DEPARTMENT OF ACTUARIAL STUDIES RESEARCH PAPER SERIES

Forecasting General Insurance Liabilities

by

Piet de Jong

piet.dejong@mq.edu.au

Research Paper No. 2004/03
February 2004

Division of Economic and Financial Studies
Macquarie University
Sydney NSW 2109 Australia

.

Abstract The forecasting of general insurance liabilities using runoff triangle data is discussed and considered. Techniques are based on time series models and methods which facilitate the calculation of forecast distributions. Models are readily estimated and permit the consideration of correlation within and between triangles. These correlations are critical to proper reserving. Models are assessed using diagnostics. Examples illustrate procedures. Calculations are implemented in Excel linked to specialized algorithms.

1 Introduction

Table 1 displays a typical runoff triangle. Runoff triangles form the basis for predicting outstanding liabilities. Outstanding liabilities correspond to the lower unfilled portion of the rectangle. This article deals with models and methods for performing the prediction. For example using the data in Table 1 and the methods of this paper leads to the forecast liability distribution displayed in Figure 1. The mean and standard deviation of the displayed distribution differ appreciably from estimates derived with a method currently regarded as sound actuarial practice.

Table 1: the AFG data – cumulative incurred claim amounts

accident	development year j									
year i	0	1	2	3	4	5	6	7	8	9
1	5012	8269	10907	11805	13539	16181	18009	18608	18662	18834
2	106	4285	5396	10666	13782	15599	15496	16169	16704	
3	3410	8992	13873	16141	18735	22214	22863	23466		
4	5655	11555	15766	21266	23425	26083	27067			
5	1092	9565	15836	22169	25955	26180				
6	1513	6445	11702	12935	15852					
7	557	4020	10946	12314						
8	1351	6947	13112							
9	3133	5395								
10	2063									

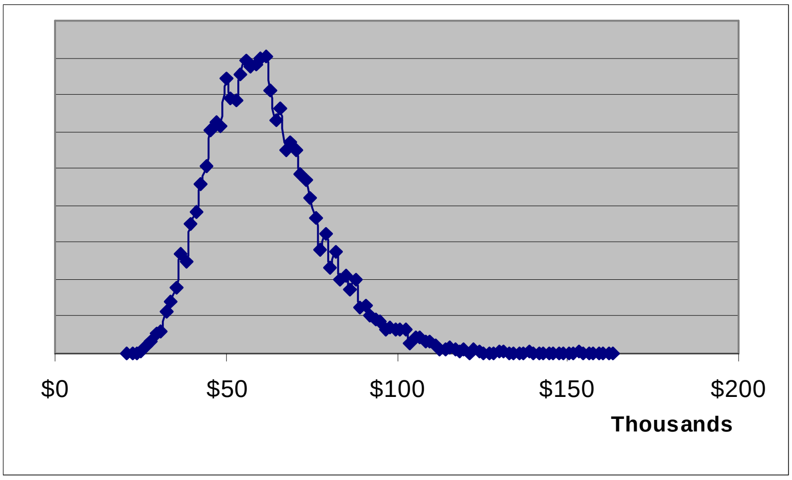

Figure 1: Estimated histogram of forecast incurred liabilities for the AFG data using DC model

Entries in the body of Table 1 are denoted c_{ij} and are cumulative paid out amounts with respect to accident year i up to and including development year j. Entries in each row generally increase with j indicating that as time progresses, incurred liabilities with respect to an accident year increase. Each calendar year leads to an oblique diagonal of observations. In Table 1 there are $n = 10$ calendar years of observations. The data relate to Automatic Facultative General Liability (excluding Asbestos and Environmental) from the Historical Loss Development study. These data, hereafter called the AFG data, were considered by both Mack (1994) and England and Verrall (2002) and are used in this paper to illustrate methods, facilitating comparison to the earlier papers.

Of interest is the conditional distribution of each accident year's liability through to development year $n-1$:

$$c_{i,n-1} - c_{i,n-i} = c_{i,n-i}\left(e^{g_i} - 1\right) \qquad \Rightarrow \qquad g_i \equiv \ln\left(\frac{c_{i,n-1}}{c_{i,n-i}}\right) , \qquad i = 2, \ldots, n . \tag{1}$$

Thus g_i is the continuously compounded percentage growth in future cumulative claims with respect to accident year i, through to development year $n-1$ while $e^{g_i} - 1$ is the actual percentage increase. The g_i are unknown and must be predicted. Total liabilities through to development year $n-1$ are

$$\sum_{i=2}^{n}(c_{i,n-1} - c_{i,n-i}) = \sum_{i=2}^{n} c_{i,n-i}\left(e^{g_i} - 1\right) , \tag{2}$$

The approach of this paper to forecasting (2) is to analytically construct the means, standard deviations and correlations of the joint distribution of the future growth rates and the g_i given the observed runoff triangle c_{ij}. Simulation from the conditional distributions yields a detailed view of projected future liabilities including distributions of sums across development, accident or calendar years.

Forecast distributions are constructed using time series methods and techniques, tailored to the runoff triangle situation. There are many advantages to such an approach. First, there is a rich and widely studied range of available models. Second, there is no need to rethink or reinvent optimal forecasting technology. The available technology is robust, exhaustively studied as to properties, issues, limitations, and efficient coding. Third, fit diagnostics are readily available. Fourth, by employing the conventional framework the actuary is free to tackle actuarially meaningful tasks including uncovering and dealing with features such as correlations between accident and calendar years or between different triangles. Thus the actuary need not spend time on estimation, forecasting and diagnostic design issues which have already been resolved. The situation is akin to using a car: the driver focusses on where to go and how to negotiate obstacles without bothering with the car's engineering details. The framework imposes broad constraints (keep to roads) but within the constraints, the setup is known to be "best." A (car) model is chosen according to the particular job, and having a broad yet well placed appreciation for its driving or actuarial capabilities and limitations.

There is an extensive literature on claims reserving: see for example the bibliography in the recent paper to the Institute of Actuaries by England and Verrall (2002) or the book by Taylor (2000). Papers of particularly of interest include Zehnwirth (1985), Wright (1990), Verrall (1990) , Goovaerts and Redant (1999) and Barnett and Zehnwirth (2000), and Collins and White (2001). Further references include: Sanders (1990), Verrall (1991), Verrall (1996) and De Jong and Zehnwirth (1983).

The further layout of this paper is as follows. The next section compares our approach, using the AGF data of Table 1, to other current methods. Section 3 sets out the "basic model" which is the starting point of our approach and the basis for subsequent generalizations. An example analysis, using the AFG data and basic model, is given in the subsections of 3 which discuss estimation, calculation of the forecast distribution, basic model assessment and assessment of correlations. Section 4 discusses extensions to the basic model to incorporate development, accident or calendar year correlation. Examples illustrate methods and calculation procedures. Further forecasting issues are discussed in §5 including forecasting beyond the latest development year in §5.2 and the robustifying of forecasts is discussed in §5.3. Section 6 gives a detailed comparison between the basic model and the methods based on the "chain ladder" method. Correlation between different runoff triangles is dealt with in §7. Appendices deal implementation and technical issues as well as a detailed critique of the model proposed by Mack (1993) to justify the chain ladder method.

2 Example forecast using the AFG data

Figure 1 displays the estimated conditional distribution of (2) using the AFG data of Table 1 and "Development Correlation" (DC) model described in §4. The mean and coefficient of variation of (2) are estimated to be around \$63 000 and 26%, respectively. Existing techniques focus on the first two moments and it is useful to compare them to our moments. Mack (1994) gives estimates of \$52 135 and 52%. A breakdown of our estimates and the Mack estimates according to accident year is given in Table 2. Also displayed are estimates based on the over dispersed Poisson model with a Hoerl curve (England and Verrall 2002).

Table 2 emphasizes the material differences between estimates derived from differing methods. Some insight into the relative merits of the methods is gained by examining the estimates. The Mack method gives an estimate of the standard deviation of the prediction error for accident year 2 of 206 while the Poisson model gives 486. The data directly relevant to this estimate is the 0.92% or 172 growth in liabilities in accident year 1 between development years 8 and 9. Using the Mack estimates and the arguably conservative normal approximation lead us to expect liabilities in accident year 2 to grow in excess of $154 + 1.65 \times 206 = 500$ with a probability of around 5% while the Poisson approach would give an even more surprising 5% limit of 1045. These conclusions seem at odds with the data. Furthermore consider the standard deviation of total liabilities under the Mack method. The figure of 26 909 is only marginally higher the standard deviation associated with the estimated accident year 10 liability. This has the counter intuitive implication that the individual accident year liability estimates are at most marginally positively correlated and probably negatively correlated. These issues and the material differences between the estimates warrant further examination and discussion. The Mack (1993) approach is critiqued in Appendix D.

It must be emphasized that the methods of this paper focus on the whole distribution of forecast liabilities, not just the first two moments. For the AFG data the estimated distribution is displayed in Figure 1. The quartiles of the estimated distribution are estimated as \$54, \$63 and \$73 thousand indicating a slightly right skewed distribution. The upper quartile has recently been suggested as an appropriate measure of risk. Further the 10%, 5% and 1% upper percentiles of the distribution are estimated to be \$85, \$92, \$101 and \$111 thousand respectively.

3 Development factors and the basic model

The models of this paper are stated in terms of the development factors

$$\delta_{ij} \equiv \ln\left(\frac{c_{ij}}{c_{i,j-1}}\right) \qquad \Rightarrow \qquad g_i \equiv \ln\left(\frac{c_{i,n-1}}{c_{i,n-i}}\right) = \delta_{i,n-i+1} + \cdots + \delta_{i,n-1} \ .$$

Thus δ_{ij} is the (continuously compounded) percentage growth in accident year's i cumulative in development year j. The δ_{ij} measure the rates of growth in liabilities moving across rows of the runoff triangle. The δ_{ij} approximate the "chain ladder" link ratios $c_{ij}/c_{i,j-1} \approx 1 + \delta_{ij}$.

The methods in this paper model the δ_{ij} using the available runoff data. These models are then used to forecast future δ_{ij} from which the forecast distribution of, for example, the g_i are derived.

3.1 Basic model

The basic model illustrates the overall features of our approach without the complications of extensions. The model assumes the development factors δ_{ij}, $i = 1, \ldots, n$ have a common mean μ_j and standard deviation σ_j:

$$\delta_{ij} \sim (\mu_j \ , \ \sigma_j^2) \ , \qquad i = 1, \ldots, n \ , \qquad j = 0, \ldots \ . \tag{3}$$

In other words the basic model states that the cumulatives c_{ij} are uncorrelated geometric "modulated" random walks for each accident year in the development year direction. By a modulated random walk is meant that the drifts μ_j and variances σ_j^2 fluctuate across the development years but are common across accident years.

Hertig (1985) was the first to introduce the model (3) although, unfortunately, not considered were the developments δ_{i0} in the first development year and their relationship to subsequent years' developments. Aspect associated with the estimation of the model were considered by Murphy (1994). Taylor (2000, pp 196-203) gives a good overview of the model as proposed by Hertig (1985).

