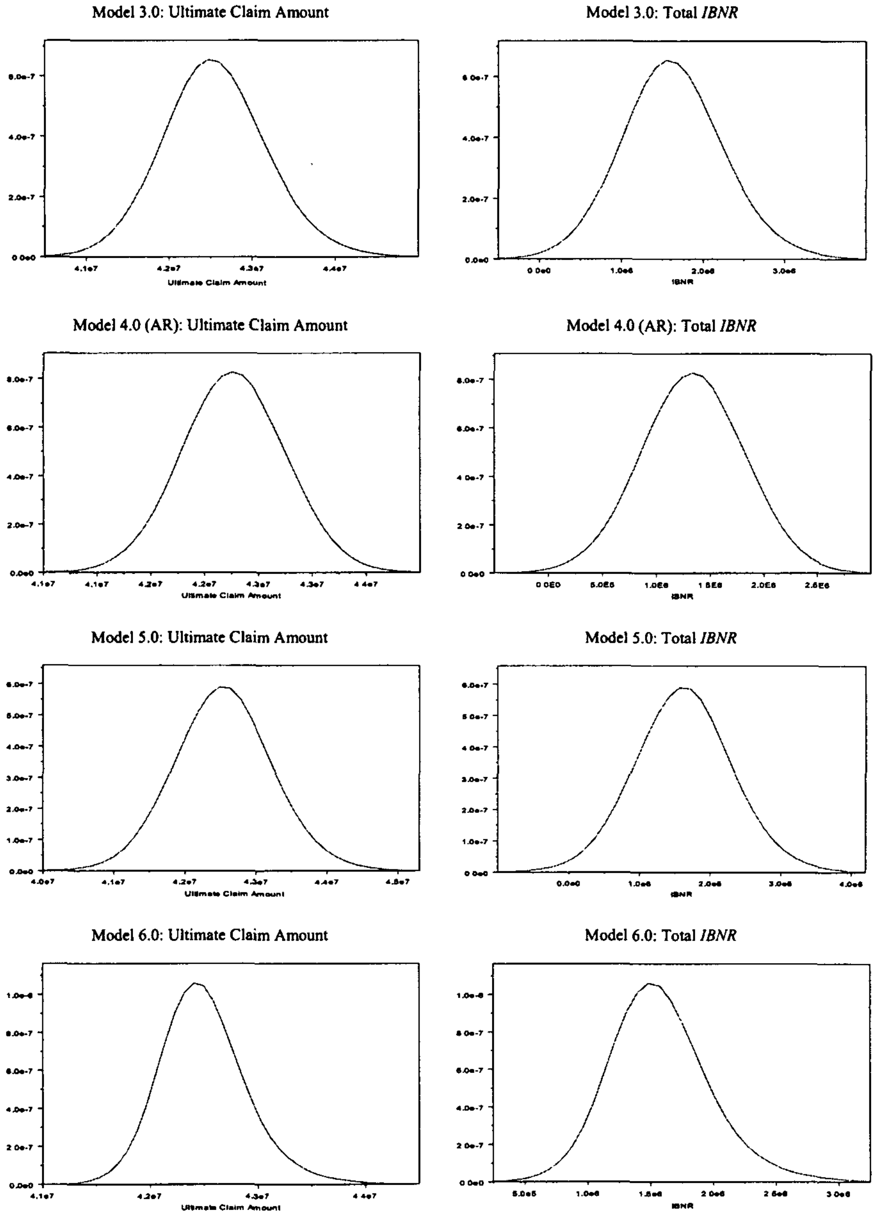
	Model 2.0		Model 3.0		Subset Membership For Model 3.0
	Fixed parameters		Fixed parameters		
	L	ϑ	L	ϑ	
	14.890	-3.471	14.290	-6.861	
Book Year	Random parameters		Random parameters		
	l_w	ϑ_w	l_w	ϑ_w	
1	0.4214	-0.5061	1.0300		1
2	0.3865	-0.1910	0.9940		1
3	-0.3060	-1.4670	0.3015		2
4	-0.3818	-0.8611	0.2259		2
5	0.8153	-0.0693	1.4240		3
6	1.2400	1.4680	1.8620		3
7	1.4700	1.6780	2.0660		3
Book Year	Combined Effect		Combined Effect		
	$L+l_w$	$\exp(\vartheta+\vartheta_w)$	$L+l_w$	$\exp(\vartheta+\vartheta_w)$	
1	15.3114	-3.9771	15.3200		1
2	15.2765	-3.6620	15.2840		1
3	14.5840	-4.9380	14.5915		2
4	14.5082	-4.3321	14.5159		2
5	15.7053	-3.5403	15.7140		3
6	16.1300	-2.0030	16.1520		3
7	16.3600	-1.7930	16.3560		3
σ^2	5.87E+09				
Subset 1 - σ_1^2			7.496E+09		
Subset 2 - σ_2^2			3.328E+09		
Subset 3 - σ_3^2			3.182E+10		
Deviance	2,936		2,937		

Table 3.4.5.1 Scale and deviance values and parameter estimates for models 2.0 and 3.0.

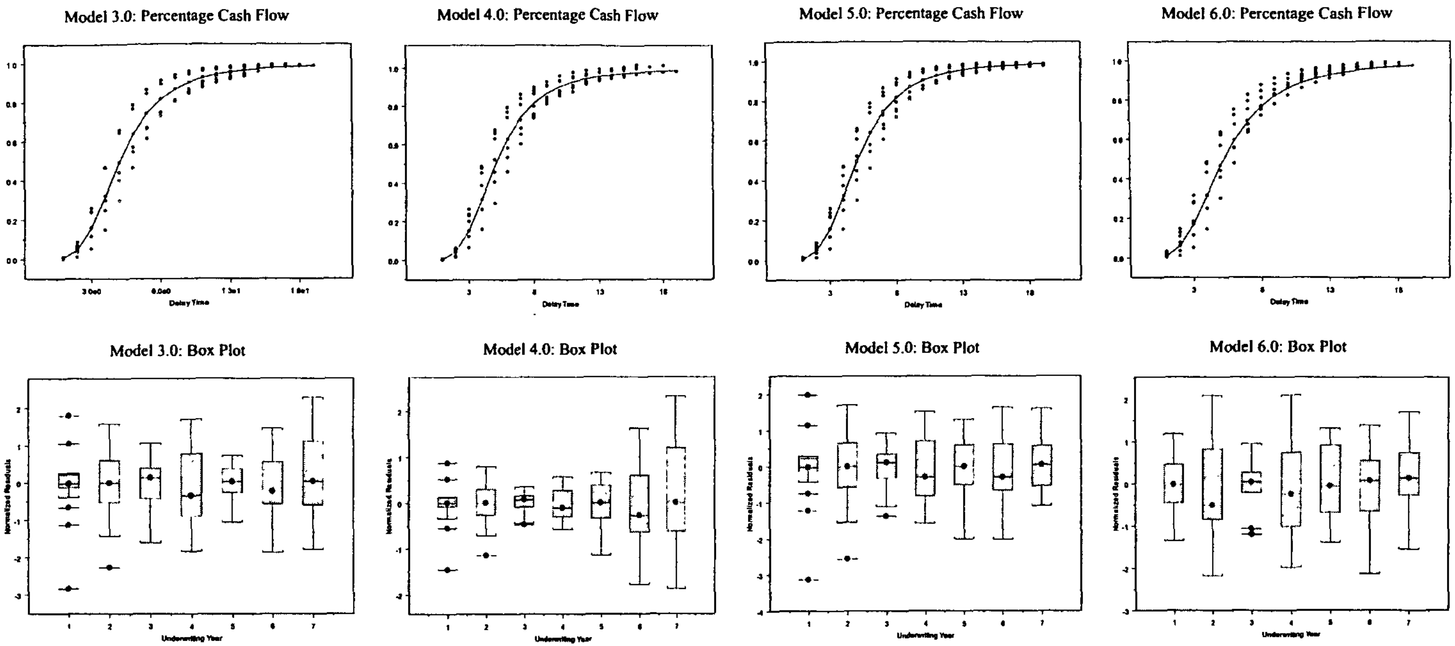
The table compares parameters L, l_w and $L+l_w$ for both models. Variance function power for model 3.0 is very small. For a model with variance σ_k^2, instead of $\sigma_k^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta)}$, the values of $L+l_w$ are not significantly different. Although that model version is excluded from this paper, it is observed that its results indicate that in model 3.0 information on the data variability structures is mainly contained in σ_k^2, and that neither model successfully deals with claim volume differences between underwriting years.

		Ultimate Claim Amount				Reported *IBNR* (1.9)				Subset Membership For Model 3.0
		Mean	Standard Mean Sq. Pred. Error	Confidence Interval 2.50%	Confidence Interval 97.50%	Mean	Standard Mean Sq. Pred. Error	Confidence Interval 2.50%	Confidence Interval 97.50%	
Model 3.0		Validation Model								
Und. Year	1	4,480,000	122,100	4,239,000	4,721,000	115,200	122,100	-125,400	356,600	1
	2	4,324,000	115,800	4,097,000	4,556,000	-121,600	115,800	-347,900	110,200	1
	3	2,163,000	76,730	2,012,000	2,315,000	-21,340	76,730	-172,300	130,700	2
	4	2,004,000	80,470	1,847,000	2,164,000	-86,010	80,470	-243,900	73,750	2
	5	6,648,000	280,700	6,103,000	7,214,000	358,600	280,700	-186,000	924,700	3
	6	10,300,000	340,500	9,656,000	11,000,000	395,000	340,500	-251,900	1,088,000	3
	7	12,630,000	318,900	12,020,000	13,280,000	988,900	318,900	376,800	1,637,000	3
	Total	42,550,000	607,500	41,400,000	43,810,000	1,629,000	607,500	473,400	2,891,000	
Model 4.0		Validation Model - (AR) error structure								
Und. Year	1	4,451,000	164,000	4,131,000	4,776,000	86,840	164,000	-233,200	411,400	
	2	4,315,000	159,700	4,002,000	4,630,000	-130,300	159,700	-443,400	185,000	
	3	2,149,000	157,300	1,842,000	2,462,000	-34,560	157,300	-342,200	277,800	
	4	1,995,000	165,100	1,677,000	2,322,000	-94,990	165,100	-413,700	231,400	
	5	6,635,000	180,900	6,277,000	6,989,000	345,800	180,900	-11,950	699,800	
	6	10,130,000	211,700	9,732,000	10,570,000	226,500	211,700	-176,300	660,900	
	7	12,590,000	229,900	12,130,000	13,030,000	945,900	229,900	486,100	1,391,000	
	Total	42,270,000	469,500	41,340,000	43,190,000	1,345,000	469,500	420,800	2,264,000	
Model 5.0		Final Model								
Und. Year	1	4,477,000	114,300	4,254,000	4,703,000	112,700	114,300	-110,900	338,300	
	2	4,321,000	105,800	4,112,000	4,527,000	-124,800	105,800	-333,000	81,830	
	3	2,162,000	87,570	1,990,000	2,336,000	-22,300	87,570	-194,300	151,900	
	4	2,001,000	92,290	1,823,000	2,188,000	-89,490	92,290	-267,900	97,450	
	5	6,645,000	164,000	6,327,000	6,971,000	356,000	164,000	38,250	681,600	
	6	10,190,000	328,100	9,576,000	10,870,000	280,200	328,100	-332,600	965,000	
	7	12,750,000	513,700	11,770,000	13,800,000	1,110,000	513,700	127,000	2,162,000	
	Total	42,550,000	675,900	41,240,000	43,920,000	1,623,000	675,900	313,300	2,994,000	
Model 6.0		Final Model								
Und. Year	1	4,418,000	37,320	4,345,000	4,492,000	53,860	37,320	-19,210	127,700	
	2	4,426,000	42,000	4,338,000	4,504,000	-19,530	42,000	-107,100	58,360	
	3	2,201,000	26,940	2,146,000	2,253,000	16,980	26,940	-38,450	69,290	
	4	2,094,000	36,930	2,016,000	2,161,000	3,306	36,930	-74,000	70,160	
	5	6,525,000	70,370	6,395,000	6,673,000	235,700	70,370	106,000	384,200	
	6	10,470,000	231,800	10,040,000	10,960,000	560,000	231,800	136,600	1,054,000	
	7	12,370,000	283,300	11,880,000	13,010,000	724,400	283,300	235,500	1,365,000	
	Total	42,500,000	393,100	41,800,000	43,380,000	1,575,000	393,100	876,400	2,453,000	

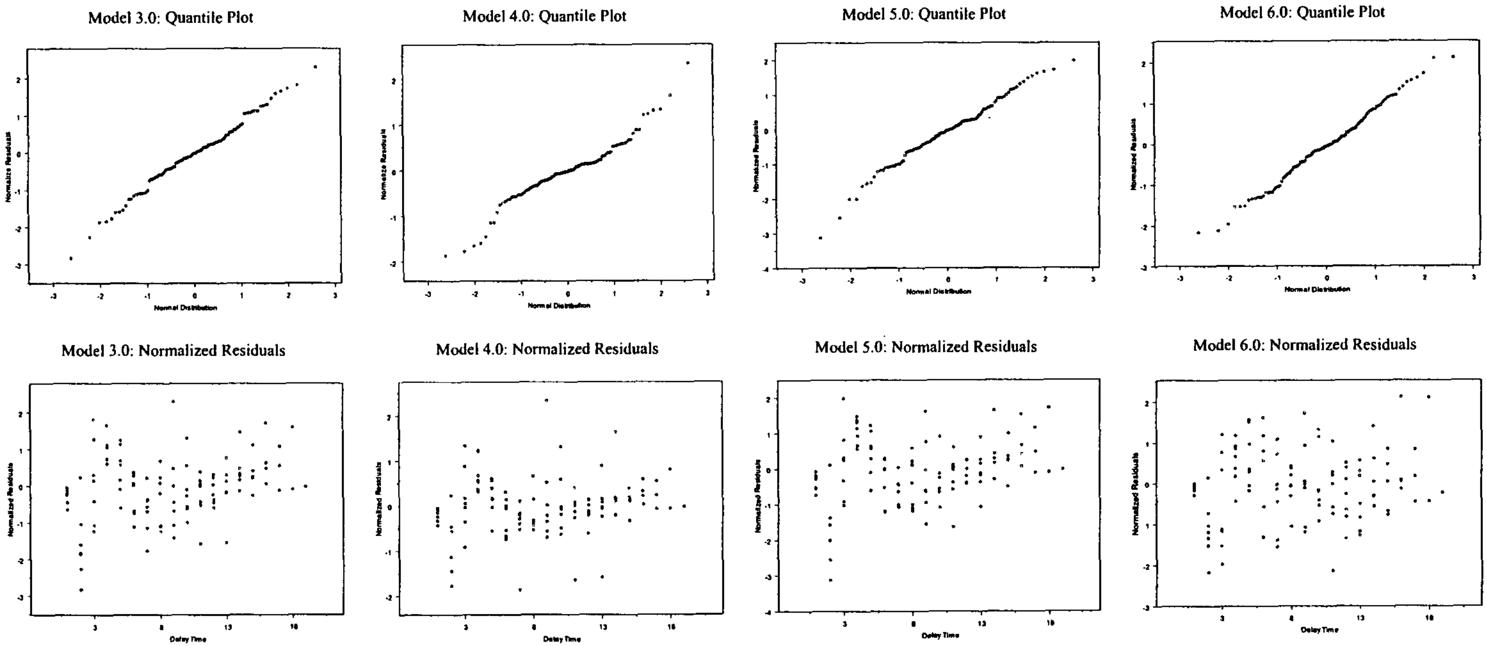
Table 3.4.5.2 Models 3.0 to 6.0: Ultimate losses and *IBNR* predictions and predictive distributions.