The basic model (3) is a first step in analyzing outstanding liabilities similar to the first step of analyzing a time series as a random walk. It can be generalized to address issues such as "process" correlation between

Table 2: Comparison between forecast liabilities

accident	mean			standard deviation		
year	DC model	Mack	Poisson	DC model	Mack	Poisson
2	155	154	243	146	206	486
3	643	617	885	375	623	984
4	1702	1636	2033	753	753	1589
5	2845	2747	3582	1271	1456	2216
6	3953	3649	3849	1462	2007	2301
7	5954	5435	5393	2290	2228	2873
8	12293	10907	11091	5463	5344	4686
9	12578	10650	10568	6747	6284	5563
10	22859	16339	17654	11551	24509	12801
Total	62982	52135	55297	16260	26909	17357

accident, development or calendar years, and correlations between different triangles. Pertinent generalizations in a particular context are suggested by basic model diagnostics discussed in §3.4.

The basic model implies the cumulatives c_{ij} are correlated within accident years but not between accident years. With (3) the minimum mean square error linear predictor of g_i and associated prediction error variance is

$$\hat{g}_i \equiv \mu_{n-i+1} + \cdots + \mu_{n-1} \ , \qquad \nu_i^2 \equiv \sigma_{n-i+1}^2 + \cdots + \sigma_{n-1}^2 \ , \qquad i = 2, \ldots, n \ . \tag{4}$$

If the δ_{ij} are normal then the g_i for $i = 2, \ldots, n$ are independent normals and, given the μ_j and σ_j it is straightforward to simulate values from the normal forecast distribution $g_i \sim N(\hat{g}_i, \nu_i^2)$. Exponentiating the simulated g_i values, subtracting 1 and multiplying by $c_{i,n-i}$ as in (1) then gives the simulated values from the conditional distribution of $c_{i,n-1} - c_{i,n-i}$, conditioning on the known data.

In applications the μ_j and σ_j, and hence $\hat{g}_i$ and ν_i, are unknown. They may be set using actuarial judgement or estimated from the runoff data. Estimating the μ_j induces "estimation" correlation between forecasts for different accident years since the same μ_j estimates are used in the forecasts for different accident years. The forecast of $\hat{g}_i$ for different i are thus correlated in a known way and the distribution of (2) will be that of sum of correlated lognormals. Simulating from such a distribution is straightforward and Figure 1 is an example of such a simulated distribution.

The expressions in (4) do not assume normality of the development factors δ_{ij}. However if normality applies then $\hat{g}_i$ and ν_i are the conditional mean and error standard deviation of g_i given the observed data, respectively. Further, it then follows that the expected value of $c_{i,n-1}$ and associated coefficient of variation are

$$\hat{c}_{i,n-1} = c_{i,n-i} e^{\hat{g}_i + \nu_i^2/2} \ , \qquad \sqrt{e^{\nu_i^2} - 1} \ , \qquad i = 2, \ldots, n \ ,$$

respectively. Accordingly the forecast liabilities with respect to accident year i and associated coefficient of variation are

$$\hat{c}_{i,n-1} - c_{i,n-i} \ , \qquad \frac{\hat{c}_{i,n-1}}{\hat{c}_{i,n-1} - c_{i,n-i}} \sqrt{e^{\nu_i^2} - 1} \ , \qquad i = 2, \ldots, n \ .$$

Thus normal distribution assump tions are convenient but not essential. The examples below illustrate the appropriateness of the normal assumption and the applicability or otherwise of log normality of forecast liabilities

3.2 Basic model estimates

The bottom two rows of Table 3 displays estimates of the basic model parameters μ_j and σ_j for the AFG data:

$$\hat{\mu}_j = \frac{1}{n-j} \sum_{i=1}^{n-j} \delta_{ij} \ , \qquad \hat{\sigma}_j = \sqrt{\frac{1}{n-j} \sum_{i=1}^{n-j} (\delta_{ij} - \hat{\mu}_j)^2} \ , \qquad j = 0, \ldots, n-1 \ , \tag{5}$$

where $c_{i,-1} \equiv 1$ for $i = 1, \ldots, n$. The $\hat{\mu}_j$ for $j = 1, \ldots, n-1$ are the approximate percentage changes in cumulative incurred liabilities moving from development year $j-1$ to j. Thus for example $\hat{\mu}_4 = 0.17$ indicates that moving from the development year 3 to year 4 there is an average increase in cumulative incurred liabilities of approximately 17%. The standard deviations quantify the variability in the observed development factors for a given development year and across the accident years. For example $\hat{\sigma}_4 = 0.05$ indicates that a percentage increase from development year 3 to 4 as high as 0.17+2(0.05)=27% is unlikely.

3.3 Basic model forecast distribution

The forecast growth in accident year's i log-cumulative payments is $\hat{g}_i$ given in (4) with the μ_j replaced by the $\hat{\mu}_j$. The second column of Table 4 displays these estimates for the AFG data of Table 1. Note that the first displayed accident year is $i = 2$ emphasizing that, at this stage, no forecast is attempted beyond the latest development year.

The estimate of the"process" variance is the second equation in (4) with the σ_j^2 replaced by the estimates $\hat{\sigma}_j^2$. This is the estimated forecast error variance of $\hat{g}_i$ if the μ_j were known exactly. An additional source of error is "estimation" variance which quantifies the extra error due to estimating the μ_j. Under (3), the estimate $\hat{\mu}_j$ has variance $\sigma_j^2/(n-j)$ and the $\hat{\mu}_j$ are uncorrelated. Thus the estimation variance associated with the forecast of $\ln(c_{i,n-1})$ is

$$\frac{\sigma_{n-i+1}^2}{i-1} + \cdots + \frac{\sigma_{n-1}^2}{1} \ , \qquad i = 2, \ldots, n \ . \tag{6}$$

The sum of the process and estimation error variance gives the total variance. Its square root for the AFG data is displayed in the "std dev" column in Table 4 and where estimates $\hat{\sigma}_j$ are used instead of the σ_j. As expected

Table 3: Development factors and basic model estimates for the AFG data

accident	development year j									
year i	0	1	2	3	4	5	6	7	8	9
1	8.52	0.50	0.28	0.08	0.14	0.18	0.11	0.03	0.00	0.01
2	4.66	3.70	0.23	0.68	0.26	0.12	-0.01	0.04	0.03	
3	8.13	0.97	0.43	0.15	0.15	0.17	0.03	0.03		
4	8.64	0.71	0.31	0.30	0.10	0.11	0.04			
5	7.00	2.17	0.50	0.34	0.16	0.01				
6	7.32	1.45	0.60	0.10	0.20					
7	6.32	1.98	1.00	0.12						
8	7.21	1.64	0.64							
9	8.05	0.54								
10	7.63									
$\hat{\mu}_j$	7.35	1.52	0.50	0.25	0.17	0.12	0.04	0.03	0.02	0.01
$\hat{\sigma}_j$	1.12	0.96	0.24	0.20	0.05	0.06	0.04	0.01	0.01	0.00

Table 4: Basic model forecasting parameters for the AFG data

accident	forecast	growth	forecast	coef	correlation matrix of forecast log–error								
year i	growth $\hat{g}_i$	sd $\hat{\nu}_i$	liability	var	2	3	4	5	6	7	8	9	10
2	0.01	0.00	154	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.03	0.02	643	0.68	0.00	1.00	0.31	0.12	0.07	0.06	0.03	0.02	0.01
4	0.06	0.02	1698	0.33	0.00	0.31	1.00	0.13	0.08	0.06	0.03	0.02	0.01
5	0.10	0.05	2853	0.51	0.00	0.12	0.13	1.00	0.13	0.11	0.05	0.03	0.01
6	0.22	0.08	3968	0.42	0.00	0.07	0.08	0.13	1.00	0.15	0.07	0.05	0.01
7	0.39	0.10	5901	0.31	0.00	0.06	0.06	0.11	0.15	1.00	0.07	0.05	0.02
8	0.64	0.23	12416	0.49	0.00	0.03	0.03	0.05	0.07	0.07	1.00	0.09	0.03
9	1.14	0.34	12445	0.50	0.00	0.02	0.02	0.03	0.05	0.05	0.09	1.00	0.04
10	2.66	1.07	50033	1.53	0.00	0.01	0.01	0.01	0.01	0.02	0.03	0.04	1.00

the standard deviation associated with latter accident years are much higher reflecting the fact that there is much more uncertainty associated with the less developed accident years.

The middle columns of Table 4 indicates the correlations between the forecasts of the $\ln(c_{i,n-1})$ for the different accident years i. With the basic model (3) these correlations arise solely because the estimation of μ_j. The exact expressions for the underlying covariance matrix is given in Appendix C.3. Table 4 indicates that these correlations are generally higher for near accident years. The nonzero correlations imply that the sum of the forecasts of the $c_{i,n-1}$ is not the forecast of the sum.

The forecast liability column in Table 4 contains the expected values of $c_{i,n-1} - c_{i,n-i}$ computed assuming log normality. These appear plausible except for $i = 10$ which is unreasonably large. The estimated growth $\hat{g}_{10}$ appears appropriate but the large value of the standard deviation, $\hat{\nu}_{10} = 1.07$, is excessive. This latter value is a result of the large value of $\hat{\sigma}_1 = 0.96$ in Table 3, driven by the outlying value of $\delta_{21} = 3.70$. Thus the basic model does not seem to fit accident year $i = 2$ and the lack of fit results in an unreasonable forecast distribution. We address this issue later with an appropriate basic model extension. An ad–hoc approach is to reduce 1.07 to a more reasonable value. This is equivalent to discounting the experience of accident year 2.

Given the means, standard deviations and correlations displayed in Table 4, simulation is used to derive the estimated liability distribution. In particular, draws are made from the multivariate normal distribution with parameters displayed in Table 4. Each multivariate draw is added to $(\ln c_{2,n-2}, \ldots, \ln c_{n0})'$ and then exponentiated to arrive at a random draw of $(c_{2,n-1}, \ldots, c_{n,n-1})'$. Subtracting current liabilities $(c_{2,n-2}, \ldots, c_{n0})'$ yields estimates of the outstanding liabilities each accident year. Repeated drawing leads to different estimates and a detailed picture of the conditional liability distribution. For the runoff triangle Table 1 the simulated liability distribution based on 10 000 draws is displayed in Figure 2.

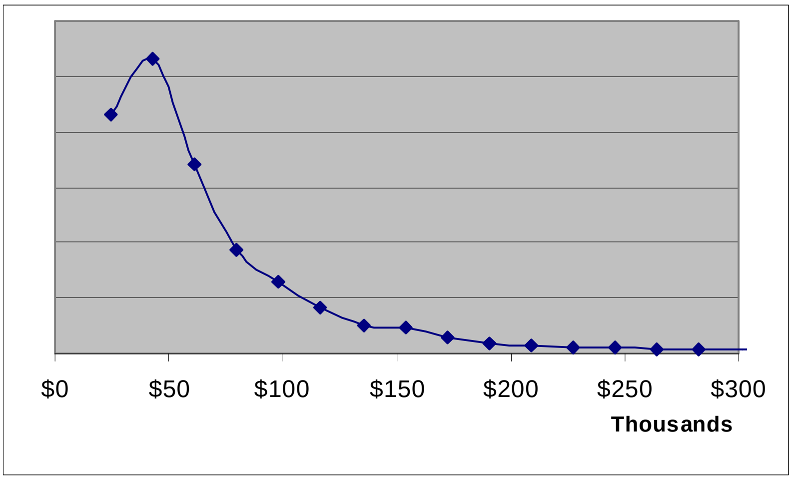

Figure 2: Liability distribution for the AFG data using basic model

The simulated forecast distribution synthesizes the information provided by the model (3) and the Table 1 data. The distribution Figure 2 is heavily skewed to the right. Assessment of the basic model as discussed in the next two subsections indicate that the basic model requires modification to reflect correlation within the triangle. These modifications are discussed in §4.1. The effect of these modifications is to alter the estimated liability distribution to yield the forecast liability distribution displayed in Figure 1 and Table 2.

3.4 Basic model assessment

The appropriateness of the basic model (3) is assessed using the standardized development factors

$$z_{ij} \equiv \frac{\delta_{ij} - \hat{\mu}_j}{\hat{\sigma}_j} \ . \tag{7}$$

These quantities measure the number of standard deviations that the observed development factor in accident year i is away from the average development $\hat{\mu}_j$ in development year j. Standardized development factors are, given the basic model (3), approximately normally distributed with zero mean zero and unit standard deviation. Thus large positive or negative z–scores suggests the inappropriateness of the constant mean and variance assumption for a given development year.

Table 5 displays the standardized development factors z_{ij} corresponding to the AFG data of Table 1. The first few entries in the second row in Table 5 suggest that accident year 2 is aberrant in that values as extreme as the ones observed should happen less than 1% of time if (3) holds. Trends across accident, development or calendar years are readily assessed with the standardized development factors.

Table 5: Standardized runoff triangle for the AFG data

accident	development year j									
year i	0	1	2	3	4	5	6	7	8	9
1	1.04	-1.06	-0.94	-0.87	-0.58	1.00	1.59	-0.15	-1.00	0.00
2	$\boxed{-2.39}$	$\boxed{2.27}$	-1.14	$\boxed{2.17}$	1.76	0.10	-1.17	1.29	1.00	
3	0.70	-0.57	-0.28	-0.51	-0.35	0.87	-0.31	-1.14		
4	1.15	-0.83	-0.80	0.24	-1.38	-0.17	-0.11			
5	-0.31	0.68	0.02	0.42	-0.18	$\boxed{-1.79}$				
6.	-0.02	-0.07	0.42	-0.77	0.72					
7	-0.91	0.48	$\boxed{2.14}$	-0.68						
8	-0.12	0.12	0.58							
9	0.62	-1.01								
10	0.25									

Framed numbers indicate significant (10%) departures from the basic model

Figure 3 plots the standardized development factors z_{ij} of Table 5 versus accident year i, development year j and accident year $i+j$. Evidence against the basic model is suggested when there is structure in these plots. From (7), the z_{ij} for each development year j have mean zero and standard deviation 1 and hence the top left panel of Figure 3 will display a centered and homoscedastic scatter of points. The accident year panel however can display mean and standard deviation departures from the basic model. For example the standardized development factors for accident year 2 have excessive variability. Similarly the bottom left calendar year panel indicates that calendar year 2 has consistently less than average development.

The final panel in Figure 3 is the normal probability plot of standardized development factors. Thus theoretical lower tail probabilities of the z_{ij} based on the standard normal distribution are plotted against empirical lower tail probabilities. Conformance to normality is indicated by all the points lying close to the 45^o line. For these data normality appears reasonably well supported.

Formal tests for significant departures from 0 and 1 of the estimated means and standard deviations of the z_{ij} are displayed Table 6. The "mean" rows in the body of the table are sample means of the z_{ij} for the indicated accident or calendar year with associated p-values appearing underneath. If the basic model holds then

$$\bar{z}_{i.} \equiv \frac{1}{n-i+1}\sum_{j=0}^{n-i} z_{ij} \sim N\left(0, \frac{1}{n-i+1}\right) , \qquad i = 1, \ldots, n . \tag{8}$$

Near zero p-values indicate a mean for that accident or development year significantly lower than expected while p-values near 1 indicate significantly higher development factors. Standard deviation rows in Table 6 indicate the standard deviations of the z_{ij} values across the corresponding accident or calendar years and the associated p-values. The p-values test whether or not development factors in the given year are unusually variable. For accident years, the p-values are determined on the basis of

$$\sum_{j=0}^{n-i} (z_{ij} - \bar{z}_{i.})^2 \sim \chi^2_{n-i} , \qquad i = 1, \ldots, n . \tag{9}$$

Low p-values indicate relative little variability in the development factors for that accident year while values near 1 suggest excessive variability. For these data, accident year 2 appears to have excessively high variability in the development ratios. The equivalent calendar year variability p-values are also based on (9) except that summation is over each calendar year. Calendar years do not appear to have extreme behavior in terms of variability.

3.5 Assessment of correlation within a runoff triangle

Table 7 displays correlation diagnostics for the AGF data of Table 1. The displayed correlations are correlations calculated from the z_{ij} and hence relate to the development factors δ_{ij}. The displayed development correlations are multiple correlations for predicting δ_{ij} from prior developments $\delta_{i,j-1}, \ldots, \delta_{i0}$ in the same accident year. Thus the cases for estimation are the accident years. For these data the values in development year $j = 1$ are highly predictable from those in year $j = 0$, violating the basic model assumptions. A plot of the relationship

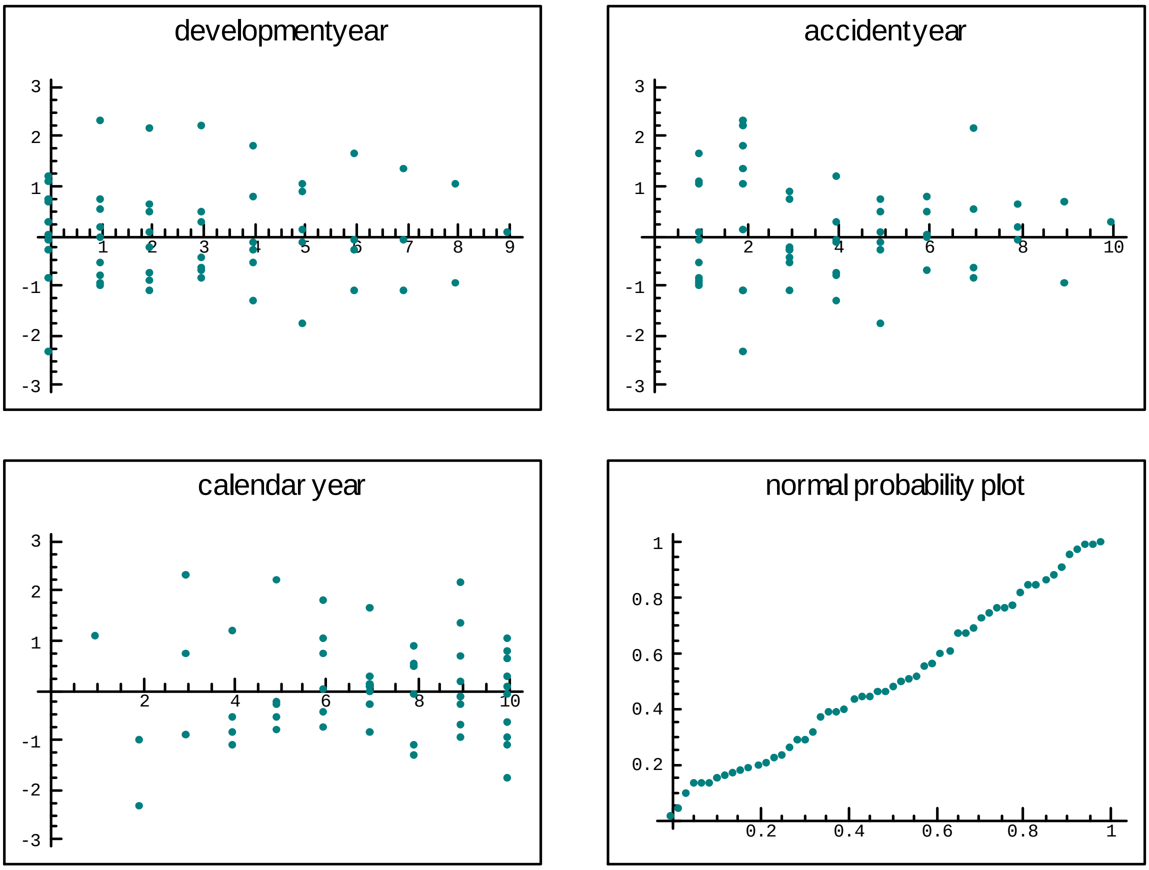

Figure 3: Standardized development factors plotted against development, accident and calendar year, and against log cumulatives.

Table 6: Basic model mean and variance diagnostics for the AFG data

	year									
	1	2	3	4	5	6	7	8	9	10
accident										
mean	-0.10	0.43	-0.20	-0.27	-0.19	0.05	0.26	0.19	-0.19	0.25
p-value	0.37	0.79	0.18	0.17	0.28	0.60	0.66	0.87	0.37	
std dev	0.93	$\boxed{1.57}$	0.62	0.77	0.79	0.50	1.21	0.29	0.82	
p-value	0.53	1.00	0.12	0.34	0.42	0.13	0.88	0.12	0.75	
calendar										
mean	1.04	$\boxed{-1.72}$	0.67	-0.36	0.03	0.35	0.09	-0.08	0.20	-0.22
p-value		0.00	0.81	0.21	0.53	0.83	0.63	0.38	0.73	0.21
std dev		0.67	1.31	0.89	1.09	0.89	0.71	0.76	0.95	0.86
p-value		0.65	0.92	0.64	0.79	0.55	0.26	0.29	0.57	0.41

All p-values are lower tail values. Framed numbers indicate significant (10% - two tailed) values.

between δ_{i0} and δ_{i1} is given in Figure 4 indicating δ_{i1} is almost perfectly negatively correlated with δ_{i0}. An adjustment to the basic model that addresses this correlation is given in §4.1. Other significant development year correlation occurs for development year 4.

Table 7: Basic model correlation diagnostics for the AFG data

	year									
	1	2	3	4	5	6	7	8	9	10
development	$\boxed{0.97}$	0.59	$\boxed{0.96}$	$\boxed{1.00}$	0.88	0.81	0.41			
p-value	1.00	0.66	0.96	0.94	0.42	0.42	0.27			
accident		$\boxed{0.67}$	0.63	0.64	0.97	$\boxed{1.00}$	0.86	0.99		
p-value		0.95	0.72	0.38	0.66	0.92	0.49	0.93		
calendar					0.81	0.94	0.77	0.66	0.20	0.04
p-value					0.60	0.66	0.59	0.58	0.08	0.01

All p-values are lower tail values. Framed numbers indicate significant (10% - upper tailed) values.

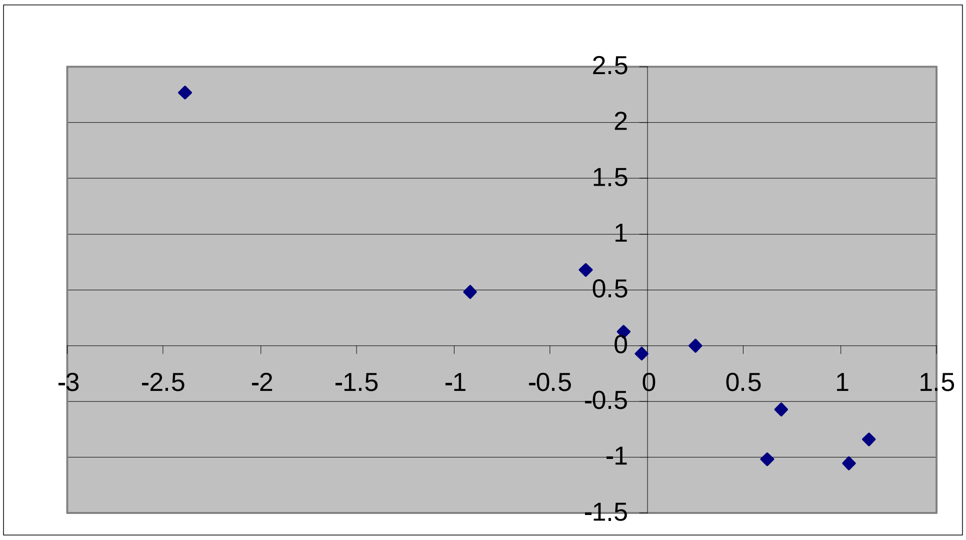

Figure 4: Relation between δ_{i0} and δ_{i1} for the AFG data

The accident correlations displayed in Table 7 are multiple correlation between δ_{ij} and the previous values $\delta_{i-1,j}, \delta_{i-2,j}, \ldots, \delta_{1j}$ for accident years $i = 2, \ldots, n-2$. Thus implicitly each δ_{ij} is predicted from previous accident year values in the same development year j and where the development years serve as cases for estimation. The p–values suggest that only accident year 2 is unusually well predicted.