Graph 3.4.5.1 Models 3.0 to 6.0: Kernel densities and predictive distributions for ultimate losses and *IBNR* totals.



Graph 3.4.5.2 Models 3.0 to 6.0: Scatter plots and average array curve for percentage cash flow versus delay time and Box plots of normalized residuals.



Graph 3.4.5.3 Models 3.0 to 6.0: Quantile and scatter plots of normalized residuals.

Model	Preliminary models																		Devi.
	Predictive distributions								Log Likelihood			AIC			BIC				
	Ultimate Claim Amount				*IBNR* (1.8)														
	Mean	Standard Deviation	Predictive Interval 2.5%	97.5%	Mean	Standard Deviation	Predictive Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%		
1.0	42,600,000	607,100	41,940,000	44,830,000	1,912,000	499,400	1,546,000	3,915,000											2,980
2.0	42,570,000	471,100	41,720,000	43,590,000	1,873,000	268,600	1,421,000	2,502,000											2,930

Model	Validation models 3.0 and 4.0 and final models 5.0 and 6.0																	Devi.
	Predictive distributions								Log Likelihood			AIC			BIC			
	Ultimate Claim Amount				Reported *IBNR* (1.9)													
	Mean	Standard Deviation	Predictive Interval 2.5%	97.5%	Mean	Standard Deviation	Predictive Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	
3.0	42,550,000	607,500	41,400,000	43,810,000	1,629,000	607,500	473,400	2,891,000	56.0	42.3	71.6	181.9	154.7	213.1	277.1	249.8	308.3	2,936
4.0	42,270,000	469,500	41,340,000	43,190,000	1,345,000	469,500	420,800	2,264,000	56.0	42.6	71.7	182.0	155.1	213.5	277.2	250.3	308.6	2,902
5.0	42,550,000	675,900	41,240,000	43,920,000	1,623,000	675,900	313,300	2,994,000	56.0	42.3	71.6	183.9	156.6	215.1	281.8	254.4	313.0	2,928
6.0	42,500,000	393,100	41,800,000	43,380,000	1,575,000	393,100	876,400	2,453,000	56.0	42.3	71.5	187.9	160.5	219.0	291.2	263.8	322.3	2,887

Table 3.4.5.3 Comparison of results for models 1.0 to 6.0.

The ultimate claim amount and *IBNR* predictions for final models 5.0 and 6.0 and preliminary models 3.0 and 4.0 are compared in table 3.4.5.2. The boxplots for model 5.0 are the most consistent with those of model 2.0. (see graphs 3.4.5.2 and 3.3.3), but the predictive intervals are slightly wider than for models 3.0 and 4.0. The autoregressive error structure in model 4.0 is insufficient to deal with scale variability. In contrast to model 6.0, models 2.0 to 5.0 do not resolve the downwards pattern in the quantile plots (see graph 3.4.5.3). From the percentage cash flow plots and the array average percentage cash flow curve it is evident that the curve is representative of the array. The additional variance parameters increase the AIC and BIC values with respect to model 3.0, but decrease the deviance (table 3.4.5.3). The slight skewness of the *IBNR* and the ultimate claim amount kernel densities for model 2.0 is no longer so evident in models 3.0 to 6.0 (see graphs 3.3.2 and 3.4.5.1).

4. MULTI-ARRAY MODELS

To explore data variability structures and illustrate the process of designing multiple-array models, two mean response functions are used. For the preliminary models the variance functions considered are σ^2, σ_r^2 and σ_{rw}^2, denoting the three model versions by a, b, and c respectively. In section 4.2, observations on the models and numerical examples highlight the motivation for their inclusion. In section 4.3 the values of σ_{rw}^2 are analysed and the final multi-array models are introduced. Numerical examples and assessment of the final models are given in sections 4.5 and 4.6.

4.1 Examples Of Preliminary Multi-Array Models

4.1.1 Models 7.0 (a), (b) and (c)

Model 7.0 is proposed as example of hierarchical reserving models with a limited number of parameters in the percentage cash flow function. It is followed by two amended versions selected to further explore variability patterns in the data.

Model 7.0(a)

For claims array r and underwriting year w, the first homoscedastic model at delay time t is defined as follows:

$$Y_{r,w,t} = \mu_{r,w,t}(\phi_{rw}) + \varepsilon_{r,w,t}$$

such that,

$$\mu_{r,w,t}\left(\phi_{rw}\right)=\exp\left(L+l^{*}_{rw}\right)\left\{1-\frac{\left(1+2\left(\frac{\exp\left(D+d_r+d^{*}_{rw}\right)}{\exp\left(Kc+kc_r+kc^{*}_{rw}\right)}\right)\ln\left(t+\frac{\exp\left(Ks_1\right)}{t^{\exp\left(Ks_2\right)}}\right)\right)}{\dfrac{\exp\left(\left(\exp\left(Kc+kc_r+kc^{*}_{rw}\right)+1\right)\ln\left(t+\frac{\exp\left(Ks_1\right)}{t^{\exp\left(Ks_2\right)}}\right)\right)}{\exp\left(-\exp\left(D+d_r+d^{*}_{rw}\right)\left(\ln\left(t+\frac{\exp\left(Ks_1\right)}{t^{\exp\left(Ks_2\right)}}\right)\right)^2\right)}}\right\} \tag{4.1}$$

and

$$\begin{aligned}&\phi_{rw}=A_{rw}\beta+B_{r,w}b_r+B_{rw}b_{rw}\\&\varepsilon_{r,w}\left|\phi_{rw},\sigma\sim N\left(0,\sigma^2 I_{n_{rw}\times n_{rw}}\right)\right.\end{aligned} \tag{4.2}$$

where

$$\begin{aligned}b_{rw}&=\left[l^{*}_{rw},d^{*}_{rw},kc^{*}_{rw}\right]^T, &\quad b_{rw}\left|b^{*}_{rw},\Sigma^{\circ}_r\sim MVN\left(b^{*}_{rw},\Sigma^{\circ}_r\right)\right.\\b_r&=\left[d_r,kc_r\right]^T, &\quad b_r\left|b^{*}_r,\Sigma_r\sim MVN\left(b^{*}_r,\Sigma_r\right)\right.\\\beta&=\left[Ks_1,Ks_2L,D,Kc\right]^T, &\quad \beta\left|\beta^{*},\Sigma_0\sim MVN\left(\beta^{*},\Sigma_0\right)\right.\end{aligned}$$

The configuration of the design matrices is determined by the order of the parameters in the fixed and random effect parameter vectors. For known parameters $\beta^{**},b^{**}_{rw},b^{**}_r,\Sigma^{*}_0,\Sigma^{*\circ}_r,\Sigma^{*}_r,\Sigma^{**}_0,\Sigma^{\sigma**}_r,\Sigma^{**}_r$, the hyperprior distributions are:

$$\begin{aligned}&\beta^{*}\left|\beta^{**},\Sigma^{**}_0\sim MVN\left(\beta^{**},\Sigma^{**}_0\right)\right. &\quad &\left(\Sigma_0\right)^{-1}\left|\Sigma^{*}_0\sim Wi\left(\left(5\Sigma^{*}_0\right)^{-1},5\right)\right.\\&b^{*}_{rw}\left|b^{**}_{rw},\Sigma^{\sigma**}_r\sim MVN\left(b^{**}_{rw},\Sigma^{\sigma**}_r\right)\right. &\quad &\left(\Sigma^{\circ}_r\right)^{-1}\left|\Sigma^{*\circ}_r\sim Wi\left(\left(3\Sigma^{*\circ}_r\right)^{-1},3\right)\right.\\&b^{*}_r\left|b^{**}_r,\Sigma^{**}_r\sim MVN\left(b^{**}_r,\Sigma^{**}_r\right)\right. &\quad &\left(\Sigma_r\right)^{-1}\left|\Sigma^{*}_r\sim Wi\left(\left(2\Sigma^{*}_r\right)^{-1},2\right)\right.\end{aligned}$$

and $\frac{1}{\sigma^2}\sim Ga(0.001,0.001)$. The claims process functions $C_{r,w}$ and $P_{r,w,t}$ for model 4.1 and the related survival and hazard functions $S_{r,w,t}$ and $h_{r,w,t}$ are:

$$C_{r,w} = \exp\left(L + l_{rw}^{*}\right)$$

$$P_{r,w,t} = \left\{ 1 - \frac{\dfrac{\left(1 + 2\left(\dfrac{\exp\left(D + d_r + d_{rw}^{*}\right)}{\exp\left(Kc + kc_r + kc_{rw}^{*}\right)}\right)\ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\right)}{\exp\left(\left(\exp\left(Kc + kc_r + kc_{rw}^{*}\right) + 1\right)\ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\right)}}{\exp\left(-\exp\left(D + d_r + d_{rw}^{*}\right)\left(\ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\right)^2\right)} \right\}$$

$$S_{r,w,t} = 1 - P_{r,w,t}$$

$$h_{r,w,t} = \frac{-\left(1 - \dfrac{\left(2\exp\left(D + d_r + d_{rw}^{*}\right)\ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right) + \left(\exp\left(Kc + kc_r + kc_{rw}^{*}\right) + 1\right)\right)}{\left(\dfrac{\exp\left(Kc + kc_r + kc_{rw}^{*}\right)}{2\exp\left(D + d_r + d_{rw}^{*}\right)} + \ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\right)^{-1}}\right)}{\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\left(\dfrac{\exp\left(Kc + kc_r + kc_{rw}^{*}\right)}{2\exp\left(D + d_r + d_{rw}^{*}\right)} + \ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\right)\left(1 - \dfrac{\exp\left(Ks_1 + Ks_1\right)}{t^{\exp(Ks_2)+1}}\right)^{-1}} \qquad (4.3)$$

Amended Versions Of Model 7.0(a)

In the alternative versions of model 7.0(a), denoted by 7.0(b) and 7.0(c), $\varepsilon_{r,w} \mid \phi_{rw}, \sigma_r \sim N\left(0, \sigma_r^2 I_{n_{rw} \times n_{rw}}\right)$ and $\varepsilon_{r,w} \mid \phi_{rw}, \sigma_{rw} \sim N\left(0, \sigma_{rw}^2 I_{n_{rw} \times n_{rw}}\right)$ replace $\varepsilon_{r,w} \mid \phi_{rw}, \sigma \sim N\left(0, \sigma^2 I_{n_{rw} \times n_{rw}}\right)$.

4.1.2 Models 8.0 (a), (b) and (c)

In model 8.0 the percentage cash flow function has more parameters than model 7.0 to assess if a more flexible percentage cash flow function could produce more reliable *IBNR* predictions. As with model 7.0, three versions are considered.

Model 8.0(a)

For claims array *r*, underwriting year *w* and development time *t* the model is given by:

$$Y_{r,w,t} = \mu_{r,w,t}\left(\phi_{rw}\right) + \varepsilon_{r,w,t}$$

such that,

$$\mu_{r,w,t}(\phi_{rw}) = \exp\left(L + l_{rw}^{*}\right)\left(1 + \frac{\exp\left(D + d_r + d_{rw}^{*}\right)}{\exp\left(\frac{\ln t}{\exp\left(-Kc - kc_r - kc_{rw}^{*}\right)} + \frac{t}{\exp\left(-Kd - kd_r - kd_{rw}^{*}\right)}\right)}\right)^{-1} \tag{4.4}$$

and

$$\begin{gathered}
\phi_{rw} = A_{rw}\beta + B_{r,w}b_r + B_{rw}b_{rw} = d\left(a_{rw}, \beta, b_r, b_{rw}\right) \\
\varepsilon_{r,w}\left|\phi_{rw}, \sigma \sim N\left(0, \sigma^2 I_{n_{rw} \times n_{rw}}\right)\right. \\
\begin{aligned}
b_{rw} &= \left[l_{rw}^{*}, d_{rw}^{*}, kc_{rw}^{*}, kd_{rw}^{*}\right]^T, & b_{rw}\left|b_{rw}^{*}, \Sigma_{rw} \sim MVN\left(b_{rw}^{*}, \Sigma_{rw}\right)\right. \\
b_r &= \left[d_r, kc_r, kd_r\right]^T, & b_r\left|b_r^{*}, \Sigma_r \sim MVN\left(b_r^{*}, \Sigma_r\right)\right. \\
\beta &= \left[L, D, Kc, Kd\right]^T, & \beta\left|\beta^{*}, \Sigma_0 \sim MVN\left(\beta^{*}, \Sigma_0\right)\right.
\end{aligned}
\end{gathered} \tag{4.5}$$