Calendar year correlations displayed in Table 7 are the multiple correlation of $\delta_{i,k-i}$ with $\delta_{i,k-i-1}$ and $\delta_{i-1,k-i}$. The correlations are for calendar years $k = 4, \ldots, n$ since it is only for these calendar years there are sufficient cases for estimation.

4 Modelling correlations within runoff triangles

This section extends the basic model to incorporate correlation between development factors δ_{ij} falling in different development, accident or calendar years. Thus the possibility of "process" correlation is introduced with correlations modelled using relatively few parameters. Predicting with such models factor in process correlation making for more robust forecasts. The three dimensional nature of the data: accident, development and calendar years, admits considerable flexibility. Postulating correlation in any two of these directions generally implies correlation in the third direction.

4.1 Development year correlation

Development year correlation refers to the correlation between development factors at two development years in the same accident year. To introduce such correlation into the basic model, first rewrite the basic model (3) as

$$\delta_{ij} = \mu_j + h_j \epsilon_{ij} \ , \qquad i = 1, \dots n \ , \qquad j = 0, \dots, n-1, \tag{10}$$

where

$$h_j = \frac{\sigma_j}{\sigma} \ , \qquad \sigma \equiv \sigma_0 \ , \qquad \epsilon_{ij} \equiv \frac{\delta_{ij} - \mu_j}{h_j} \sim (0, \sigma^2) \ .$$

The ϵ_{ij} are standardized development factors: standardized on mean zero, variance σ^2, and uncorrelated across both accident years i and development years j. We now adjust (10) to

$$\delta_{i1} = \mu_1 + h_1(\epsilon_{i1} + \theta \epsilon_{i0}) \ , \qquad i = 1, \dots, n \ . \tag{11}$$

This is called the development correlation model and allows for correlation between δ_{i0} and δ_{i1} of

$$r = \frac{\theta h_1}{\sqrt{h_1^2(1+\theta^2)}} = \frac{\theta}{\sqrt{1+\theta^2}} \qquad \Rightarrow \qquad \theta = \frac{r}{\sqrt{1-r^2}} \ .$$

If $r < 0$ then a low value of δ_{i0} relative its mean μ_0 will be associated with a large value of δ_{i1} relative to its mean μ_1. Such correlation was observed in the AFG runoff triangle in Table 1.

The unknown correlation r can be estimated using moment estimates. Thus r is estimated by $\hat{r}$, the correlation between the δ_{i0} and δ_{i1}, $i = 1, \dots, n-1$. Then $\hat{r}$ serves to define $\hat{\theta}$. In turn, $\hat{h}_j = \hat{\sigma}_j / \hat{\sigma}_0$ for $j = 2, \dots, n-2$ while $\hat{h}_1 = \hat{\sigma_1}/(\hat{\sigma}_0\sqrt{1+\hat{\theta}^2})$. A smooth geometric decline in the σ_j can be imposed by regressing the $\ln \hat{\sigma}_j$ on $j = 2, \dots, n-2$ yielding least squares coefficients a and b say and $\hat{h}_j = \mathrm{e}^{a+bj}/\hat{\sigma}_0$, $j = 2, \dots, n-1$. A more refined approach is where all these so–called hyperparameters are estimated simultaneously using maximum likelihood.

Table 8 reports moment and maximum likelihood estimates of the correlation model parameters using the AGF data in Table 1 and indicates virtually no difference between the two sets of estimates.

Table 8: Parameter estimates for correlation model and Table 1 data

parameter	r	h_1	a	b	σ	ℓ_0
initial estimates	-0.97	3.30	-0.17	-0.54	0.31	-151.96
maximum likelihood estimates	-0.97	3.17	-0.08	-0.53	0.29	-152.72

Given the moment estimates, the maximum likelihood estimates of the development factors μ_j follow via closed form formulas. These, together with the derived estimates of the σ_j are used to generate the parameters of the forecasting distribution displayed in Table 9. Simulations from this distribution yield the forecast liability distribution displayed in Figure 1.

Table 9: Parameters of forecast distribution for the AFG data with the development correlation model

	forecast	std	forecast	coef	correlation matrix of forecast log–error								
i	growth $\hat{g}_i$	dev $\hat{\nu}_i$	liability	var	2	3	4	5	6	7	8	9	10
2	0.01	0.01	154	0.95	1.00	0.28	0.17	0.10	0.06	0.03	0.02	0.01	0.01
3	0.03	0.01	642	0.58	0.28	1.00	0.23	0.14	0.08	0.05	0.03	0.02	0.01
4	0.06	0.02	1701	0.44	0.17	0.23	1.00	0.18	0.11	0.06	0.04	0.02	0.02
5	0.10	0.04	2843	0.45	0.10	0.14	0.18	1.00	0.14	0.08	0.05	0.03	0.02
6	0.22	0.07	3948	0.37	0.06	0.08	0.11	0.14	1.00	0.11	0.07	0.04	0.03
7	0.39	0.12	5941	0.38	0.03	0.05	0.06	0.08	0.11	1.00	0.09	0.05	0.04
8	0.64	0.20	12243	0.44	0.02	0.03	0.04	0.05	0.07	0.09	1.00	0.08	0.06
9	1.14	0.35	12475	0.54	0.01	0.02	0.02	0.03	0.04	0.05	0.08	1.00	0.09
10	2.40	0.45	22957	0.50	0.01	0.01	0.02	0.02	0.03	0.04	0.06	0.09	1.00

Further θ parameters can be introduced to model correlations between adjacent higher order development years. For example

$$\delta_{i2} = \mu_2 + h_2(\epsilon_{i2} + \theta^* \epsilon_{i1}) \ , \qquad i = 1, \dots, n \ ,$$

in which case the correlation between δ_{i1} and δ_{i2} is

$$r^* = \frac{\theta^*}{\sqrt{(1+\theta^2)(1+\theta^{*2})}} \qquad \Rightarrow \qquad \theta^* = \frac{r^*\sqrt{1+\theta^2}}{\sqrt{1-r^{*2}(1+\theta^2)}} = \frac{r^*}{\sqrt{1-r^2-r^{*2}}} \ ,$$

which can again be estimated using the method of moments. However it is unlikely that such higher order correlations are likely to be important or significantly alter the forecasts.

4.2 Accident year correlation

Accident year correlation refers to the correlation between development factors in different accident years of the same development year. Accident year correlation is induced by assuming δ_{ij} has mean μ_{ij}, depending on both accident year i and development year j. Since year to year changes are likely to be small a reasonable setup is where, moving down each development year, the mean μ_{ij} is a random walk leading to

$$\delta_{ij} = \mu_{ij} + h_j \epsilon_{ij} \ , \qquad \mu_{i+1,j} = \mu_{ij} + \lambda_j \eta_{ij} \ , \qquad \epsilon_{ij},\ \eta_{ij} \sim (0, \sigma^2) \ . \tag{12}$$

Thus the average growth rate in claims evolves smoothly with accident year i. The basic model is the special case $\lambda_j = 0$ for $j = 0, \ldots, n-1$ implying the mean in each development year is constant. In this case the optimal estimate of the δ_{ij} given the data up to calendar year n is the simple average of $\delta_{1j}, \ldots, \delta_{n-j,j}$. If $h_j = 0$ then expected development ratios μ_{ij} are a random walk in the accident year direction and the best estimate of δ_{ij} for $i + j > n$ is $\delta_{n-j,j}$, the latest observed development factor in development year j. Intermediate cases are where the ratio h_j/λ_j is neither 0 nor infinite. If

$$c = \frac{2}{2 + \left(\frac{h_j}{\lambda_j}\right)^2 + \left(\frac{h_j}{\lambda_j}\right)\sqrt{4 + \left(\frac{h_j}{\lambda_j}\right)^2}} \ ,$$

then under (12) the optimal forecast of $\delta_{i+1,j}$ given the observed development ratios up to calendar year $i + j$ is approximately[1] an exponentially weighted moving average of observed δ_{ij} in the given development year j:

$$\tilde{\delta}_{i+1,j} \approx \tilde{\delta}_{ij} + c(\delta_{ij} - \tilde{\delta}_{ij}) = c\delta_{ij} + c(1-c)\delta_{i-1,j} + c(1-c)^2\delta_{i-2,j} + \cdots \ .$$

The forecast $\tilde{\delta}_{ij}$ equals $\tilde{\delta}_{i-1,j}$ or $\delta_{i-1,j}$ according as $c = 0$ or $c = 1$. When $c = 0$ then $h_j/\lambda_j \to \infty$ while $c = 1$ corresponds to $h_j/\lambda_j = 0$. Actuarial judgment can be used to set c or h_j and λ_j. Alternatively, these parameters can be estimated using maximum likelihood.

The effect of introducing accident year variation in the μ_j is to induce accident year correlation since for example a high value of μ_{ij} is likely to be followed by a high value of $\mu_{i+1,j}$. Correlations across accident years induce correlations amongst all future δ_{ij} in a complicated yet transparent way. For example suppose there is a general increase in the δ_{ij}'s with accident year i. Then δ_{ij}'s in the same accident year but different development years will appear to be positively related.

An examination of the AFG data of Table 1 revealed no significant evidence in favor of accident year variation in underlying rates and hence for the AFG data it is reasonable to assume $\lambda_j = 0$ for $j = 0, \ldots, n-1$. To illustrate methods we instead use Lumley's data previously considered in Mack ... . A three dimensional surface plot of the runoff triangle is given in Figure 5. The triangle is clearly growing in the accident year direction which is the axis heading south east.

For Lumley's data assume $h_j = \mathrm{e}^{bj}$ and $\lambda_j = rh_j$, $j = 0, \ldots, n-1$. Maximum likelihood estimation yielded $\hat{r} = 0.48$. The latter value is significantly different from zero and indicates significant fluctuation in the underlying development factors as we move down each development year column. The fact that development development factors fluctuate leads to increased spread in the forecast liability distribution since future development factors may be significantly different from the the average experienced over the past. Figure 6 compares results from the basic and accident correlation model and indicates exceedance probabilities under the two models. The basic model liability distribution is less spread out with a median of around 180 000 compared to around 210 000 for the distribution computed under the development year correlation model. The latter model induces a liability distribution which has a much heavier right hand tail with a 10% probability of liabilities exceeding 350 000.

A special case the of the accident year correlation model is where $\lambda_j = 0$ for $j = 1, 2, \ldots$. The model then postulates constant growth rates in liabilities (apart from random variation) for development years $j = 1, 2, \ldots$ while development year 0 growth rates follows a random walk in the accident year direction. This apparent extension of the basic model however has no direct impact on forecasts. This is because forecasts take as their base the latest cumulative $c_{i,n-i}$ and utilize development factors associated with development years $j = 1, 2, \ldots$. Thus the basic model allows for growth in claims across accident years.

4.3 Calendar year correlation

A model yielding correlation across calendar years is

$$\delta_{ij} = \mu_j + h_j(\tau_{i+j} + \epsilon_{ij}) \ , \qquad \tau_{i+j+1} = \tau_{i+j} + \lambda\eta_{i+j} \ , \qquad \epsilon_{ij},\ \eta_{i+j} \sim (0, \sigma^2) \ .$$

[1] The formula is approximate because of the limited number of observations in each column. In the examples below all smoothing is done exactly based on the model.

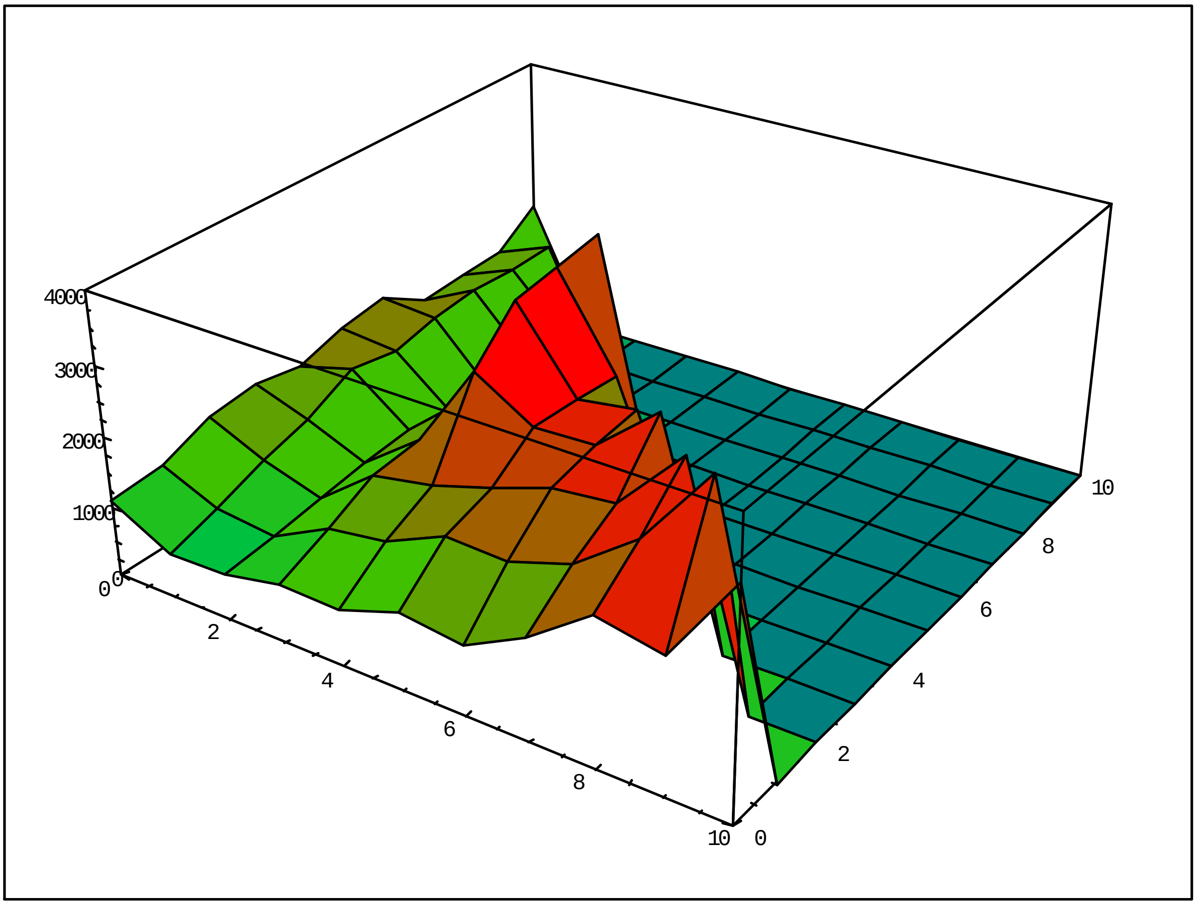

Figure 5: Runoff triangle for Lumley's data

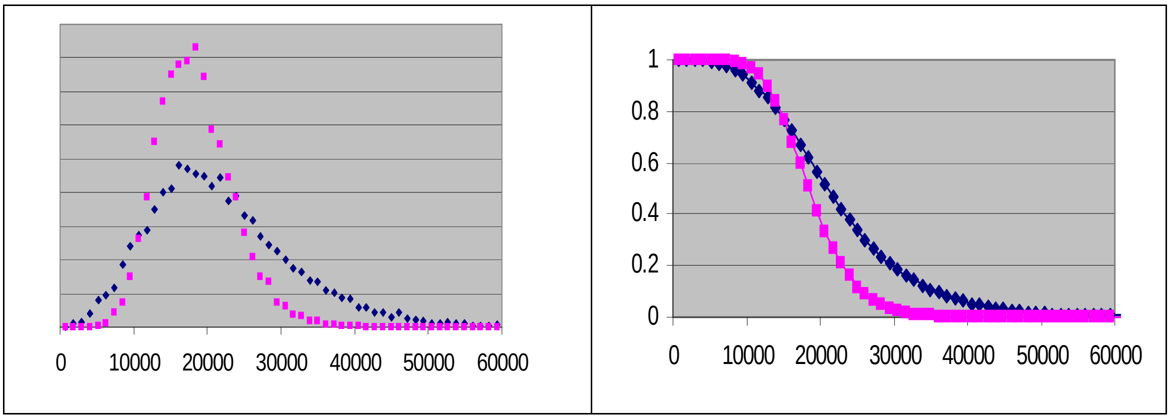

Figure 6: Simulated liability distribution and exceedance probabilities for Lumley's data under basic and accident year correlation model. Less variable distribution is from basic model analysis

Thus the τ_k operates in calendar time. Each τ_k serves to increase or decrease all the development factors falling in calendar year k with calendar year effects evolving as a random walk in calendar time. The effect of τ_k on a particular development factor is scaled by h_j and hence the effect is assumed proportional to the standard deviation associated with the development factor. The basic model is the case $\lambda = 0$. The other extreme $\lambda \to \infty$ corresponds to where the τ_k are unrelated and, from a technical point of view, fixed but unknown (see Appendix C.1). In this case the τ_k can be estimated using regression.

Initial estimates of the parameters of the calendar correlation model are obtained from the

$$\delta_{ij} - \delta_{i-1,j} = h_j(\epsilon_{ij} - \epsilon_{i-1,j} + \lambda\eta_{i+j-1}) \ ,$$

which have mean zero and variance $\sigma^2(2+\lambda^2)h_j^2$. Further, the average products of adjacent differences in the same development year have expectation $-\sigma^2 h_j^2$. Thus

$$r_j \equiv \frac{(n-i-1)\sum_{i=3}^{n-i}(\delta_{ij} - \delta_{i-1,j})(\delta_{i-1,j} - \delta_{i-2,j})}{(n-i-2)\sum_{i=2}^{n-i}(\delta_{ij} - \delta_{i-1,j})^2} \approx -\frac{1}{2+\lambda^2} \ , \qquad j = 0, \ldots, n-2$$

suggesting estimating λ by averaging the r_j to yield $\bar{r}$ and putting $\hat{\lambda} = \sqrt{-2 - 1/\bar{r}}$. This is a valid estimate if $\bar{r} < -1/2$ with $\bar{r}$ near -1/2 indicating lack of calendar correlation. The initial estimate can serve as a starting point for maximum likelihood estimation using the Kalman filter to evaluating the likelihood.

To illustrate, consider again the AFG runoff triangle of Table 1 and its standardized form in Table 5. These data do not suggest time trends but despite this we use them to illustrate methods. The coefficients variation $h_j\sigma/\mu_j$ is almost constant and we exploit this property in the determination of calendar year effects. Thus δ_{ij}/h_j are regressed on dummy variables corresponding to $i+j$ leading to the results in Table 10. To avoid parameter confounding $\tau_1 \equiv 0$ and hence $\mu_0 + \tau_k$ is the mean in calendar year k, development year 0, and $\tau_k - \tau_s$ is the difference in development year 0 means, between calendar year k and s. Further $\mu_j + \tau_k$ is the average percentage change between development year $j-1$ and j supposing the latter falls in calendar year k. Thus the difference between the average percentage changes in a given development year is $\tau_k - \tau_s$.

Table 10: Least squares estimates of calendar effects for the AFG data

	development year j or calendar year $k-1$									
	0	1	2	3	4	5	6	7	8	9
$\hat{\mu}_j$	8.52	2.82	1.58	1.45	1.37	1.30	1.24	1.18	1.18	1.28
$\hat{h}_j\hat{\sigma}$	0.59	0.70	0.45	0.29	0.12	0.09	0.14	0.08	0.12	
$\hat{\tau}_{k-1}$	0	-3.09	-0.27	-1.11	-1.35	-1.11	-1.33	-1.13	-1.06	-1.27

4.4 Calendar year inflation

Calendar year inflation is a form of correlation and can be dealt with as follows. Since the δ_{ij} are growth rates, calendar year inflation of rate r_{i+j} in calendar year $i+j$ implies δ_{ij} is increased to $\delta_{ij}(1+r_{i+j})$. Thus the basic model can be modified to

$$\delta_{ij} = (1+r_{i+j})\mu_j + (1+r_{i+j})h_j\epsilon_{ij} \ , \qquad \epsilon_{ij} \sim (0, \sigma^2) \ .$$

A constant rate of calendar year inflation increases both the mean and standard deviation for all development factors.

Trends in claims over over the accident years is allowed for with the basic model since each accident year's development starts off from from the relevant c_{i0}. Hence a high or low value in c_{i0} automatically shifts up or down the subsequent development profile for that accident year. Thus the assumption that the δ_{i0} all have the same mean μ_0 is of no import from the forecasting point of view since the forecast liability for each accident year takes off from $c_{i,n-i}$, the latest observed cumulative for that accident year.

4.5 General discussion of correlation modelling in runoff triangles

To further understand correlation modelling it is useful to write

$$\delta_k^{\rightarrow} \equiv (\delta_{k0}, \delta_{k1}, \ldots)' \ , \qquad \delta_k^{\downarrow} \equiv \begin{pmatrix} \delta_{1k} \\ \delta_{2k} \\ \vdots \end{pmatrix} \ , \qquad \delta_k^{\searrow} \equiv \begin{pmatrix} \ddots & & \\ & \delta_{k-1,k-1-i} & \\ & & \delta_{k,k-i} \\ & & & \ddots \end{pmatrix} \ , \qquad \delta_k^{\nearrow} \equiv \begin{pmatrix} & & \iddots \\ & \delta_{k-1,1} & \\ \delta_{k0} & & \end{pmatrix} \ , \quad (13)$$

where in each case the vector is considered as a column vector. Then $\delta_k^{\rightarrow}$ and $\delta_k^{\downarrow}$ are the vectors of development factors falling in accident year k and development year k, respectively. Development correlation models specify the covariance matrix $\text{cov}(\delta_k^{\rightarrow})$ while accident correlation models deal with the covariance matrix $\text{cov}(\delta_k^{\downarrow})$. In the development correlation models of §4.1 it is assumed $\text{cov}(\delta_k^{\rightarrow})$ is zero except on the diagonal and in the top 2×2 or 3×3 portion of the covariance matrix. Further it is assumed that the $\delta_k^{\rightarrow}$ are uncorrelated for different k. In the accident correlation model of §4.2, $\text{cov}(\delta_k^{\downarrow})$ is that of a random walk plus noise and the $\delta_k^{\downarrow}$ are assumed uncorrelated for different k.