The hyperprior distributions are:

$$\begin{aligned}
\beta^{*}\left|\beta^{**}, \Sigma_0^{**} \sim MVN\left(\beta^{**}, \Sigma_0^{**}\right)\right. \qquad & \left(\Sigma_0\right)^{-1}\left|\Sigma_0^{*} \sim Wi\left(\left(4\Sigma_0^{*}\right)^{-1}, 4\right)\right. \\
b_{rw}^{*}\left|b_{rw}^{**}, \Sigma_{rw}^{**} \sim MVN\left(b_{rw}^{**}, \Sigma_{rw}^{**}\right)\right. \qquad & \left(\Sigma_{rw}\right)^{-1}\left|\Sigma_{rw}^{*} \sim Wi\left(\left(4\Sigma_{rw}^{*}\right)^{-1}, 4\right)\right. \\
b_r^{*}\left|b_r^{**}, \Sigma_r^{**} \sim MVN\left(b_r^{**}, \Sigma_r^{**}\right)\right. \qquad & \left(\Sigma_r\right)^{-1}\left|\Sigma_r^{*} \sim Wi\left(\left(3\Sigma_r^{*}\right)^{-1}, 3\right)\right.
\end{aligned}$$

and $\frac{1}{\sigma^2} \sim Ga(0.001, 0.001)$, such that $\beta^{**}, b_{rw}^{**}, b_r^{**}, \Sigma_0^{*}, \Sigma_r^{*o}, \Sigma_r^{*}, \Sigma_0^{**}, \Sigma_r^{o**}, \Sigma_r^{**}$ are known. Functions $C_{r,w}$ and $P_{r,w,t}$ for model 8.0(a) and the related survival and hazard functions are:

$$C_{r,w} = \exp\left(L + l_{rw}^{*}\right)$$

$$P_{r,w,t} = \left\{1 + \frac{\exp\left(D + d_r + d_{rw}^{*}\right)}{\exp\left(\exp\left(Kc + kc_r + kc_{rw}^{*}\right)\ln t + \exp\left(Kd + kd_r + kd_{rw}^{*}\right)t\right)}\right\}^{-1}$$

$$S_{r,w,t} = 1 - \left\{1 + \frac{\exp\left(D + d_r + d_{rw}^{*}\right)}{\exp\left(\exp\left(Kc + kc_r + kc_{rw}^{*}\right)\ln t + \exp\left(Kd + kd_r + kd_{rw}^{*}\right)t\right)}\right\}^{-1}$$

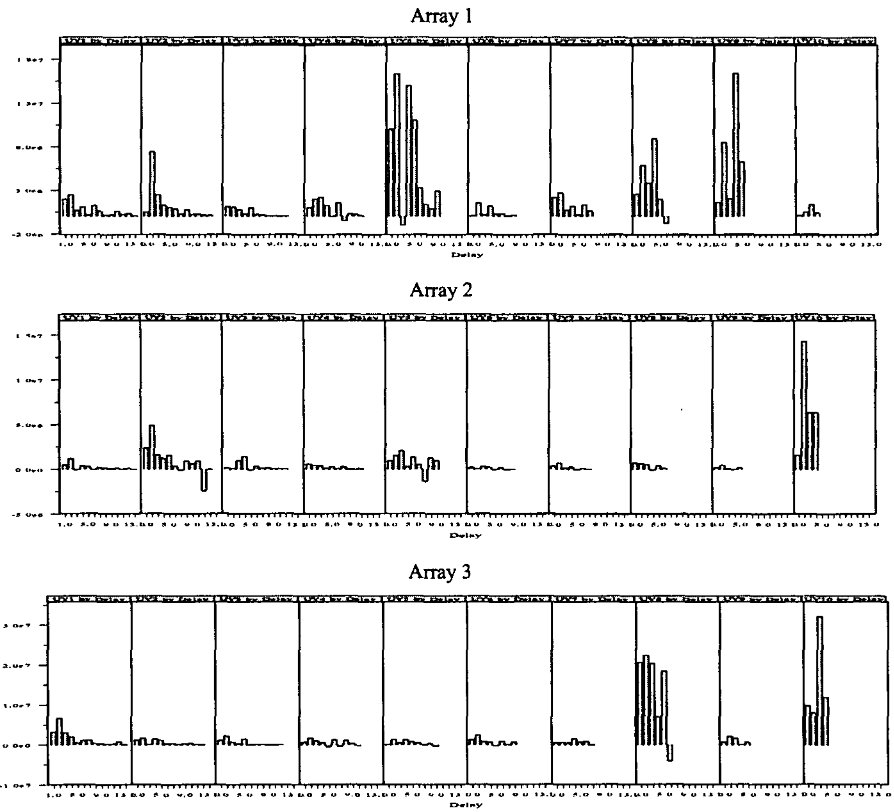
$$h_{r,w,t} = \left(\frac{\exp\left(Kc + kc_r + kc_{rw}^{*}\right)}{t} + \exp\left(Kd + kd_r + kd_{rw}^{*}\right)\right)P_{r,w,t} \tag{4.6}$$

Amended Versions Of Model 8.0(a)

Model versions 8.0(b) and 8.0(c) are derived from 8.0(a) as 7.0(b) and 7.0(c).

4.2 Numerical Examples And Discussion For Preliminary Models 7.0 And 8.0

The claims data selected to illustrate the models in section 4 are reported in tables B.1 to B.3. The data have been obtained through simulations based on a marine portfolio consisting of hull, cargo and aviation hull claims, labelled in graphs and tables as arrays 1, 2 and 3 respectively. Evident from graph 4.2.1 are the data variability and a large number of negative entries in the incremental claims data. Claims reserving models for multiple-array claims portfolios have to explain the variability emerging from the different array characteristics, settlement speeds and exposure levels. The broad range of the cumulative claim totals, from 1,013,800 to 85,287,218, suggest that such claim volume variability may not be effectively captured by the random effects parameters for the mean response model alone.

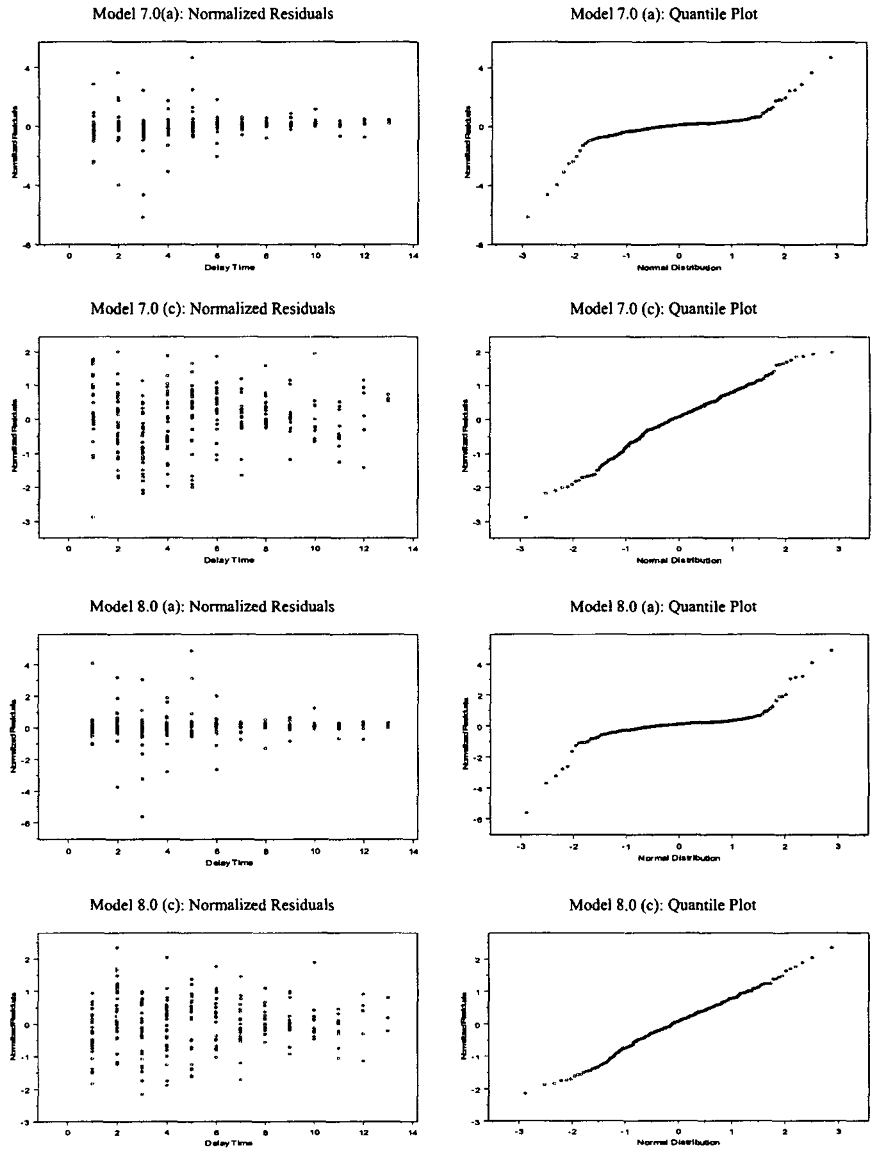


Graph 4.2.1 Incremental data bar plots by array and underwriting year for tables B.1, B.2 and B.3.

	Under. Year	Model 7.0 (c)								σ^2_{rw}
		Ultimate Claim Amount				*IBNR* (1.8)				
		Mean	Standard Mean Sq. Predict Error	Predictive Interval		Mean	Standard Mean Sq. Predict Error	Predictive Interval		
				2.50%	97.50%			2.50%	97.50%	
Array 1	1	11,460,000	560,300	10,250,000	12,530,000	2,130,000	367,600	1,338,000	2,827,000	2.63E+11
	2	18,660,000	880,100	16,790,000	20,390,000	3,103,000	592,200	1,836,000	4,246,000	5.31E+11
	3	5,703,000	363,700	4,959,000	6,413,000	1,006,000	249,000	489,400	1,498,000	8.09E+10
	4	10,520,000	699,000	8,984,000	11,830,000	2,400,000	493,100	1,318,000	3,321,000	2.42E+11
	5	79,440,000	9,587,000	59,530,000	96,770,000	21,190,000	6,833,000	7,172,000	33,700,000	3.25E+13
	6	5,361,000	646,100	3,836,000	6,453,000	1,707,000	479,000	531,200	2,518,000	1.02E+11
	7	10,340,000	1,073,000	8,789,000	12,900,000	2,199,000	788,100	1,256,000	4,081,000	4.53E+11
	8	37,200,000	8,030,000	22,320,000	51,550,000	14,460,000	6,217,000	3,977,000	25,830,000	1.13E+13
	9	57,000,000	22,430,000	20,850,000	100,700,000	28,180,000	17,890,000	4,647,000	63,910,000	7.32E+13
	10	4,479,000	1,919,000	1,427,000	8,180,000	2,451,000	1,601,000	413,800	5,634,000	4.28E+11
Array 2	1	3,259,000	160,000	2,928,000	3,555,000	460,700	105,900	236,100	657,100	2.01E+10
	2	14,920,000	718,600	13,810,000	16,590,000	1,661,000	455,000	1,153,000	2,778,000	7.57E+11
	3	5,504,000	456,900	4,520,000	6,358,000	1,787,000	311,800	1,108,000	2,369,000	1.24E+11
	4	2,888,000	213,500	2,450,000	3,298,000	558,300	148,800	259,500	846,300	2.23E+10
	5	9,255,000	1,209,000	7,091,000	11,620,000	2,261,000	864,500	819,200	3,993,000	5.82E+11
	6	1,530,000	232,800	1,112,000	1,964,000	422,500	171,100	143,200	747,500	1.59E+10
	7	1,896,000	66,610	1,800,000	2,063,000	293,300	44,510	251,100	413,400	4.42E+09
	8	2,892,000	237,600	2,557,000	3,517,000	575,600	168,300	419,200	1,064,000	3.98E+10
	9	1,264,000	133,600	1,083,000	1,612,000	299,500	99,010	213,300	579,900	8.60E+09
	10	50,860,000	12,980,000	28,600,000	72,890,000	23,640,000	10,970,000	7,198,000	42,520,000	1.51E+13
Array 3	1	22,570,000	379,300	21,730,000	23,270,000	3,176,000	271,900	2,555,000	3,660,000	8.60E+10
	2	10,270,000	529,900	9,141,000	11,250,000	2,549,000	351,300	1,813,000	3,210,000	2.12E+11
	3	7,549,000	421,000	6,755,000	8,380,000	1,181,000	288,200	646,200	1,759,000	1.15E+11
	4	7,465,000	748,000	5,903,000	8,828,000	1,741,000	523,900	647,200	2,698,000	2.56E+11
	5	8,308,000	570,900	7,056,000	9,387,000	2,950,000	408,700	2,060,000	3,712,000	1.22E+11
	6	8,368,000	648,500	7,269,000	9,715,000	1,642,000	477,200	917,700	2,646,000	1.67E+11
	7	8,719,000	1,810,000	4,564,000	11,610,000	3,735,000	1,376,000	683,200	5,940,000	5.23E+11
	8	115,700,000	15,020,000	92,600,000	148,100,000	30,480,000	11,610,000	15,520,000	56,200,000	5.84E+13
	9	7,997,000	1,149,000	6,085,000	10,260,000	2,621,000	925,900	1,228,000	4,485,000	2.07E+11
	10	94,890,000	42,140,000	22,790,000	188,100,000	43,200,000	33,030,000	7,345,000	121,300,000	3.62E+14
By Array	Array 1	240,200,000	26,480,000	191,200,000	291,100,000	78,830,000	20,810,000	42,460,000	119,300,000	
	Array 2	94,270,000	13,240,000	71,450,000	117,200,000	31,960,000	11,160,000	14,750,000	51,400,000	
	Array 3	291,800,000	45,850,000	215,600,000	389,100,000	93,280,000	35,830,000	48,710,000	175,000,000	
Total		626,200,000	54,430,000	530,900,000	740,400,000	204,100,000	43,170,000	134,800,000	298,700,000	
Deviance		7,372								
Iterat.: Start +Sample		31,000								

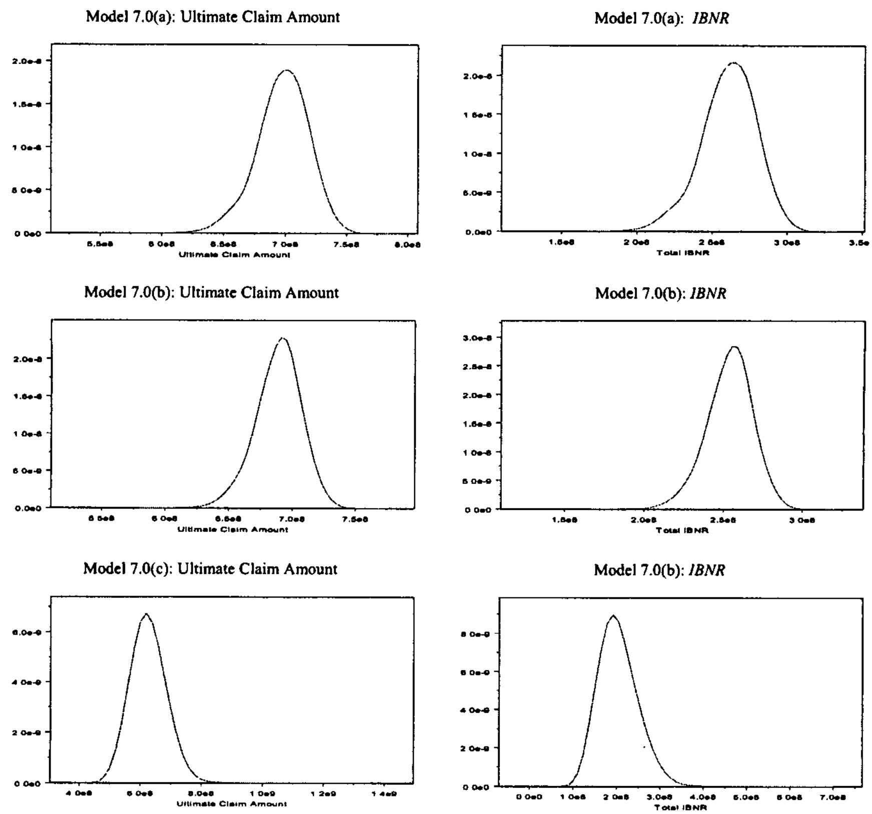
Table 4.2.1 Model statistics, ultimate losses and *IBNR* predictions, and respective predictive distributions for Model 7.0 (c).