The calendar year correlation model of §4.3 parametrizes both $\text{cov}(\delta_k^{\searrow})$ and $\text{cov}(\delta_k^{\nearrow})$. An alternative specification parametrizing $\text{cov}(\delta_k^{\searrow})$ is where standardized development factors are related across calendar years as follows:

$$z_{k,k-i} = \phi z_{k-1,k-1-i} + \eta_{k,k-i} \qquad \Rightarrow \qquad \delta_{k,k-i} = \mu_{k-i} + \phi\sigma_{k-i}\frac{\delta_{k-1,k-1-i} - \mu_{k-1-i}}{\sigma_{k-i-1}} + \sigma_{k-i}\eta_{k,k-i} \tag{14}$$

Preliminary estimates of μ_j and σ_j and can be those derived from the basic model estimates. More refined estimates utilize maximum likelihood. If $\phi = 1$ then we have random walks in the calendar year direction. The model can be extended with extra autoregressive and moving average terms.

Alternatively development factors within a calendar year may be correlated and different calendar years may be uncorrelated. Thus $\text{cov}(\delta_k^{\nearrow})$ is specified in terms of unknown parameters while the $\delta_k^{\nearrow}$ for different k are assumed uncorrelated. For example if from calendar year to calendar year there are common causes for increases or decreases. The accident year correlation model (12) can accommodate such correlation by imposing correlation amongst the $\epsilon_k^{\nearrow}$ or $\eta_k^{\nearrow}$ which are the disturbances associated with a diagonal of the runoff triangle. For example it may be assumed that the correlation matrix of $\epsilon_k^{\nearrow}$ is banded with 0.5 on the first off diagonal and 0.25 on the second off diagonal and zero on all the other off diagonals. This model implies that the increase or decrease in developments in a given calendar year have common causes but which impact for just that calendar year. A similar correlation structure on the $\eta_k^{\nearrow}$ implies similar common causes with the additional feature that their effects persists across calendar years since the increase in the level μ_{ij} serves to increase subsequent levels.

5 Further forecasting issues

This section deals with a number of further forecasting issues.

5.1 Forecasting individual terms or calendar year liabilities

The previous development has mainly focussed on forecasting future growth rates $g_i \equiv \delta_{i,n-i+1}+\cdots+\delta_{i,n-1}$ and hence liabilities associated with each accident year up to development year $n-1$ and total liabilities summing across all accident years. However all the methods can be applied to forecasting individual growth rates δ_{ij} and hence different sums of liabilities. For example

$$\hat{c}_{ij} = c_{i,n-i}\mathrm{e}^{\hat{\delta}_{i,n-i+1}+\hat{\delta}_{ij}} \ , \qquad c_{i,n-i}\mathrm{e}^{\hat{\delta}_{i,n-i+1}+\hat{\delta}_{i,j-1}}\left(\mathrm{e}^{\hat{\delta}_{ij}}-1\right) \ , \tag{15}$$

are forecast of c_{ij} and $c_{ij} - c_{i,j-1}$ respectively. Forecasts of liabilities falling due in future calendar years t are sums terms given by the second expression in (15) with $i + j = t$. Given the joint distribution of the forecasts $\hat{\delta}_{ij}$ and the associated errror covariance matrix, simulation can be used to derive the distribution of such sums of liabilities.

5.2 Forecasting beyond the latest development year

The approach of the previous sections has focussed on forecasting up to and including the latest observed development year $n-1$. Forecasting beyond the latest development year requires assumptions about claims development in this region of the runoff triangle. In this paper, these assumptions are couched in terms of models which may be estimated and assessed using the available data. This section discusses one approach and illustrates methods using the AFG data.

The standardized runoff Table 5 suggests that both the μ_j and h_j are decreasing in j, and do so at a decreasing rate. A least squares fit of $\ln\hat{\mu}_j$ and $\ln\hat{\sigma}_j$ for the development years $j = 0, 1, \ldots, 9$ yields

$$\ln\mu_j \approx 0.668 - 0.597j \ , \qquad \ln\sigma_j \approx 0.098 - 0.613j \ . \tag{16}$$

These relations can be used to extrapolate out to unobserved development years. The predicted μ_j and σ_j are then used in the forecasting formulas for the log of claims.

The above relations are a potential source of both increasing and decreasing confidence in the predictions. Forecasts which utilize predicted development factors beyond development year $n-1$, are likely to be subject to more sampling error than those utilizing development factors on which there are actual observations. On the other hand, the strength of the fitted relations in (16) indicate that the independent estimation of the development factors can be improved upon by pooling information across development years analogous to (16). Further the fitted relations in (16) indicate that the development years' means and standard deviations imply a virtual constant coefficient of variation

$$\frac{\sigma_j}{\mu_j} \approx \mathrm{e}^{0.098-0.668} = 0.566 \; .$$

Using such features in the estimation leads to more precise parameter estimation and runoff forecasting.

5.3 Stress testing forecasts and the Bornheutter-Ferguson method

Varying the forecast distribution parameters in Table 4 indicates the sensitivity of the forecast to different estimates. For example analysing the AFG data with the basic model led to a liability calculation of 50 033 for accident year 10. This estimate is critically influenced by the large estimate of σ_1, which in turn is the result of an unexpectedly large development factor $\delta_{21} = 3.70$. Table 5 shows this development factor is well outside the expected range if (3) applies. Adjusting the estimate $\hat{\sigma}_1$ from 0.96 to 0.61, the standard deviation computed from the other development factors in development year 1, only affects the accident year 10 forecast liability, reducing it from 50 033 to

$$\mathrm{e}^{10.287+0.61^2/2} - 2063 = 35\;350 - 2063 = 33\;287 \; .$$

The resulting simulated aggregate liability distribution is as displayed in Figure 7 and factors in the Table 4 estimation correlations between accident years. The revised aggregate distribution has a median of about \$69 000 and a 10% upper percentile of about \$104 000. Thus although the distribution is less skewed it retains considerable upside risk.

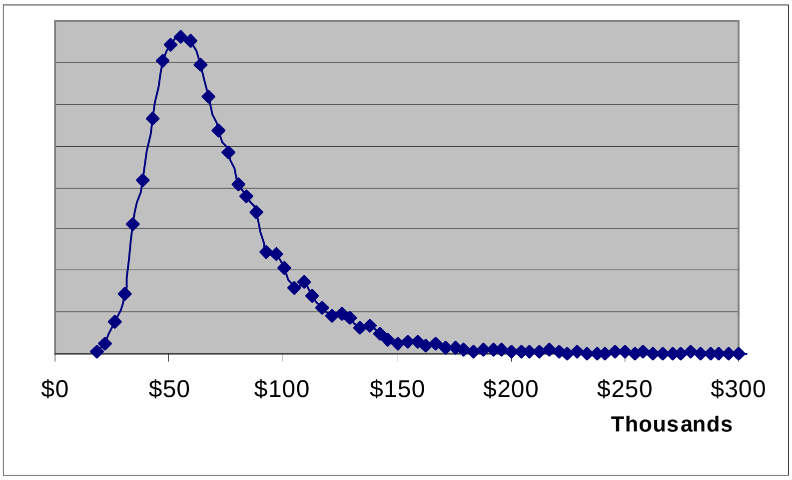

Figure 7: Liability distribution for the AFG data after adjusting $\hat{\sigma}_1$ from 0.96 to 0.61

The above adjustment technique is related to the well known Bornheutter-Ferguson method. This method is based on the idea that ultimate claims for each accident year can be predicted with increased precision using external information. The liability with respect to accident year i is

$$c_{i,n-i}(\mathrm{e}^{g_i} - 1) = c_{i,n-1}(1 - \mathrm{e}^{-g_i}) \; .$$

Given an external forecast $\hat{c}_{i,n-1}$ of $c_{i,n-1}$ then the Bornheutter-Ferguson estimate of outstanding liabilities with respect to accident year i is

$$\hat{c}_{i,n-1}(1 - \mathrm{e}^{-g_i}) \; .$$

Thus we discount the forecast $\hat{c}_{i,n-1}$ by the percentage estimates provided by the basic model analysis. This is analogous to the above approach where components driving the forecast are analysed and adjusted to make for a more reasonable forecast.

A further robustifying adjustment is to smooth the development factors μ_j or σ_j. To illustrate methods, consider the AFG data of Table 1. For these data suppose the basic model with the additional assumption that

$\sigma_j = e^{a+bj}$ for $j = 1, \ldots, n-2$. Thus the σ_j for $j > 0$ are assumed to decline geometrically as a function of development year j.

Correlations and means can be similarly varied to reflect actuarial judgement and assess sensitivity of forecasts to changes in parameters. However a number of cautions are in order:

- The approach is ad–hoc in that there is limited scope for assessing the suitability of the imposed parameters in terms of the data.
- Correlations cannot be set arbitrarily adjusted since the correlation matrix must be positive semi–definite.
- There is no attempt to distinguish between process and estimation correlation. Correlation induced by estimating parameters from the given runoff triangle are best dealt with, if possible mechanically. On the other hand process correlations are a modelling issue wherein experience gained from elsewhere is imposed on the general structure of the development process.

6 Comparison of the basic model to other models

The basic model (3) is related to the chain ladder method and it is useful to compare the two approaches. To facilitate comparison, the second column of Table 11 summarizes the structure of the basic model. The mean and standard deviation rows indicate that the estimates the μ_j and volatilities σ_j are the sample mean and standard deviation of the observed development forces $\ln(c_{ij}/c_{i,j-1})$, $i = 1, \ldots, n-j$. The final row indicates that the distribution of each term in (2) is known analytically but simulation is required to evaluate the distribution of the sum across accident years.

The third column of Table 11 displays the implied form of the basic model in terms of the development ratios $c_{ij}/c_{i,j-1}$. This implied form permits direct comparison to the chain ladder method and model and assumes the normal distribution in (3). The results follow from properties of the lognormal distribution as discussed in for example (Aitchison and Brown 1957, p.87). The basic model thus implies that the observed development ratios have a coefficient of variation independent of δ_j and directly proportional to σ_j. The "mean estimate" row indicates that the average development ratio is, ignoring the bias adjustment, estimated using the geometric mean of the observed development ratios.

The "chain ladder – Mack" column in Table 1 deals with the simplest form of the chain ladder method and the related model proposed by Mack (1993). The chain ladder method was developed without reference to a model, proposing the development ratio mean estimates

$$\hat{m}_j = \frac{\sum_{i=1}^{n-j} c_{ij}}{\sum_{i=1}^{n-j} c_{i,j-1}} = \frac{\sum_{i=1}^{n-j} \left(\frac{c_{ij}}{c_{i,j-1}}\right) c_{i,j-1}}{\sum_{i=1}^{n-j} c_{i,j-1}} \ , \qquad j = 0, \ldots, n-1 \ . \tag{17}$$

This is weighted average of observed link ratios. Other entries in the "chain ladder – Mack" column of the Table 11 summarize the model proposed by Mack (1993) and which can be used to justify (17). With this model the volatility of the development ratios are assumed to inversely proportional to $\sqrt{c_{i,j-1}}$.