Portfolios displaying large differences in exposure levels or claims magnitudes are not at all unusual, even in treaties where underwriting contracts remain unaltered. Cost limitations or timing restrictions may impede exploring methods, possibly able to deal with high variability in exposure volumes, such as analyses at transaction level. In the models proposed, a good fit to historical data as assessment criterion of the preliminary models, is as important as suitable variance models, as the latter determines the stability of *IBNR* and ultimate claim predictions. This is more likely to be achieved by models 7.0(c) and 8.0(c), as inspection of graph 4.2.2 and of actual and fitted claims confirm.

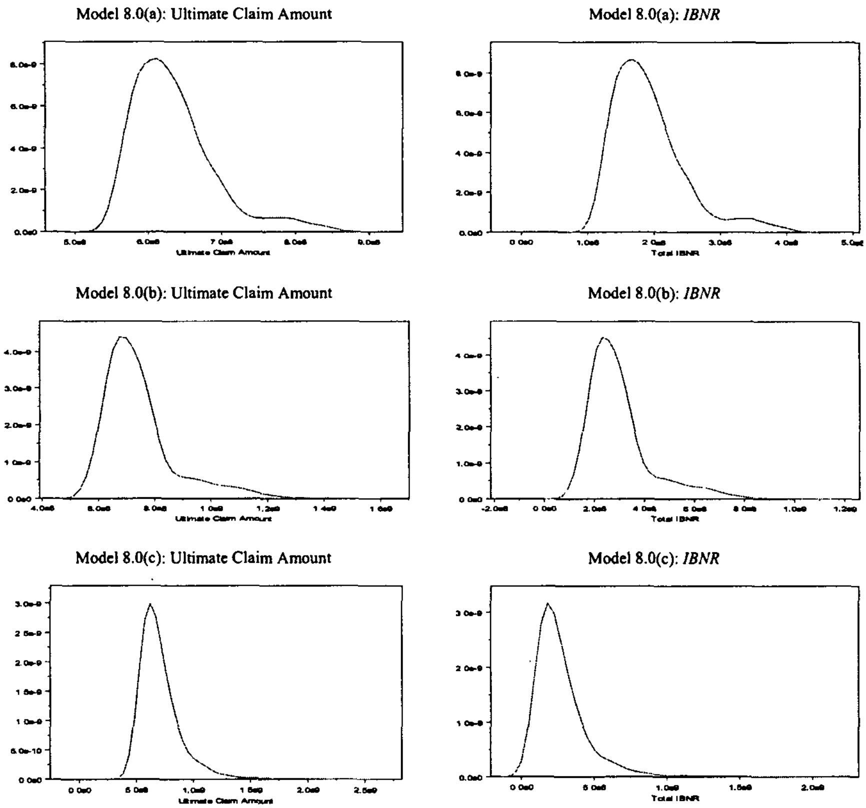


Graph 4.2.2 Normalized residuals and quantile plots for models 7.0 and 8.0 (a) and (c).

Plots for 7.0(b) and 8.0(b) were found to be uninformative, and for this reason were excluded from graph 4.2.2. While the rankings of σ_r^2 in models 7.0(b) and 8.0(b) are consistent, with $\sigma_2^2 < \sigma_3^2 < \sigma_1^2$, the claims volume variability within each array present similar problems to those encountered with model 1.0. According to the quantile plots only the residuals from models with variance function σ_{rw} may satisfy the Shapiro-Wilk test W for near-normality (Shapiro and Wilk (1965)). Model 7.0(c) gives narrower intervals for the mean *IBNR* at underwriting year, array levels and overall. (see table 4.5.3). The close equivalence of ranking orders for σ_{rw}^2 and $c_{r,w}$ (table 4.2.1) confirms the expectation that either $\zeta_{rw} = (\sigma_{rw}, \vartheta, \rho)$ or $\zeta_{rw} = (\sigma, \vartheta_{rw}, \rho)$ could reveal scale variability structures in the data. They do so more effectively than $\zeta_{rw} = (\sigma_{rw}, \vartheta_{rw}, \rho)$. In a variance model $\sigma_r^2 \left(\mu_{r,w,t}(\phi_{rw})\right)^{\exp(\vartheta + \vartheta_r)}$, parameters σ_r and ϑ_r are less informative.



Graph 4.2.3 Preliminary model 7.0: Kernel densities for ultimate losses and *IBNR* totals.



Graph 4.2.4 Preliminary model 8.0: Kernel densities for ultimate losses and *IBNR* totals.

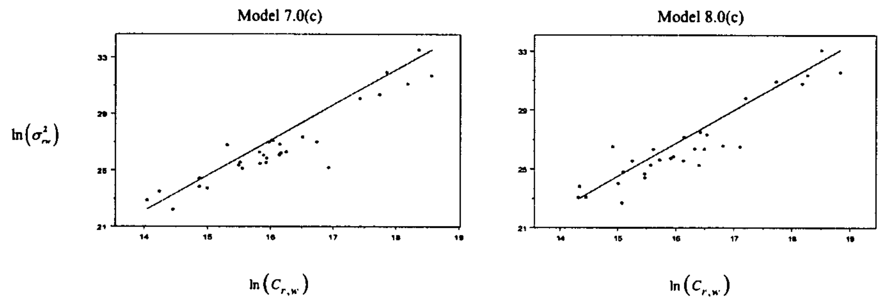
Graphs 4.2.3 and 4.2.4 and table 4.5.3 show that the kernels for mean *IBNR* and ultimate claim predictions are skewed. In the next section it is shown how σ^2_{rw} and $C_{r,w}$ can be used to construct the variance function for the final models.

4.3 Final Multi-Array Models

The preliminary models demonstrate that the data variability can be explored more freely when $\text{var}\left(Y_{r,w}\right) = \sigma^2_{rw}$. The values of σ^2_{rw} and $C_{r,w}$ suggest a variability structure associated to scale differences between underwriting year data sets, around which a cluster structure could be constructed for analytical purposes. However, some management decisions, such as commutations, would require more precise *IBNR* and

$C_{r,w}$ predictions at underwriting year or contract levels. Reconciliation of reserves would be more difficult if the data of interest were not part of the same cluster. A better approach to deal with scale variability, and one that is totally coherent with the generic model, may involve formulating $C_{r,w}$ into the variance model. To assess this, the following regression is applied to the output of models 7.0(c) and 8.0(c):

$$\ln\left(\sigma^2_{r,w}\right) = \delta_1 + \delta_2 \ln\left(C_{r,w}\right) \tag{4.7}$$



Graph 4.3.1 Lines $\delta_1 + \delta_2 \ln\left(C_{r,w}\right)$ and scatter plots of $\ln\left(C_{r,w}\right)$ vs $\ln\left(\sigma^2_{rw}\right)$ on the y-axis.

(4.7) gives $[\delta_1, \delta_2] = [-8.7068, 2.1934]$ for model 7.0(c) and $[\delta_1, \delta_2] = [-6.355, 2.0347]$ for model 8.0(c). Graph 4.3.1 displays the regression lines and the scatter plots of $\ln\left(C_{r,w}\right)$ versus $\ln\left(\sigma^2_{rw}\right)$ for both models. Equation (4.7) suggests that the final models should be

$$\left(Y_{r,w,t} \middle| C_{r,w}, \zeta\right) \sim N\left(\mu_{r,w,t}\left(\phi_{rw}\right), \sigma^2 C^2_{r,w}\right) \tag{4.8}$$

such that $\zeta = \sigma^2$ and $\mu_{r,w,t}\left(\phi_{rw}\right)$ is given by equations (4.1) and (4.4) for models 7.0(d) and 8.0(d) respectively. From regression model (4.7) for model 7.0(c), $\exp(\delta_1) = 0.000165$ could set the initial value for σ^2. The outcome of the analysis is not unexpected. In fact, the inclusion of $C^2_{r,w}$ in the variance function has the effect of normalising the data, hence, reducing the reserving analysis with random effects models to a type of problem that is more consistent with the typical published examples, concerned with the analysis of repeated observations on subjects or trials that share some common characteristics.

See for instance Elashoff et al. (1982) and Aziz et al. (1978). For the final models the mean *IBNR* and ultimate claim amount predictions are replaced by estimates generated by their predictive distributions. The reported *IBNR* values are calculated along the lines of (1.9). Extending the definition of section 3.2, the portfolio average model for the percentage cash flow is given below.

4.4 Portfolio And Array Average Models For The Percentage Cash Flow

Section 3.2 identifies the percentage cash flow as the most suitable function in the reserving model where concepts on inferences on marginal distribution or population average models could be applied. Comparisons across a claims portfolio are more meaningful at percentage cash flow level. As observed in section 3, a model may be able to fit the data well even when the percentage cash flow function converges to a value different to 1. However, in such cases the ultimate claim amount and *IBNR* predictions would be incorrect.

To formulate the average models for the percentage cash flow the parameter vectors for the portfolio and array average models, ϕ_p and ϕ_A, are respectively defined:

$$\phi_p = A\beta + B_1\left(\frac{1}{r_1}\sum_{r=1}^{r_1} b_r\right) + B_2\left(\left(\sum_{r=1}^{r_1} u_r\right)^{-1}\sum_{r=1}^{r_1}\sum_{w=1}^{u_r} b_{rw}\right)$$

and

$$\phi_{A_r} = A\beta + B_1 b_r + B_2\left((u_r)^{-1}\sum_{w=1}^{u_r} b_{rw}\right)$$

such that, design matrices A_{rw}, $B_{r.w}$ and B_{rw} are replaced respectively by A, B_1 and B_2. Consider for example the mean response function for model 7.0(d). If

$$\begin{aligned} D_p &= D + \frac{1}{r_1}\sum_{r=1}^{r_1} d_r + \left(\sum_{r=1}^{r_1} u_r\right)^{-1}\sum_{r=1}^{r_1}\sum_{w=1}^{u_r} d^{*}_{rw} \\ Kc_p &= Kc + \frac{1}{r_1}\sum_{r=1}^{r_1} kc_r + \left(\sum_{r=1}^{r_1} u_r\right)^{-1}\sum_{r=1}^{r_1}\sum_{w=1}^{u_r} kc^{*}_{rw} \end{aligned} \tag{4.9}$$

such that $D, Kc, d_r, kc_r, d^{*}_{rw}$ and kc^{*}_{rw} are the percentage cash flow function parameters, then the portfolio average for the percentage cash flow function at time t is given by

$$P_{p,t}=1-\frac{\left(1+2\left(\frac{\exp(D_p)}{\exp(Kc_p)}\right)\ln\left(t+\frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)\right)}{\exp\left(\left(\exp(Kc_p)+1\right)\ln\left(t+\frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)\right)}\exp\left(-\exp(D_p)\left(\ln\left(t+\frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)\right)^2\right) \quad (4.10)$$

Additional insight may be gained by including in the plots a curve for the percentage cash flow average model for each array. Continuing with the example, for the array average model for array r in model 7.0(d)

$$D_{A_r}=D+d_r+\frac{1}{u_r}\sum_{w=1}^{u_r}d_{rw}^{*}$$

$$Kc_{A_r}=Kc+kc_r+\frac{1}{u_r}\sum_{w=1}^{u_r}kc_{rw}^{*}$$

should replace D_p Kc_p in equation (4.10).

4.5 Numerical Examples And Discussion For Models 7.0(d) And 8.0(d)

Models 7.0(d) and 8.0(d) provide close fit to the data. The portfolio reported *IBNR* and ultimate claim predictions for 7.0(d) and 8.0(d) are given on tables 4.5.1 and 4.5.2 and summarised on table 4.5.3. They show that the final models' predictive intervals are narrower than for their earlier model versions. At underwriting year level, the mean response function of model 7.0 is still the most useful of the two (see table 4.5.1).

Graph 4.5.2 compares scatter plots for the percentage cash flow values for both models and shows that model 8.0(d) is the least successful in separating the volume and development pattern elements in the data. Note that the graphs' scales are not the same and that the projection period for model 7.0(d) is longer than for model 8.0(d). The predictive interval for $\sum_{r=1}^{3}\sum_{w=1}^{10}C_{r,w}$ for model 8.0(d) is also wider. (See tables 4.5.2 and 4.5.3 and graph 4.5.1). Evident from graph 4.5.2 is the settlement speeds variability. A reduction in the reported *IBNR* predictive intervals is consistent with a reduction of the normalized residuals and the Bayesian Information Criterion. Particularly relevant to the claims process is the systematic correction of historical errors as claims evolve, since negative incremental entries frequently adjust earlier overstated claim entries. The box, scatter and quantile plots make apparent data anomalies generated by negative adjustments to paid claims and by large claim volume differences. As in section 3, neither can be addressed with autoregressive error structures. The negative incremental claim entries are responsible for the some of the outliers and, in particular, for the slight depression in the quantile plots, between –2 and –1 of the horizontal axis.

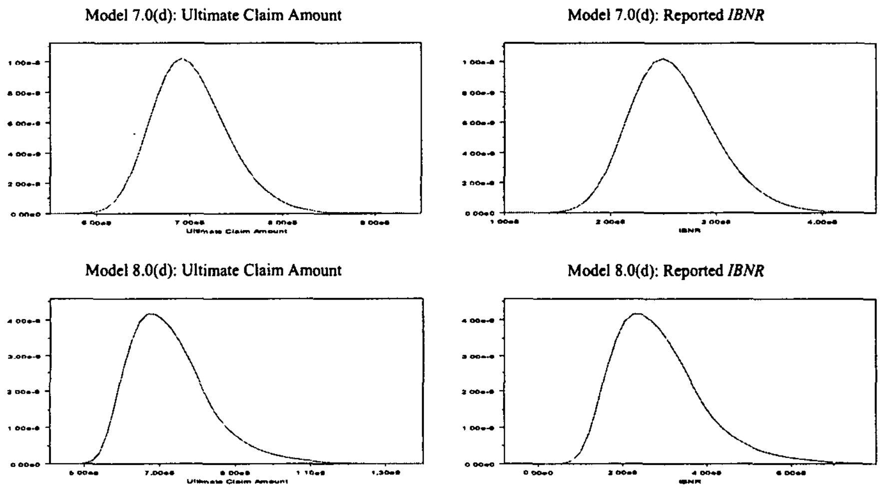
		Model 7.0(d)							
		Ultimate Claim Amount				Reported *IBNR* (1.9)			
	Under. Year	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.50%	Predictive Interval 97.50%	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.50%	Predictive Interval 97.50%
Array 1	1	11,210,000	769,200	9,777,000	12,780,000	1,529,000	769,200	95,310	3,099,000
	2	17,710,000	1,249,000	15,420,000	20,320,000	2,353,000	1,249,000	56,370	4,962,000
	3	5,444,000	413,800	4,704,000	6,318,000	872,000	413,800	131,300	1,746,000
	4	10,080,000	837,600	8,526,000	11,800,000	2,233,000	837,600	681,600	3,954,000
	5	82,510,000	7,098,000	69,530,000	97,320,000	22,950,000	7,098,000	9,971,000	37,750,000
	6	5,419,000	523,300	4,480,000	6,556,000	1,814,000	523,300	874,500	2,951,000
	7	9,677,000	708,000	8,497,000	11,350,000	864,700	708,000	-315,500	2,534,000
	8	41,090,000	5,069,000	32,640,000	52,560,000	18,800,000	5,069,000	10,350,000	30,270,000
	9	77,960,000	12,940,000	57,800,000	108,600,000	43,250,000	12,940,000	23,090,000	73,890,000
	10	5,935,000	1,329,000	3,974,000	9,123,000	3,690,000	1,329,000	1,730,000	6,879,000
Array 2	1	2,954,000	159,900	2,655,000	3,284,000	82,640	159,900	-217,200	412,100
	2	14,730,000	867,600	13,140,000	16,560,000	2,603,000	867,600	1,004,000	4,431,000
	3	5,511,000	416,500	4,745,000	6,370,000	1,968,000	416,500	1,202,000	2,828,000
	4	2,795,000	234,300	2,377,000	3,289,000	388,300	234,300	-30,250	881,400
	5	9,708,000	836,200	8,198,000	11,490,000	1,928,000	836,200	418,600	3,708,000
	6	1,610,000	165,700	1,319,000	1,967,000	510,000	165,700	218,700	867,300
	7	1,903,000	110,600	1,695,000	2,130,000	315,000	110,600	106,900	541,700
	8	2,878,000	206,000	2,527,000	3,345,000	489,700	206,000	139,200	956,900
	9	1,256,000	130,000	1,089,000	1,527,000	242,700	130,000	74,990	513,500
	10	55,000,000	12,350,000	37,090,000	86,390,000	26,260,000	12,350,000	8,350,000	57,640,000
Array 3	1	21,060,000	1,339,000	18,570,000	23,870,000	1,494,000	1,339,000	-1,002,000	4,302,000
	2	9,372,000	709,500	8,070,000	10,840,000	1,601,000	709,500	298,900	3,064,000
	3	7,253,000	509,800	6,333,000	8,330,000	1,012,000	509,800	92,520	2,089,000
	4	7,430,000	611,500	6,300,000	8,685,000	1,447,000	611,500	316,900	2,702,000
	5	7,978,000	683,900	6,742,000	9,428,000	2,968,000	683,900	1,731,000	4,417,000
	6	7,794,000	585,200	6,785,000	9,090,000	509,400	585,200	-499,700	1,806,000
	7	9,066,000	992,900	7,317,000	11,150,000	3,852,000	992,900	2,103,000	5,936,000
	8	115,400,000	14,440,000	94,040,000	148,400,000	30,070,000	14,440,000	8,752,000	63,110,000
	9	7,423,000	1,083,000	6,060,000	10,240,000	1,912,000	1,083,000	549,400	4,731,000
	10	144,200,000	30,660,000	100,000,000	217,200,000	82,050,000	30,660,000	37,860,000	155,100,000
By Array	Array 1	267,000,000	15,930,000	239,600,000	302,600,000	98,350,000	15,930,000	70,950,000	133,900,000
	Array 2	98,350,000	12,500,000	80,030,000	130,000,000	34,790,000	12,500,000	16,460,000	66,480,000
	Array 3	336,900,000	34,040,000	284,200,000	416,100,000	126,900,000	34,040,000	74,190,000	206,100,000
Total		702,300,000	39,690,000	636,000,000	790,200,000	260,100,000	39,690,000	193,700,000	347,900,000
σ^2		0.002308							
Deviance		7,365							
Iterat.: Start +Sample		29,500							

Table 4.5.1 Model 7.0(d): statistics, ultimate loss and reported *IBNR* predictions, and predictive intervals.

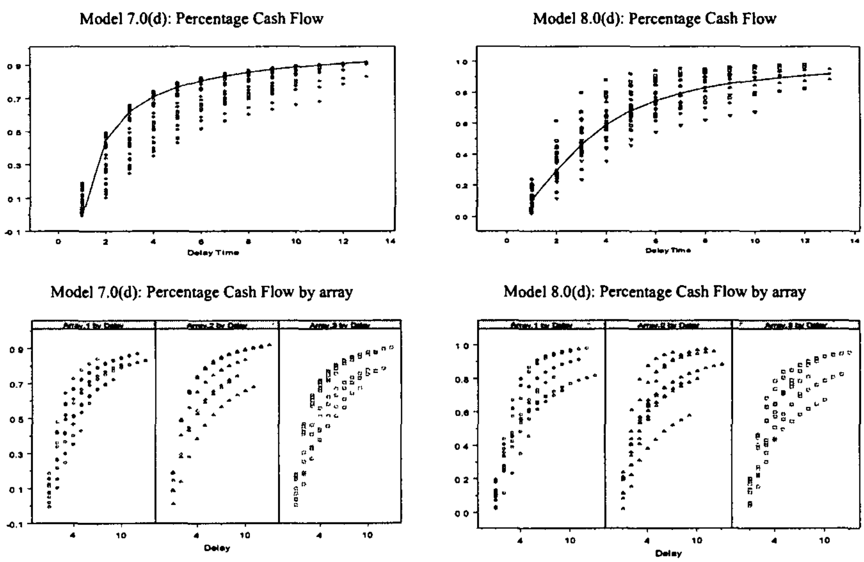
	Under. Year	Model 8.0(d)							
		Ultimate Claim Amount				Reported *IBNR* (1.9)			
		Mean	Standard Mean Sq. Predict. Error	Predictive Interval		Mean	Standard Mean Sq. Predict. Error	Predictive Interval	
				2.50%	97.50%			2.50%	97.50%
Array 1	1	11,650,000	1,272,000	9,664,000	14,610,000	1,964,000	1,272,000	-17,750	4,926,000
	2	15,330,000	925,100	13,650,000	17,290,000	-33,230	925,100	-1,713,000	1,931,000
	3	5,257,000	529,300	4,491,000	6,648,000	684,300	529,300	-81,330	2,075,000
	4	8,169,000	518,100	7,230,000	9,284,000	324,400	518,100	-614,400	1,439,000
	5	88,310,000	24,820,000	62,930,000	159,600,000	28,750,000	24,820,000	3,371,000	100,100,000
	6	5,174,000	2,195,000	3,529,000	11,800,000	1,569,000	2,195,000	-76,040	8,196,000
	7	12,640,000	4,024,000	8,691,000	24,190,000	3,827,000	4,024,000	-121,500	15,380,000
	8	43,080,000	20,980,000	24,940,000	94,960,000	20,790,000	20,980,000	2,646,000	72,670,000
	9	100,800,000	56,840,000	49,440,000	261,700,000	66,120,000	56,840,000	14,730,000	226,900,000
	10	3,732,000	2,999,000	2,169,000	11,610,000	1,488,000	2,999,000	-75,410	9,362,000
Array 2	1	3,378,000	642,300	2,746,000	5,221,000	506,300	642,300	-125,300	2,349,000
	2	13,820,000	1,035,000	12,190,000	16,230,000	1,690,000	1,035,000	58,680	4,094,000
	3	3,567,000	214,600	3,174,000	4,022,000	24,380	214,600	-368,000	479,200
	4	3,102,000	537,100	2,413,000	4,534,000	694,700	537,100	5,700	2,126,000
	5	13,150,000	5,018,000	7,527,000	26,000,000	5,372,000	5,018,000	-252,300	18,220,000
	6	1,679,000	615,700	1,125,000	3,443,000	579,000	615,700	24,490	2,343,000
	7	1,726,000	386,700	1,429,000	2,971,000	137,900	386,700	-159,100	1,383,000
	8	3,693,000	1,482,000	2,337,000	7,675,000	1,305,000	1,482,000	-51,290	5,287,000
	9	1,718,000	1,104,000	880,500	4,346,000	703,800	1,104,000	-133,300	3,332,000
	10	45,620,000	29,960,000	26,490,000	132,500,000	16,880,000	29,960,000	-2,253,000	103,800,000
Array 3	1	19,970,000	1,316,000	17,640,000	22,840,000	407,000	1,316,000	-1,924,000	3,271,000
	2	9,249,000	996,600	7,773,000	11,710,000	1,477,000	996,600	1,074	3,937,000
	3	7,050,000	1,353,000	5,927,000	11,690,000	809,000	1,353,000	-314,100	5,446,000
	4	9,812,000	2,500,000	6,652,000	16,390,000	3,829,000	2,500,000	668,500	10,410,000
	5	5,939,000	756,800	5,010,000	7,855,000	928,500	756,800	-607	2,844,000
	6	9,454,000	2,978,000	6,651,000	17,840,000	2,169,000	2,978,000	-633,000	10,560,000
	7	8,073,000	2,031,000	5,703,000	13,670,000	2,859,000	2,031,000	489,100	8,458,000
	8	125,800,000	37,370,000	89,280,000	229,700,000	40,550,000	37,370,000	3,992,000	144,400,000
	9	6,229,000	1,520,000	4,957,000	10,410,000	717,800	1,520,000	-554,400	4,897,000
	10	141,100,000	60,410,000	75,900,000	293,600,000	78,980,000	60,410,000	13,760,000	231,500,000
By Array	Array 1	294,200,000	66,190,000	217,800,000	475,200,000	125,500,000	66,190,000	49,160,000	306,500,000
	Array 2	91,450,000	30,130,000	68,590,000	177,000,000	27,890,000	30,130,000	5,030,000	113,400,000
	Array 3	342,700,000	72,200,000	252,700,000	525,000,000	132,700,000	72,200,000	42,750,000	315,000,000
Total		728,300,000	102,600,000	584,700,000	984,300,000	286,100,000	102,600,000	142,500,000	542,000,000
σ^2		0.002137							
Deviance		7,323							
Iterat.: Start +Sample		29,500							

Table 4.5.2 Model 8.0(d): statistics, ultimate loss and reported *IBNR* predictions, and predictive intervals.

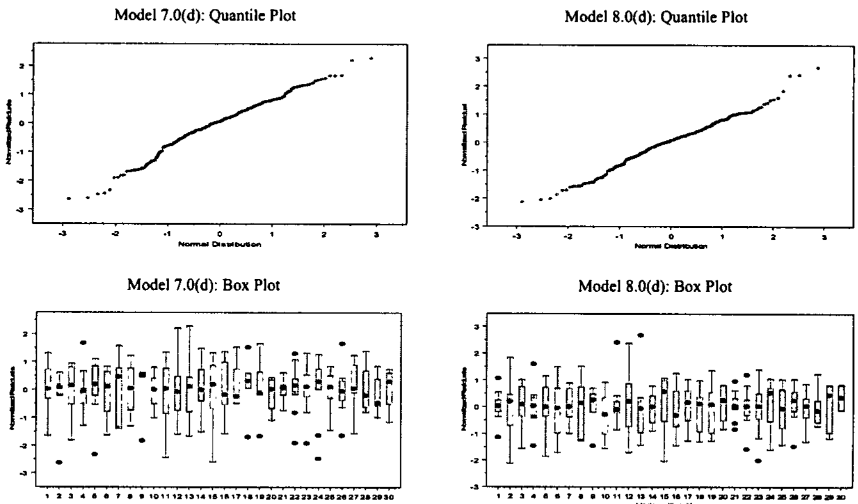
Historical claims add to 442,249,345. The difference between the ultimate claim amounts and the reported *IBNR* predictions for the final models are approximately 442 million. The order of accuracy in the WinBugs system prevents an exact reconciliation with the total claim amount to date. When model 7.0(d) is appraised for consistency with an analysis by array, the ultimate claim amount and reported *IBNR* predictions show respectively 1.1% and 3.1% overall difference from the predictions on table 4.5.1. In section 4.6 the hazard rate profile extracted from the model is discussed.



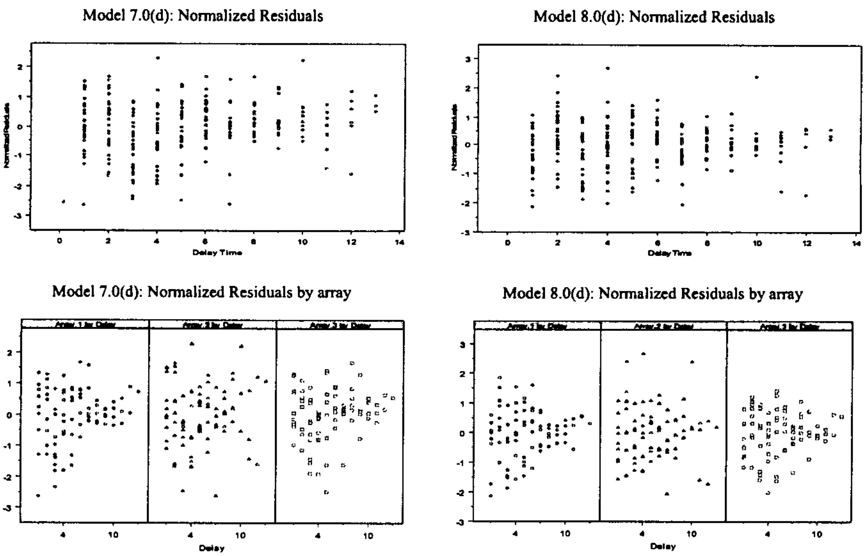
Graph 4.5.1 Kernel densities and predictive distributions for ultimate losses and reported *IBNR*.