The stochastic chain ladder method Renshaw and Verrall (1994) , Collins and White (2001) also employs weighted averages of observed ratios to estimate an implicit model average development ratio. Future link ratios are then simulated from appropriate distributions with this estimated averages and estimated volatilities. The weightings are determined using actuarial judgement and need not be the same as in (17). The simulated distributions reflect reflect both estimation uncertainty and future assumed process uncertainty. The approach is a flexible alternative to more analytically based methods. The statistical model underlying the method is left implicit. The method permits a considerable amount of subjective input into the claims forecasting process. This is both an advantage and disadvantage. It is in an advantage in that subjective information can be a very useful and important for improved claims forecasting. It is a potential disadvantage in that the subjective information may be inappropriate.

The forecast distribution under the different approaches is derived differently. With the basic model, the means, variances and correlations associated with future forces of development are analytically determined using minimum mean square error prediction. In these expressions, estimates replace model parameters. The joint distribution of future $\ln(c_{ij})$ across different accident and development years is also analytically determined assuming log-normality of future forces of development. Thus (2) is the sum of analytically determined correlated log normals. The location and shape of this final, easily described distribution are determined from simulation.

Mack (1993) derives estimates of the first two moments of the forecast distribution under the stated model assumptions and using ad–hoc estimates of individual terms. With the stochastic chain ladder method all future development ratios are individually simulated from distributions constructed using actuarial judgement and these simulated values are used to "project out" the likely actual liabilities. Repeated simulations yield estimates of the likely distribution of future liabilities.

Table 11: Comparison of Basic model and chain ladder approaches

	Basic Model		Chain ladder	
	Development factors	Implied ratio form	Mack	Stochastic
Model components	$\delta_{ij} = \ln \frac{c_{ij}}{c_{i,j-1}}$	$\mathrm{e}^{\delta_{ij}} = \frac{c_{ij}}{c_{i,j-1}}$	$r_{ij} = \frac{c_{ij}}{c_{i,j-1}}$	$r_{ij} = \frac{c_{ij}}{c_{i,j-1}}$
Mean	μ_j	$\mathrm{e}^{\mu_j+\sigma_j^2/2}$	m_j	implicit
Standard deviation	σ_j	$\mathrm{e}^{\mu_j+\sigma_j^2/2}\sqrt{\mathrm{e}^{\sigma_j^2}-1}$	$\frac{\nu_j}{\sqrt{c_{i,j-1}}}$	implicit
Coefficient of Variation	$\frac{\sigma_j}{\mu_j}$	$\sqrt{\mathrm{e}^{\sigma_j^2}-1} \approx 0.71\sigma_j$	$\frac{m_j\sqrt{c_{i,j-1}}}{\nu_j}$	implicit
Distribution	normal or other	log–normal	“distribution free”	any
Mean estimate	$\hat{\mu}_j = \frac{1}{n-j}\sum_{i=1}^{n-j}\delta_{ij}$	$\mathrm{e}^{\hat{\mu}_j+\hat{\sigma}_j^2/2}$	$\hat{m}_j = \sum_i w_i r_{ij}$	$\sum_i w_i r_{ij}$
Standard deviation estimate	$\hat{\sigma}_j = \sqrt{\frac{1}{n-j}\sum_{i=1}^{n-j}(\delta_{ij}-\hat{\mu}_j)^2}$	$\mathrm{e}^{\hat{\mu}_j+\hat{\sigma}^2/2}\sqrt{\mathrm{e}^{\hat{\sigma}_j^2}-1}$	$\sqrt{\sum_i w_i(r_{ij}-\hat{m}_i)^2}$	weighted sd
Forecast of $(c_{2,n-1},\cdots,c_{n,n-1})'$	multivariate log normal		approximate first two moments	simulation
Forecast of $\sum_{i=2}^{n}(c_{i,n-1}-c_{i,n-i})$	simulation from correlated lognormals		approximate first two moments	simulation

7 Correlation between runoff triangles

Table 12 displays three runoff triangles, in standardized form, corresponding to three classes of business transacted by an insurance company. The framed number in each panel indicates the correlation between development years 0 and year 1.

This section considers the cross correlation between runoff triangles its impact on liability valuation. Generally speaking, positive correlation between different classes of business, will require higher prudential margins and vice versa.

The standardized runoff triangles in Table 12 suggest serial correlation moving down development year 0. The first class of business in particular suggests growth. We ignore issues of serial correlation and instead focus on the correlation between the different classes of business.

A broad picture of the overall correlation in the triangles is gained by computing the correlations between the z–scores of the three standardized triangles. The resulting correlation matrix[2] is

$$\begin{pmatrix} 1.000 & 0.078 & -0.066 \\ 0.078 & 1.000 & -0.077 \\ -0.066 & -0.077 & 1.000 \end{pmatrix} \tag{18}$$

A more detailed picture emerges in computing correlations between the same development year of the different triangles. These are displayed in Table 13.

APPENDIX

A Excel implementation: The `fcast` workbook

Calculations described in this paper are implemented in an Excel workbook called `fcast.xls`. All tables and figures in this article are copies from Excel output. The `fcast.xls` workbook contains a number of spreadsheets:

- **Cumulatives.** Copy the triangle to be analyzed to this spreadsheet so that the top left entry of the triangle is in cell B2. It is assumed that the triangle contains cumulatives with as many rows (accident periods) as columns (development periods). Entries below the diagonal are cleared.

  Perform calculations on the triangle by pressing the keys described below. Results are placed into the spreadsheets enumerated below.

  - Ctrl–a. Accident correlation model of §4.2.
  - Ctrl–b. Basic model of §3.1.
  - Ctrl–d. Development correlation model of §4.1.
  - Ctrl–p. Plot of cumulatives c_{ij} and the standardized residuals z_{ij} from the basic model. Also plotted are the average development factors $\hat{\mu}_j$ and standard deviations $\hat{\sigma}_j$ against j. To avoid scale issues $\hat{\mu}_0$ is set to $\hat{\mu}_1$. The plot is copied to the clipboard and hence available for pasting into other documents
  - Ctrl–s. Simulate the liability distribution. Parameters used in the simulation depend on the active worksheet as described below.
  - Ctrl–z. The z–scores or standardized residuals plotted as in Figure 3. The plot is copied to the clipboard and hence available for pasting into other documents.

- **Developments.** The first table in this spreadsheet corresponds to the display in Table 3. The body of the table contains the δ_{ij}. Under the basic model, the last two rows of the table correspond to the $\hat{\mu}_j$ and $\hat{\sigma}_j$ defined in (5). With development or accident correlation model, the $\hat{\mu}_j$ correspond to the generalized least squares estimates where $\ln \hat{\sigma}_j \approx a + bj$ for $j = 1, \ldots, n-1$ while $\hat{\sigma}_0$ is xx . The parameters of the development or accident correlation model are determined using maximum likelihood.

  The second table is the incremental form of the runoff triangle "filled out" with the forecasts as computed with the last used model. The final row contains sums along each future diagonal of the filled out triangle and hence displays the expected liabilities falling due in each future calendar year.

  Press Ctrl–s to simulate the basic model from the given development factors $\hat{\mu}_j$ and the associated standard deviations $\hat{\sigma}_j$. Thus changing the $\hat{\mu}_j$ or $\hat{\sigma}_j$ and simulating explores the sensitivity of the basic model forecasts to the estimates. An example is given in §5.3.

[2]The three top right entries in the standardized triangles are ignored since they are constrained to ± 1 or 0.

Table 12: Runoff triangles for three classes of business ($ million)

accident	development year j						
year i	0	1	2	3	4	5	6
1	48.052	50.652	50.734	50.766	50.772	50.810	50.848
	-0.897	1.503	0.814	0.341	-1.404	1.000	0.000
2	49.490	51.967	51.845	51.825	51.884	51.921	
	-0.786	1.152	-1.362	-1.042	0.553	-1.000	
3	49.150	50.562	50.701	50.775	50.841		
	-0.812	-0.723	1.442	1.480	0.851		
4	51.622	53.514	53.480	53.469			
	-0.628	-0.021	-0.432	-0.779			
5	65.080	66.645	66.598				
	0.239	-1.136	-0.463			$\boxed{-0.422}$	
6	80.850	83.123					
	1.051	-0.775					
7	99.672						
	1.834						
$\hat{\mu}_j$	17.927	0.036	0.000	0.000	0.001	0.001	0.001
$\hat{\sigma}_j$	0.267	0.011	0.002	0.001	0.001	0.000	0.000
1	28.956	30.945	30.900	30.940	30.930	30.927	30.908
	-1.646	0.344	-1.543	1.173	0.867	-1.000	0.000
2	33.175	35.152	35.342	35.366	35.348	35.357	
	-0.596	0.016	1.479	0.243	0.534	1.000	
3	36.203	37.608	37.730	37.712	37.650		
	0.078	-0.744	0.537	-1.592	-1.401		
4	32.626	36.410	36.474	36.498			
	-0.725	2.004	-0.111	0.176			
5	38.885	40.310	40.359				
	0.630	-0.824	-0.363			$\boxed{-0.482}$	
6	38.985	40.444					
	0.650	-0.795					
7	44.142						
	1.608						
$\hat{\mu}_j$	17.395	0.057	0.002	0.001	-0.001	0.000	-0.001
$\hat{\sigma}_j$	0.130	0.026	0.002	0.001	0.001	0.000	0.000
1	6.953	8.189	8.301	8.286	8.408	8.446	8.484
	0.517	1.028	-0.513	-0.325	0.870	1.000	0.000
2	6.927	8.310	8.680	8.620	8.611	8.592	
	0.497	1.375	1.272	-0.955	0.531	-1.000	
3	8.939	9.810	9.770	9.909	9.055		
	1.877	-0.314	-1.577	1.680	-1.401		
4	4.966	5.140	5.258	5.246			
	-1.303	-1.428	0.024	-0.400			
5	5.202	5.530	5.730				
	-1.052	-0.921	0.794			$\boxed{0.526}$	
6	6.104	6.904					
	-0.187	0.259					
7	5.925						
	-0.348						
$\hat{\mu}_j$	15.659	0.110	0.022	0.001	-0.026	0.001	0.005
$\hat{\sigma}_j$	0.185	0.053	0.017	0.008	0.046	0.003	0.000

Framed numbers are correlations between development years 0 and 1

Table 13: Correlations between portfolio runoffs

development year											
0			1			2			3		
1.000			1.000			1.000			1.000		
0.874	1.000		0.431	1.000		-0.456	1.000		-0.586	1.000	
-0.378	-0.200	1.000	0.705	-0.283	1.000	-0.961	0.307	1.000	0.921	-0.839	1.000

- **Forecasts.** The simulated liability distribution corresponding to the chosen model. The first row corresponds to different values of the total liability x say. The second row contains simulated histogram values denoted $f(x)$. The final row contains $1 - F(x)$ where $F(x)$ is the simulated cumulative distribution. This row thus indicates inferred probabilities of liabilities exceeding the given x value. Each row contains 100 entries and the simulation is based on 10,000 simulations from the multivariate lognormal with parameters displayed in the third table of the spreadsheet. The histogram $f(x)$ is plotted and the figures in this paper are copies of those charts.

  The second table displays the mean, standard deviation and coefficient variation and upper percentiles: 75%, 50% (median), 25%, 10%, 5%, 2.5% and 1%.