Graph 4.5.2 Models 7.0(d) and 8.0(d): Scatter plots and average portfolio curve for percentage cash flow versus delay time.



Graph 4.5.3 Quantile plots and box plots by underwriting year. Underwriting years are labelled 1 to 30. The first 10 correspond to the marine hull, the next 10 to marine cargo and the last 10 to aviation cargo.



Graph 4.5.4 Models 7.0(d) and 8.0(d): Scatter plots versus delay time, overall and by array.

Model	Preliminary Models																	Devi.
	Distributions								Log Likelihood			AIC			BIC			
	Ultimate Claim Amount				*IBNR* (1.8)													
	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.5%	97.5%	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	
7.0 (a)	697,800,000	20,900,000	651,600,000	735,100,000	261,000,000	18,360,000	219,900,000	293,800,000										7942
7.0 (b)	688,900,000	17,810,000	650,100,000	721,400,000	253,800,000	14,350,000	222,600,000	280,000,000										7,866
7.0 (c)	626,200,000	54,430,000	530,900,000	740,400,000	204,100,000	43,170,000	134,800,000	298,700,000										7,372
8.0 (a)	635,000,000	56,760,000	558,600,000	793,200,000	192,800,000	55,590,000	120,200,000	350,300,000										7,912
8.0 (b)	742,600,000	126,600,000	587,000,000	1,107,000,000	298,100,000	125,500,000	144,500,000	661,200,000										7,816
8.0 (c)	711,200,000	186,900,000	505,300,000	1,190,000,000	274,700,000	182,000,000	88,270,000	746,500,000										7,335

Model	Final Models																	Devi.
	Predictive distributions								Log Likelihood			AIC			BIC			
	Ultimate Claim Amount				Reported *IBNR* (1.9)													
	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.5%	97.5%	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	Mean	Confidence Interval 2.5%	97.5%	
7.0 (d)	702,300,000	39,690,000	636,000,000	790,200,000	260,100,000	39,690,000	193,700,000	347,900,000	127.0	106.0	150.0	452.1	409.5	498.8	802.7	760.1	849.4	7,365
8.0 (d)	728,300,000	102,600,000	584,700,000	984,300,000	286,100,000	102,600,000	142,500,000	542,000,000	127.0	106.0	150.0	520.1	477.6	565.7	991.1	948.6	1,037.0	7,323

Table 4.5.3 Comparison of results for models 7.0 and 8.0.

4.6 Average Hazard Rate For Model 4.1(d)

As a pure loss measure, hazard rate can help comparing underwriting year contracts, to formulate portfolio management strategies, determine future premiums, portfolio composition, commutation or closure policies, etc. Hazard rates by underwriting year, or weighted average hazard rates for each array or for the whole claims portfolio can be derived from a reserving analysis. For payment year τ these are respectively:

$$h_{r,w,\tau-w+1} = -\left(\frac{\partial\left(\ln\left(1-P_{r,w,z}\right)\right)}{\partial z}\right)_{z=\tau-w+1} = \frac{\left(\dfrac{\partial P_{r,w,z}}{\partial z}\right)_{z=\tau-w+1}}{1-P_{r,w,\tau-w+1}} = \frac{C_{r,w}}{IBNR_{(r,w,\tau-w+1)}}\left(\frac{\partial P_{r,w,z}}{\partial z}\right)_{z=\tau-w+1}$$

$$A\bar{h}_{r,\tau} = \sum_{w=1}^{u_r} h_{r,w,\tau-w+1}\left(\frac{IBNR_{(r,w,\tau-w+1)}}{\sum_{k=1}^{u_r} IBNR_{(r,k,\tau-k+1)}}\right) = \sum_{w=1}^{u_r} h_{r,w,\tau-w+1}\left(\frac{IBNR_{(r,w,\tau-w+1)}}{IBNR_r(\tau)}\right)$$

$$G\bar{h}_\tau = \sum_{r=1}^{\eta}\sum_{w=1}^{u_r} h_{r,w,\tau-w+1}\left(\frac{IBNR_{(r,w,\tau-w+1)}}{\sum_{r=1}^{\eta}\sum_{k=1}^{u_r} IBNR_{(r,k,\tau-k+1)}}\right) = \sum_{r=1}^{\eta}\sum_{w=1}^{u_r} h_{r,w,\tau-w+1}\left(\frac{IBNR_{(r,w,\tau-w+1)}}{IBNR(\tau)}\right)$$

Given in terms of the *IBNR*, the above equations make explicit the changing nature of the average hazard as claims evolve. Table 4.6.1 lists the hazard rate values for payment years 13, 15 and 17 for model 7.0(d). Since underwriting year losses are at different stages in their development, a similar table to 4.6.1 can be used to assess the impact on the claims portfolio of, for example, excluding from it underwriting year contracts related to underwriting year j_1 of array i_1 and underwriting year j_2 of array i_2. The average hazard rate for the reduced portfolio becomes:

$$g\bar{h}_\tau = G\bar{h}_\tau + \frac{\sum_{n=1}^{2}\left(G\bar{h}_\tau - h_{i_n,j_n,\tau-j_n+1}\right)IBNR_{(i_n,j_n,\tau-j_n+1)}}{IBNR(\tau) - \sum_{n=1}^{2} IBNR_{(i_n,j_n,\tau-j_n+1)}}$$

The exclusion of the contracts from the claims portfolio reduces the portfolio hazard rate only when

$$\sum_{n=1}^{2}\left(G\bar{h}_\tau - h_{i_n,j_n,\tau-j_n+1}\right)IBNR_{(i_n,j_n,\tau-j_n+1)} < 0$$

Table 4.6.1 shows that the exclusion of underwriting year 10 from any of the three arrays would reduce the portfolio average hazard rate. The removal of underwriting year data sets belonging to any of the first seven underwriting years would have the opposite effect. While not included in table 4.6.1, from the reserving analysis the full distribution for the hazard rates can be obtained.

Model 7.0(d)					
			Hazard rates		
	Underwriting Year	Ultimate Loss	Payment Year 13	Payment Year 15	Payment Year 17
Array 1	1	11,230,000	0.0515	0.0456	0.0410
	2	17,660,000	0.0548	0.0482	0.0430
	3	5,458,000	0.0589	0.0513	0.0455
	4	9,996,000	0.0638	0.0550	0.0484
	5	82,800,000	0.0695	0.0593	0.0516
	6	5,340,000	0.0760	0.0640	0.0552
	7	9,662,000	0.0816	0.0681	0.0583
	8	40,970,000	0.0932	0.0761	0.0640
	9	77,950,000	0.1046	0.0842	0.0697
	10	5,835,000	0.1161	0.0933	0.0762
Array 2	1	2,963,000	0.0502	0.0446	0.0401
	2	14,650,000	0.0543	0.0478	0.0427
	3	5,535,000	0.0593	0.0517	0.0458
	4	2,787,000	0.0635	0.0548	0.0482
	5	9,692,000	0.0694	0.0592	0.0515
	6	1,623,000	0.0759	0.0639	0.0551
	7	1,900,000	0.0810	0.0677	0.0580
	8	2,866,000	0.0895	0.0739	0.0626
	9	1,238,000	0.0990	0.0812	0.0679
	10	55,250,000	0.1152	0.0929	0.0759
Array 3	1	21,130,000	0.0551	0.0489	0.0440
	2	9,844,000	0.0604	0.0532	0.0476
	3	7,251,000	0.0626	0.0548	0.0487
	4	7,515,000	0.0653	0.0564	0.0497
	5	7,943,000	0.0699	0.0597	0.0520
	6	7,859,000	0.0755	0.0641	0.0555
	7	9,352,000	0.0856	0.0711	0.0607
	8	115,200,000	0.0914	0.0756	0.0640
	9	7,536,000	0.1000	0.0818	0.0682
	10	148,000,000	0.1147	0.0928	0.0760
By Array	Array 1	266,900,000	0.0906	0.0738	0.0620
	Array 2	98,510,000	0.1034	0.0833	0.0685
	Array 3	341,600,000	0.1041	0.0846	0.0700
Overall		707,000,000	0.0992	0.0804	0.0668

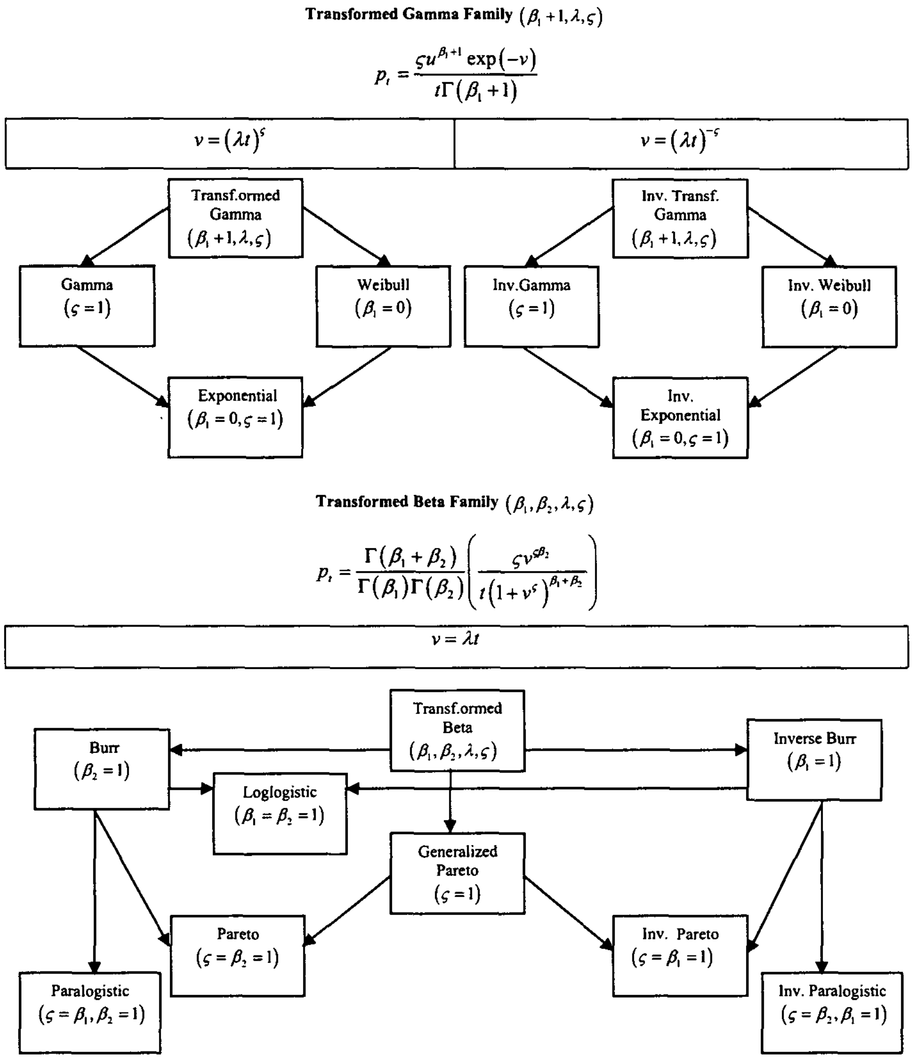
Table 4.6.1 Model 7.0(d): hazard rates for payment years 13,15 and 17.

5. Concluding Remarks

Reflective of the practical issues involved in the analysis of reserves, the related literature is extensive and explores a variety of theoretical frameworks. In general, having identified the salient data characteristics and gathered information on specific events that could have contributed to claims numbers and magnitudes, at the outset of every analysis a suitable analytical approach for the problem at hand has to be selected. Apart from any academic interest, it is likely that this search could have motivated some of the developments in reserving analysis, and will continue to do so. Hence, establishing the scope and limitations of each is important.

Through the generic model it is possible to give a functional interpretation to the claims data variability structure. As settlement speeds and scale variability increase, the assumptions and model structures encompassed by GLM models have to be replaced by more complex ones. The examples support remarks by Carroll (2003) with respect to the importance of the variance model. An inadequate variance model could lead to incorrect conclusions. The purpose of reserving analysis is not just to model historical claims data, but, more importantly, to predict *IBNR* and ultimate claim amounts. Both are strongly reliant on adequate variance definitions. Since claims records have to fulfil accounting requirements, corrections and adjustments to original entries are recorded as new transactions, and at unpredictable time lags. This could justify regarding measures of cumulative claims as repeated observations of an ongoing process. In this context, normal errors assumptions could be made tenable through suitable transformations or expectation functions, hence availing analytical approaches such as outlined in Lindstrom and Bates (1990). In the examples presented, and with the selected data, autoregressive error structures cannot be successfully used.

The generic model makes random effects models accessible to the problem of reserving. With the different variance model structures, it exponentially increases the analytical resources that can result from constructing families of reserving models around families of distributions. Graph 5.1 is an example of a template that can be used to identify the most suitable model structure for the data of interest and formulate the percentage cash flow function. With respect to the underlying assumptions for random effect parameters other alternatives are possible. Escobar and West (1992) propose a non-parametric approach, where the random parameter is taken from a rich class of distributions provided by the Dirichlet process. Lai and Shih (2003) leave the distribution of the random effects totally unspecified. The non-linear mixed effects models library (NLME) assumes that the random effects and the errors have Gaussian distributions. Using a matrix decomposition, Bates and Pinheiro (1998) shows that the random effects distribution expressed in terms of the relative precision factors can easily deliver the likelihood for the fixed and random effects. The flexibility of Gibbs sampling methods (Geman and Geman, 1984) has influenced the decision to implement the examples with BUGS (Spiegelhalter et al., 1995), as applications of Bayesian models and MCMC estimation methods. Nevertheless, other approaches in relation to analytical platforms, model structures and assumptions, beyond those explored, should be considered.



Graph 5.1 Examples of families of models for the percentage cash flow function.

ACKNOWLEGEMENTS

The author thanks Professor R. Verrall (Cass Business School) for his helpful remarks.

REFERENCES

Alba, de, E., (2002). Bayesian Estimation of Outstanding Claims Reserves. North American Actuarial Journal, Vol. 6, No. 4, 1-20.

Aziz, N. S., Gambertoglio, J. G., Lin, E. T., Grausz, H. and Benet, L. Z., (1978). Pharmacokinetics of Cefamandole using HPLC assay. Pharmacokinetics and Biopharmaceutics, 6, 153-164

Bates, D. M. and Pinheiro, J. C., (1998). Computational methods for multilevel models, Technical Memorandum BL0112140-980226-01TM, Bell Labs, Lucent Technologies, Murray Hill, NJ.

Beal, S. and Sheiner, L. B., (1992). NONMEM User's guides. NONMEM Project Group, University of California, San Francisco.

Bornheutter, R. L. and Ferguson, R. E., (1972). The actuary and IBNR. Proceedings of the Casualty Actuarial Society, LIX, 181-195.

Carroll, R. J., (2003). Variances are not always nuisance parameters. Biometrics, 59, 211-220.

Davinian M. and Giltinan, D. M., (1993). Some general estimation methods for non-linear mixed effects models. Journal of Biopharmaceutical Statistics, 3, 23-55.

Davinian M. and Giltinan, D. M., (1996). Nonlinear Models for Repeated Measurement Data. Chapman & Hall, London.

Elashoff, J. D., Reedy, T. J. and Meyer, J. H., (1982). Analysis of gastric emptying data. Gastroenterology 83, 1306-1312.

Escobar, M. and West, M., (1992). Computing Bayesian nonparametric hierarchical models. ISDS Discussion Paper 92-A20, Duke University.

Geman, S. and Geman, D., (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Tans. Pattn. Anal. Mach. Intell., 6(6), 721-741.

Kahane, Y., (1987). Ratemaking and reserving under decelerating inflation. (Unknown source).

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

Lai, T. L. and Shih, M., (2003). Nonparametric estimation in nonlinear mixed effects models. Biometrika, 90, 1, 1-13.

Lindstrom, M. J. and Bates, D. M., (1990). Nonlinear mixed effects models for repeated measures data. Biometrics 46, 673-687.

Mack, T., (1991). A Simple Parametric Model for Rating Automobile Insurance or Estimating IBNR Claims Reserves. ASTIN Bulletin 22, 93-109.

McCullagh, P. and Nelder, J., (1989). Generalised linear models, 2nd edition. Chapman and Hall.

Pinheiro, J. C., Bates, D.M., and Lindstrom, M. J., (1994). Model Building for Nonlinear Mixed Effects Models Technical Report 931, University of Wisconsin, Madison, Dept. of Statistics.

Pinheiro, J. C. and Bates, D. M., (2000). Mixed-Effects Models in S and S-Plus. Springer-Verlag New York, Inc.

Ratkowski, D. A., (1990). Handbook of Nonlinear Regression Models. Dekker, New York.

Renshaw, A. E., (1989). Chain-ladder and interactive modelling (claims reserving and GLIM). J.I.A. 116, 559-587.

Renshaw, A. E. and Verrall, R. J., (1998). A stochastic model underlying the chain-ladder technique. B.A.J. 4, 903-923.

Shapiro, S. S. and Wilk, M. B., (1965). An analysis of variance test for normality (complete samples), Biometrika, 52, 3 and 4, 591-611.

Spiegelhalter, D. J., Thomas, A., Best, N. G. and Gilks, W. R., (1995). BUGS 0.5: Bayesian inference using Gibbs sampling, MRC Biostatistics Unit, Cambridge, U.K.

Taylor, G. C., (2000). Loss reserving: an actuarial perspective. Kluwer Academic Publishers.

Verrall, R. J., (1991). On the Unbiased Estimation of Reserves from Loglinear Models. Insurance Mathematics and Economics, Vol. 10, No. 1, 75-80.

Vera, G. (2003). A generic claims reserving model. A fundamental risk analysis. Casualty Actuarial Society forum, 361-376.

Wakefield, J. C. and Walker, S., (1994). Population models with a non-parametric random coefficient distribution. Technical report, Department of Mathematics, Imperial College.

Wakefield, J. C., (1996). The Bayesian analysis of population pharmacokinetic models. Journal of the American Statistical Association, 91, 62-75.

Wright, T. S., (1990). A stochastic method for claims reserving in general insurance. J.I.A. 117, 677-731.

Zeger, S., L., Liang, K., Y. and Albert P., S., (1988). Models for Longitudinal Data: A generalized estimating equation approach, Biometrics, 44, 1049-1060.

APPENDIX A

A.1 INCREMENTAL PAID CLAIMS DATA

Und. Year	Development Period									
	1	2	3	4	5	6	7	8	9	10
1	0	94,984	1,049,297	625,878	541,108	427,352	476,477	354,258	188,400	144,987
2	0	147,751	999,224	937,426	811,294	436,866	264,148	143,616	102,416	132,920
3	0	45,751	442,168	588,627	390,301	231,257	119,690	64,365	73,641	93,371
4	0	20,252	340,320	596,633	336,142	183,473	90,574	114,241	99,467	51,950
5	0	21,655	787,440	992,505	893,315	772,514	795,088	718,526	504,213	321,630
6	0	221,177	1,212,010	1,867,718	1,372,904	1,254,084	1,003,612	696,973	534,547	409,845
7	0	192,144	749,425	1,174,401	1,500,585	2,079,434	1,675,154	1,972,712	1,372,848	491,984

(cont.)

Und. Year	Development Period								
	11	12	13	14	15	16	17	18	19
1	124,614	111,642	56,210	64,259	33,893	15,440	8,255	22,300	25,173
2	109,996	58,163	53,679	54,255	25,631	51,443	56,702	59,857	
3	53,678	29,044	12,259	10,267	11,264	9,515	8,859		
4	45,692	21,824	36,117	54,185	52,194	47,355			
5	183,470	85,610	73,300	97,350	42,620				
6	111,090	529,552	403,242	291,414					
7	212,273	191,729	28,340						

Table A.1 Simulated data based on the claims experience of a mixed portfolio.

A.2 CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Period									
	1	2	3	4	5	6	7	8	9	10
1	0	94,984	1,144,281	1,770,159	2,311,267	2,738,619	3,215,096	3,569,354	3,757,754	3,902,741
2	0	147,751	1,146,975	2,084,401	2,895,695	3,332,561	3,596,709	3,740,325	3,842,741	3,975,661
3	0	45,751	487,919	1,076,546	1,466,847	1,698,104	1,817,794	1,882,159	1,955,800	2,049,171
4	0	20,252	360,572	957,205	1,293,347	1,476,820	1,567,394	1,681,635	1,781,102	1,833,052
5	0	21,655	809,095	1,801,600	2,694,915	3,467,429	4,262,517	4,981,043	5,485,256	5,806,886
6	0	221,177	1,433,187	3,300,905	4,673,809	5,927,893	6,931,505	7,628,478	8,163,025	8,572,870
7	0	192,144	941,569	2,115,970	3,616,555	5,695,989	7,371,143	9,343,855	10,716,703	11,208,687

(cont.)

Und. Year	Development Period								
	11	12	13	14	15	16	17	18	19
1	4,027,355	4,138,997	4,195,207	4,259,466	4,293,359	4,308,799	4,317,054	4,339,354	4,364,527
2	4,085,657	4,143,820	4,197,499	4,251,754	4,277,385	4,328,828	4,385,530	4,445,387	
3	2,102,849	2,131,893	2,144,152	2,154,419	2,165,683	2,175,198	2,184,057		
4	1,878,744	1,900,568	1,936,685	1,990,870	2,043,064	2,090,419			
5	5,990,356	6,075,966	6,149,266	6,246,616	6,289,236				
6	8,683,960	9,213,512	9,616,754	9,908,168					
7	11,420,960	11,612,689	11,641,029						

Table A.2 Cumulative data based on table A.1.

A.3 PRELIMINARY MODEL 1.0

A.3.1 MODEL 1.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	55,360	348,100	920,800	1,640,000	2,322,000	2,871,000	3,278,000	3,570,000	3,778,000	3,928,000
2	52,840	402,300	1,143,000	2,030,000	2,772,000	3,289,000	3,624,000	3,840,000	3,980,000	4,075,000
3	21,410	178,700	547,000	1,017,000	1,413,000	1,681,000	1,848,000	1,953,000	2,019,000	2,063,000
4	41,190	183,000	474,200	854,300	1,198,000	1,448,000	1,617,000	1,730,000	1,808,000	1,863,000
5	27,850	236,600	774,400	1,641,000	2,657,000	3,610,000	4,386,000	4,970,000	5,393,000	5,698,000
6	82,280	563,200	1,597,000	3,037,000	4,550,000	5,878,000	6,928,000	7,713,000	8,290,000	8,711,000
7	10,640	152,100	695,400	1,911,000	3,769,000	5,864,000	7,742,000	9,191,000	10,220,000	10,930,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,037,000	4,118,000	4,179,000	4,226,000	4,262,000	4,290,000	4,313,000	4,332,000	4,347,000
2	4,140,000	4,186,000	4,219,000	4,243,000	4,262,000	4,276,000	4,287,000	4,296,000	
3	2,093,000	2,113,000	2,128,000	2,139,000	2,147,000	2,153,000	2,157,000		
4	1,903,000	1,934,000	1,958,000	1,977,000	1,993,000	2,006,000			
5	5,917,000	6,077,000	6,195,000	6,284,000	6,351,000				
6	9,022,000	9,254,000	9,428,000	9,562,000					
7	11,410,000	11,740,000	11,970,000						

Table A.3 Fitted claims computed by Monte Carlo simulations estimated over 5000 independent samples.

A.3.2 MODEL 1.0 FIXED EFFECTS PARAMETER ESTIMATES, VARIANCE AND DEVIANCE

Und. Year	Fixed effect parameters						
	L	D	Kc	Kd	ϑ	K_{s_1}	K_{s_2}
	15.7400	4.8810	1.5470	-15.3600	-6.2470	-5.0680	-2.8080
	Underwriting year random effect parameters						
	l_w	d_w	kc_w	kd_w			
1	-0.4364	-0.4050	-0.5091	-0.1369			
2	-0.4592	-0.3512	-0.3995	-0.0313			
3	-1.1480	-0.0200	-0.3330	0.0317			
4	-1.1690	-0.1730	-0.4673	-0.4279			
5	-0.0342	0.6846	-0.3858	-0.0858			
6	0.3919	0.0078	-0.4838	-0.0484			
7	0.6101	2.3000	-0.1819	0.1548			
σ^2	1.85E+10						
Deviance	2,980						

Table A.4 Model 1.0 parameters and diagnostics.

A.4 PRELIMINARY MODEL 2.0

A.4.1 MODEL 2.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	65,520	381,400	962,800	1,665,000	2,323,000	2,855,000	3,255,000	3,547,000	3,759,000	3,914,000
2	47,310	385,000	1,127,000	2,026,000	2,778,000	3,299,000	3,634,000	3,847,000	3,985,000	4,077,000
3	15,680	155,300	519,400	1,006,000	1,420,000	1,696,000	1,864,000	1,965,000	2,028,000	2,068,000
4	17,190	143,400	444,600	847,300	1,211,000	1,473,000	1,645,000	1,756,000	1,828,000	1,876,000
5	28,240	238,800	780,300	1,649,000	2,662,000	3,612,000	4,385,000	4,967,000	5,390,000	5,695,000
6	81,750	558,800	1,592,000	3,036,000	4,555,000	5,885,000	6,932,000	7,715,000	8,288,000	8,706,000
7	14,370	176,900	753,200	1,981,000	3,810,000	5,856,000	7,699,000	9,140,000	10,180,000	10,910,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,029,000	4,115,000	4,181,000	4,232,000	4,272,000	4,303,000	4,329,000	4,350,000	4,367,000
2	4,139,000	4,183,000	4,214,000	4,237,000	4,255,000	4,268,000	4,278,000	4,286,000	
3	2,094,000	2,112,000	2,125,000	2,134,000	2,140,000	2,145,000	2,149,000		
4	1,908,000	1,931,000	1,948,000	1,960,000	1,969,000	1,976,000			
5	5,915,000	6,075,000	6,194,000	6,282,000	6,350,000				
6	9,014,000	9,244,000	9,417,000	9,550,000					
7	11,420,000	11,770,000	12,020,000						

Table A.5 Fitted claims computed by Monte Carlo simulations estimated over 7000 independent samples.

A.4.2 MODEL 2.0 DIAGNOSTICS AND PARAMETER ESTIMATES

Und. Year	Fixed effect parameters						
	L	D	Kc	Kd	ϑ	K_{κ_1}	K_{κ_2}
	14.8900	5.0140	0.2660	-13.4400	-3.4710	-5.5230	-2.5030
	Underwriting year random effect parameters						
	l_w	d_w	kc_w	kd_w	ϑ_w		
1	0.4214	-0.7378	0.7250	-0.0208	-0.5061		
2	0.3865	-0.4262	0.8981	0.0177	-0.1910		
3	-0.3060	0.0697	0.9973	0.0297	-1.4670		
4	-0.3818	-0.0873	0.9274	-0.0648	-0.8611		
5	0.8153	0.5157	0.8900	-0.0115	-0.0693		
6	1.2400	-0.1151	0.8002	-0.1619	1.4680		
7	1.4700	1.9440	1.0560	0.3943	1.6780		
σ^2	5.87E+09						
Deviance	2,930						

Table A.6 Model 2.0 parameters and diagnostics.

A.5 VALIDATION MODEL 3.0

A.5.1 MODEL 3.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	65,890	381,500	961,700	1,663,000	2,321,000	2,854,000	3,255,000	3,547,000	3,760,000	3,915,000
2	45,370	377,000	1,117,000	2,023,000	2,782,000	3,305,000	3,640,000	3,852,000	3,989,000	4,079,000
3	14,930	152,700	516,400	1,005,000	1,421,000	1,698,000	1,866,000	1,967,000	2,029,000	2,068,000
4	16,700	142,400	443,800	846,500	1,210,000	1,473,000	1,646,000	1,757,000	1,828,000	1,876,000
5	29,470	241,700	782,300	1,649,000	2,664,000	3,614,000	4,386,000	4,967,000	5,389,000	5,692,000
6	97,230	611,800	1,662,000	3,081,000	4,555,000	5,852,000	6,887,000	7,673,000	8,259,000	8,695,000
7	10,310	147,300	679,700	1,884,000	3,745,000	5,856,000	7,748,000	9,202,000	10,230,000	10,940,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,030,000	4,116,000	4,182,000	4,233,000	4,273,000	4,304,000	4,330,000	4,351,000	4,368,000
2	4,140,000	4,183,000	4,214,000	4,236,000	4,253,000	4,266,000	4,276,000	4,283,000	
3	2,094,000	2,112,000	2,124,000	2,132,000	2,139,000	2,144,000	2,147,000		
4	1,909,000	1,931,000	1,948,000	1,960,000	1,968,000	1,975,000			
5	5,911,000	6,071,000	6,190,000	6,279,000	6,347,000				
6	9,020,000	9,266,000	9,454,000	9,600,000					
7	11,410,000	11,740,000	11,970,000						

Table A.7 Fitted claims computed by Monte Carlo simulations estimated over 25,500 independent samples.

A.5.2 MODEL 3.0 DIAGNOSTICS AND PARAMETER ESTIMATES

Und. Year	Fixed effect parameters						
	L	D	Kc	Kd	ϑ	K_{s_1}	K_{s_2}
	14.2900	5.6670	0.3595	-13.6300	-6.8610	-5.3350	-2.4080
	Underwriting year random effect parameters						
	l_w	d_w	kc_w	kd_w			
1	1.0300	-1.3860	0.6305	0.1420			
2	0.9940	-1.0300	0.8152	0.0310			
3	0.3015	-0.5548	0.9109	0.0018			
4	0.2259	-0.7370	0.8359	-0.0281			
5	1.4240	-0.1191	0.7978	-0.0700			
6	1.8620	-0.9282	0.6661	-0.0521			
7	2.0660	1.5820	1.0150	0.2053			
Subset 1 - σ_1^2	9.53E+09						
Subset 2 - σ_2^2	4.21E+09						
Subset 3 - σ_3^2	4.05E+10						
Deviance	2,936						

Table A.8 Model 3.0 parameters and diagnostics.

A.6 VALIDATION MODEL 4.0

A.6.1 MODEL 4.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	68,260	386,900	967,500	1,667,000	2,323,000	2,854,000	3,253,000	3,545,000	3,757,000	3,912,000
2	47,650	385,200	1,126,000	2,026,000	2,779,000	3,300,000	3,635,000	3,848,000	3,986,000	4,077,000
3	14,160	149,000	512,300	1,005,000	1,424,000	1,702,000	1,869,000	1,969,000	2,030,000	2,069,000
4	15,730	138,100	438,100	844,200	1,212,000	1,477,000	1,649,000	1,759,000	1,830,000	1,877,000
5	28,640	236,300	771,100	1,637,000	2,656,000	3,613,000	4,389,000	4,971,000	5,392,000	5,695,000
6	80,580	552,900	1,581,000	3,023,000	4,545,000	5,881,000	6,933,000	7,719,000	8,293,000	8,713,000
7	10,650	149,800	685,900	1,892,000	3,749,000	5,853,000	7,740,000	9,195,000	10,230,000	10,930,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,028,000	4,114,000	4,181,000	4,232,000	4,272,000	4,305,000	4,331,000	4,352,000	4,369,000
2	4,139,000	4,183,000	4,214,000	4,237,000	4,254,000	4,268,000	4,278,000	4,286,000	
3	2,094,000	2,111,000	2,122,000	2,131,000	2,137,000	2,141,000	2,145,000		
4	1,909,000	1,931,000	1,946,000	1,958,000	1,966,000	1,973,000			
5	5,912,000	6,071,000	6,188,000	6,276,000	6,343,000				
6	9,021,000	9,251,000	9,424,000	9,556,000					
7	11,420,000	11,750,000	11,980,000						

Table A.9 Fitted claims computed by Monte Carlo simulations estimated over 25,500 independent samples.

A.6.2 MODEL 4.0 DIAGNOSTICS AND PARAMETER ESTIMATES

Und. Year	Fixed effect parameters					
	L	D	Kc	Kd	K_{s_1}	K_{s_2}
	15.2900	5.6850	0.6367	-14.4900	-4.4210	-2.4880
	Underwriting year random effect parameters					
	l_w	d_w	kc_w	kd_w		
1	0.0272	-1.3800	0.3547	0.0661		
2	-0.0102	-0.9854	0.5449	-0.0122		
3	-0.7086	-0.2573	0.6872	-0.0329		
4	-0.7808	-0.5506	0.5905	0.0138		
5	0.4234	-0.1890	0.5087	0.0121		
6	0.8493	-0.8225	0.4171	0.0846		
7	1.0600	1.6470	0.7479	0.0277		
σ^2	1.06E+10					
ρ	0.0025					
Deviance	2,902					

Table A.10 Model 4.0 parameters and diagnostics.

A.7 FINAL MODEL 5.0

A.7.1 MODEL 5.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	65,470	380,000	959,800	1,662,000	2,322,000	2,855,000	3,256,000	3,549,000	3,761,000	3,915,000
2	45,730	377,900	1,118,000	2,023,000	2,782,000	3,305,000	3,640,000	3,853,000	3,989,000	4,079,000
3	14,630	150,600	513,300	1,004,000	1,423,000	1,700,000	1,868,000	1,969,000	2,030,000	2,069,000
4	16,030	138,600	438,300	844,500	1,212,000	1,477,000	1,649,000	1,759,000	1,830,000	1,877,000
5	28,820	240,200	780,900	1,648,000	2,661,000	3,611,000	4,385,000	4,968,000	5,391,000	5,696,000
6	84,690	567,600	1,601,000	3,038,000	4,549,000	5,875,000	6,924,000	7,710,000	8,288,000	8,711,000
7	13,570	171,000	738,800	1,964,000	3,803,000	5,864,000	7,714,000	9,153,000	10,190,000	10,910,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,030,000	4,115,000	4,181,000	4,231,000	4,271,000	4,303,000	4,328,000	4,349,000	4,365,000
2	4,141,000	4,183,000	4,214,000	4,236,000	4,253,000	4,266,000	4,276,000	4,283,000	
3	2,094,000	2,112,000	2,124,000	2,132,000	2,138,000	2,143,000	2,146,000		
4	1,908,000	1,930,000	1,946,000	1,958,000	1,966,000	1,973,000			
5	5,916,000	6,076,000	6,195,000	6,284,000	6,351,000				
6	9,023,000	9,256,000	9,433,000	9,568,000					
7	11,410,000	11,750,000	12,000,000						

Table A.11 Fitted claims computed by Monte Carlo simulations estimated over 25,500 independent samples.

A.7.2 MODEL 5.0 DIAGNOSTICS AND PARAMETER ESTIMATES

Und. Year	Fixed effect parameters							
	L	D	Kc	Kd	ϑ_1^*	ϑ_2^*	K_{ϑ_1}	K_{ϑ_2}
	14.8300	5.2000	1.8770	-13.5100	-4.1220	1.6200	-4.4750	-2.2010
	Underwriting year random effect parameters							
	l_w	d_w	kc_w	kd_w				
1	0.4881	-0.8916	-0.8796	0.0577				
2	0.4528	-0.5451	-0.6994	-0.0371				
3	-0.2402	-0.0258	-0.5958	0.0205				
4	-0.3172	-0.1816	-0.6650	0.0317				
5	0.8829	0.3530	-0.7180	0.0166				
6	1.3100	-0.3071	-0.8147	-0.0499				
7	1.5340	1.8580	-0.5337	0.1329				
σ^2	5.05E+09							
Deviance	2,928							

Table A.12 Model 5.0 parameters and diagnostics.

A.8 FINAL MODEL 6.0

A.8.1 MODEL 6.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	66,200	383,900	964,400	1,663,000	2,317,000	2,848,000	3,249,000	3,542,000	3,755,000	3,912,000
2	135,300	641,200	1,389,000	2,138,000	2,748,000	3,197,000	3,517,000	3,745,000	3,908,000	4,027,000
3	35,860	237,300	618,100	1,046,000	1,401,000	1,652,000	1,820,000	1,931,000	2,006,000	2,057,000
4	59,230	265,600	575,400	903,400	1,188,000	1,410,000	1,575,000	1,697,000	1,787,000	1,854,000
5	24,320	220,100	747,500	1,618,000	2,651,000	3,620,000	4,404,000	4,988,000	5,407,000	5,704,000
6	165,500	810,000	1,909,000	3,252,000	4,598,000	5,790,000	6,773,000	7,554,000	8,166,000	8,642,000
7	8,110	125,600	616,800	1,796,000	3,698,000	5,889,000	7,832,000	9,285,000	10,280,000	10,940,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,028,000	4,116,000	4,182,000	4,234,000	4,275,000	4,307,000	4,334,000	4,355,000	4,372,000
2	4,116,000	4,183,000	4,235,000	4,275,000	4,308,000	4,334,000	4,355,000	4,372,000	
3	2,093,000	2,119,000	2,139,000	2,153,000	2,164,000	2,173,000	2,180,000		
4	1,905,000	1,945,000	1,975,000	2,000,000	2,019,000	2,035,000			
5	5,917,000	6,070,000	6,182,000	6,266,000	6,329,000				
6	9,015,000	9,309,000	9,543,000	9,731,000					
7	11,380,000	11,670,000	11,870,000						

Table A.13 Fitted claims computed by Monte Carlo simulations estimated over 25,500 independent samples.

A.8.2 MODEL 6.0 DIAGNOSTICS AND PARAMETER ESTIMATES

Und. Year	Fixed effect parameters									
	L	D	Kc	Kd	ϑ_1^*	ϑ_2^*	ϑ_3^*	ϑ_4^*	K_{ε_1}	K_{ε_2}
	15.8000	4.5610	1.2320	-12.9600	-1.9120	1.1780	15.6300	-0.1581	-6.3830	-1.3110
	Underwriting year random effect parameters									
	l_w	d_w	kc_w	kd_w						
1	-0.4813	-0.3208	-0.2500	-0.0844						
2	-0.4805	-1.0340	-0.3300	0.0070						
3	-1.1870	-0.3798	-0.1589	0.0487						
4	-1.2220	-0.9034	-0.3593	-0.0150						
5	-0.0970	1.0910	-0.0505	-0.0292						
6	0.3991	-0.3295	-0.3438	0.0398						
7	0.5364	2.9920	0.1889	0.0295						
σ^2	109.600									
Deviance	2.887									

Table A.14 Model 6.0 parameters and diagnostics.

B.1 ARRAYS 1 TO 3: CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Period									
	1	2	3	4	5	6	7	8	9	10
1	1,965,120	4,455,720	5,125,260	6,208,080	6,365,400	7,566,780	8,134,380	8,300,640	8,491,200	9,072,840
2	508.829	7,957,659	10.395,008	11,627,118	12,659,049	13,512,509	13,813,936	14,609,422	14,836,855	15,095,843
3	1,070,272	2,117,478	2,876,979	3,141,005	4,127,612	4,337,374	4,503,876	4,522,524	4,543,644	4,560,720
4	983,295	2,957,869	5,140,518	6.369,315	6,326,691	7.867,792	7,356,575	7,656,758	7,817.554	7,844,431
5	9,979,594	26,286,414	25,263,483	40,239,973	51,246,513	54,472,139	55,800,837	56,658.302	59,561,780	
6	55,668	1,586,037	1,764,809	2,888,328	3,158,562	3,445,626	3,459,794	3,604,995		
7	2,128,880	4,827,030	5,552,365	6,725,420	6,895,850	8,197,345	8,812,245			
8	2,528,789	8,400,695	12.219,988	21,139,396	23,109,446	22,292,555				
9	1,613,864	10,075.000	12,091,140	28,449,598	34,707,350					
10	110.580	576.687	1,887,649	2.244.074						

(cont.)

Und. Year	Development Period								
	11	12	13	14	15	16	17	18	19
1	9,298,740	9,640,380	9.681,600						
2	15,216,319	15,361,081							
3	4,572,331								
4									
5									
6									
7									
8									
9									
10									

Table B.1 Array 1: Simulated data based on the claims experience of a marine hull portfolio.

Und. Year	Development Period									
	1	2	3	4	5	6	7	8	9	10
1	445,841	1,654,609	1,605,300	2,004.723	2,299,800	2,275,241	2,470,159	2,579,168	2,641,868	2,744,127
2	2,426,373	7,352,166	8,950,532	10,167,845	11,756,358	12,137,425	12,030,761	12,970,026	13,607,600	14,532,427
3	184,480	318,830	1.296,062	2,733,415	2,650,811	3,010,199	3,168,834	3,349,023	3,431,900	3,493,316
4	601,693	1,084,468	1,510,596	1,606,829	1,910,257	1,973,043	2,274,886	2,320,886	2.304,771	2,407,211
5	968,366	2,530,871	4,608,428	4,912,525	6,271,612	6,799,211	5,466,992	6,770,634	7,779,669	
6	239,105	326,614	670,735	905,967	913,090	1,131,129	1,090,519	1,100,114		
7	361.246	1,078.435	1.205,746	1,478.083	1,461,000	1,547,928	1,587,804			
8	742.335	1.392,375	1,937,655	1,865,055	2,282,175	2,388,270				
9	228,341	704,498	798,463	815,165	1,013,800					
10	1.589.527	15.936.500	22.331.091	28.741.729						

(cont.)

Und. Year	Development Period									
	11	12	13	14	15	16	17	18	19	
1	2,786,141	2,870,495	2,871,777							
2	12,130,088	12,131,811								
3	3,542,361									
4										
5										
6										
7										
8										
9										
10										

Table B.2 Array 2: Simulated data based on the claims experience of a marine cargo portfolio.

Und. Year	Development Period									
	1	2	3	4	5	6	7	8	9	10
1	3,232,205	9,881,808	12,905,347	14,832,451	15,314,642	16,405,052	17,591,239	17,993,791	18,352,169	18,500,158
2	1,262,978	2,979,101	3,119,301	4,617,621	5,884,276	6,241,315	6,659,795	6,818,432	7,164,468	7,544,540
3	1,099,101	3,367,582	4,078,680	4,335,973	5,855,806	5,875,321	5,977,392	6,129,768	6,185,657	6,205,646
4	731,599	2,554,623	3,586,046	4,168,936	3,762,376	5,081,203	4,686,750	5,777,092	6,108,501	5,983,210
5	175,251	1,581,689	2,116,488	3,455,030	4,284,402	4,794,982	4,848,263	5,275,530	5,010,701	
6	1,339,210	3,824,400	4,704,740	5,565,181	5,412,190	6,389,658	6,517,524	7,284,369		
7	590,921	1,263,060	1,812,106	3,441,064	4,204,651	5,155,490	5,213,434			
8	20,698,911	43,203,529	63,631,429	70,713,997	89,249,817	85,287,218				
9	790,164	2,944,098	4,609,088	4,747,316	5,511,068					
10	9,900,060	17,916,636	50,167,809	62,132,660						

(cont.)

Und. Year	Development Period									
	11	12	13	14	15	16	17	18	19	20
1	18,733,437	19,496,439	19,567,234							
2	7,689,022	7,771,566								
3	6,240,897									
4										
5										
6										
7										
8										
9										
10										

Table B.3 Array 3: Simulated data based on the claims experience of an aviation hull portfolio.