  The third table displays "forecast parameters" organized similar to Table 4 or 9. The first two columns contain $\hat{g}_i$ and $\hat{\nu}_i$ for $i = 2, \ldots, n$, the predicted growth rates through to development year $n-1$ and the associated error standard deviations under the chosen model. The third and fourth column contain the estimated forecast liability $\hat{c}_{i,n-1} - c_{i,n-i}$ for each accident year and associated coefficient of variation computed under the log normality assumption. The last $n-1$ columns contain the estimated correlation matrix of $(\hat{g}_2, \ldots, \hat{g}_n)'$.

  Press Ctrl–s to simulate the liability distribution corresponding to the displayed forecast parameter table. Thus changing the forecast parameters and simulating (an example is described in §5.3) explores the sensitivity the forecast distribution to alternative forecast parameter values. You can change the $\hat{g}_i$, $\hat{\nu}_i$ or correlations. Any negative eigenvalues of the correlation matrix are set to zero.

- **Diagnostics.** The initial table displays the z_{ij} in (7), the standardized residuals from the fitted model. Table entries are in red if they lie outside ± 1.65, corresponding to 10% significance. The second table contains the diagnostics similar to those displayed in Table 6. Diagnostics with a p–value of less than 5% corresponding to a one or two sided test of significance are highlighted in red. The final third table contains correlation diagnostics as in Table 7. Correlations with a p–value of less than 10% (two–sided) are flagged in red.

The basic model is run when the workbook is opened. Spreadsheets are updated with every recalculation. Thus tables in each spreadsheet correspond to those from the last used model or calculation and initially spreadsheets contain the results from a basic model analysis. Pressing Ctrl–m copies a highlighted runoff triangle from a spreadsheet within the workbook to the "Cumulatives" spreadsheet and runs the basic model.

B Installing the `fcast` Excel workbook

Obtain the file `fcast.exe` from the author and install the Excel workbook `fcast.xls` as follows:

1. Double click on `fcast.exe` and when prompted, specify `c:\` as the installation drive (**thus overwrite the default**). Files will be placed in the `c:\fcast` subdirectory.

2. Double click on the file `c:\fcast\j406d.exe`. When prompted, install to the directory `c:\fcast` (**thus overwrite the default**). Click "OK" various times when prompted as the system installs. Finally click "yes" to "do you want to close?".

3. Copy the file `c:\fcast\jsutil.xls` to the subdirectory containing the program file `excel.exe`. Usually this is `c:\Program Files\Microsoft Office\Office10` or similar.

The Excel workbook `fcast.xls` will be contained in the subdirectory `c:\fcast`. Use it as the same as any other Excel workbook. Extra built–in functions are described in Appendix A. The first time you use the workbook, an information dialogue will display. Click the "Do not show again" box on bottom left and do a "file—exit" from the pull down menu on top left.

To conserve disk space, delete `fcast.exe`, `c:\fcast\j406d.exe` and `c:\fcast\jsetup.exe`: they are no longer needed. The `fcast.xls` workbook uses the J system as a "server" which was installed on your computer with above steps. You can ignore J but leave it installed. Information regarding J is available at `www.jsoftware.com`.

C Technical statistical details

This section outlines the technical background for the methods and models of this paper. This technical background makes explicit what is easily glossed over in the basic model. The general treatment makes the models and methods transparent and aligns the runoff triangle valuation problem with modern methods of forecasting and filtering.

C.1 Estimation

The estimating technology underlying the approach of this paper has the following features:

- Mean level parameters such as the μ_j in the basic model (3) are estimated via Generalized Least Squares (GLS). With GLS, the mean level parameters are regarded as fixed and unknown. Insofar as estimation is concerned, however, this is equivalent to regarding them as random with infinitely large variances. In the latter case the estimation variances are interpreted as mean square errors. Thus the fortunate situation is that two distinct formalisms lead to exactly the same calculations.
- Variances, such as the σ_j^2 in the basic model (3), covariances and correlations are termed hyperparameters and are estimated using maximum likelihood. The likelihood assumes normally distributed disturbances. Maximum likelihood estimates take into consideration the entire structure of the model, including implied means, variances and correlations.
- Forecasts are minimum mean square error linear predictors of the logarithms using the logarithms of the observed data and given the hyperparameter estimates. Minimum mean square error predictors are conditional means given the data and assuming normally distributed errors. The minimum mean square errors are error variances. The error covariance matrix is the conditional covariance matrix under normality.
- When discussing conditional means, variances and correlations, we are referring to minimum mean square error linear predictors, the variance of the prediction error, and the correlations computed from the prediction error covariance matrix.

C.2 State space forms

To facilitate estimation and prediction, all the models discussed in this article can be cast into the "state space" form. This form allows the application of the Kalman filter calculation engine and associated estimation apparatus. The filter obviates the need to tailor formulas or software to the specific extension: instead one casts the model in state–space form and uses the all–purpose filter applicable to this form. The Kalman filter equations are displayed in Anderson and Moore (1979) or Harvey (1989) and no purpose is served by presenting them here. The application of these formulas is illustrated and these illustrations are carried out using a basic personal computer and widely available spreadsheet programs.

The state space form is

$$y_t = X_t\beta + Z_t\alpha_t + G_t\epsilon_t \ , \qquad \alpha_{t+1} = W_t\beta + T_t\alpha_t + H_t\epsilon_t \ , \qquad t = 1, \ldots, n \ . \tag{A-1}$$

The following subsections display state space forms for the models proposed in this paper.

C.2.1 Development correlation model

For $t = 1, \ldots, n$ put $y_t = (\delta_{1,t-1}, \ldots, \delta_{t0})'$ and $\epsilon_t = (\epsilon_{1,t-1}, \ldots, \epsilon_{t0})'$. Thus y_t is diagonal t of the runoff triangle of the δ_{ij}'s. Put $X_t = 0$, $Z_t = I$ and $G_t = \text{diag}(h_{t-1}, \ldots, h_1, 1)$. These matrices are of dimension $t \times n$, $t \times t$ and $t \times t$ respectively. In the second equation in (A-1), put W_t as the last $t+1$ rows of the row permuted identity matrix of order n, $T_t = 0$ of dimension $(t+1) \times t$, and H_t a $(t+1) \times t$ matrix of zeros except in position (t, t) where it is $h_1\theta$. With this parametrization $\beta = (\mu_0, \ldots, \mu_{n-1})'$ and $\alpha_t = (\mu_{t-1}, \ldots, \mu_2, \mu_1 + h_1\theta\epsilon_{t-1,0}, \mu_0)'$. The basic model is the special case $\theta = 0$.

C.2.2 Accident correlation model

Define $y_t = (\delta_{1,t-1}, \ldots, \delta_{t0})'$, $\alpha_t = (\mu_{1,t-1}, \ldots, \mu_{t0})'$, $\epsilon_t = (\epsilon_{1,t-1}, \ldots, \epsilon_{t0}, \eta_{1,t-1}, \ldots, \eta_{t0})'$ and $\beta = (\mu_{10}, \ldots, \mu_{1,n-1})'$. Define X_t as a $t \times n$ matrix of zeros, Z_t as an identity matrix of order t and put G_t as a $t \times 2t$ matrix of zeros except on the first complete diagonal where it contains $h_{t-1}, \ldots, h_1, 1$. Further put W_t as a $(t+1) \times n$ matrix of zeros except in position $(1, t+1)$ where it is 1, T_t as an identity matrix of order t with a row of zeros on top and H_t as a $(t+1) \times 2t$ matrix of zeros except on the last complete diagonal where it contains $0, \lambda_{t-1}, \ldots, \lambda_0$. The basic model is the special case where each $\lambda_j = 0$. The development factor random walk model in the accident year direction is attained by letting each $h_j/\lambda_j \to \infty$.

C.2.3 Calendar correlation model

Define y_t as above, $\beta = (\mu_0, \ldots, \mu_{n-1})'$, $\alpha_t = (\mu_{t-1}, \ldots, \mu_0, \tau_t)'$ and $\epsilon_t = (\epsilon_{1,t-1}, \ldots, \epsilon_{t0}, \eta_t)'$, Further put X_t as a $t \times n$ matrix of zeros, Z_t as an identity matrix of order t with an extra final column containing $(h_{t-1}, \ldots, h_1, 1)'$ and G_t a $t \times (t+1)$ matrix of zeros except on the main diagonal where it contains $(h_{t-1}, \ldots, h_1, 1)$. Put W_t as the last $t+1$ rows of the row permuted identity matrix of order n augmented with a final column of zeros, T_t $(t+1) \times (t+1)$ matrix of zeros except in position $(t+1, t+1)$ where it is 1, and H_t a $(t+1) \times (t+1)$ matrix of zeros except in position $(t+1, t+1)$ where it is λ.

C.2.4 State space form for correlated triangles

To model the correlations between triangles, write the basic model for each of the triangles as

$$y_t^{(p)} = \alpha_t^{(p)} \qquad \alpha_{t+1}^{(p)} = W_t \beta^{(p)} + T_t \alpha_t^{(p)} + H_t^{(p)} \epsilon_t^{(p)} \ , \qquad t = 1, \ldots, n \ , \ p = 1, \ldots, r \ , \tag{A-2}$$

where there are r classes of business. Quantities with the superscript p are portfolio specific. The models corresponding to different portfolios p can be combined into a single model by stacking the $y_t^{(p)}$, $\alpha_t^{(p)}$ and $\beta^{(p)}$ for $p = 1, \ldots, r$ into y_t, α_t and β respectively yielding

$$y_t = \alpha_t \ , \qquad \alpha_{t+1} = (I_r \otimes W_t)\beta + (I_r \otimes T_t)\alpha_t + H_t \epsilon_t \ ,$$

where $\otimes$ denotes the kronecker product of two matrices, I_r is an identity matrix of order r and H_t is defined depending on the kind of correlation to be induced between the triangles as described below.

The simplest model for correlation between triangles is where $\mathrm{cov}(\epsilon_t^{(p)}, \epsilon_t^{(q)})$ is a constant diagonal matrix depending only on p and q and independent of t. In this case an estimate of the correlation between the components of $\eta_t^{(p)}$ and $\eta_t^{(q)}$ is the appropriate entry from the matrix (18). If the covariance between $\eta_t^{(p)}$ and $\eta_t^{(q)}$ is assumed to be diagonal matrix with non constant entries then the correlations depend on the development year and estimators are indicated in Table 13.

In the more detailed fitting of the model, the correlations between $\eta_t^{(p)}$ are estimated via maximum likelihood, simultaneously with all the other parameters and subject to constraints such as a constant diagonal. The appropriateness of constraints can be judged using generally applicable likelihood ratio testing methods.

C.3 Forecasting formulas

For the accident correlation model, the vector α_{n+1} as defined above can be modified by removing the top and bottom entry yielding $\alpha^*_{n+1} = (\mu_{2,n-1}, \ldots, \mu_{n1})' = W^*_n \beta + T^*_n \alpha_n + H^*_n \epsilon_n$ where the stars indicate the top and bottom rows of the indicated matrix have been removed. In turn define

$$\begin{aligned}\alpha^*_{n+2} &= W^*_{n+1}\beta + T^*_{n+1}\alpha_{n+1} + W^*_{n+1}\epsilon_{n+1} \\ &= (W_{n+1} + T_{n+1}W_n)^*\beta + (T_{n+1}T_n)^*\alpha_n + (T_{n+1}H_n, H_{n+1})^*(\epsilon'_n, \epsilon'_{n+1})' \ ,\end{aligned}$$

where the superscript * indicates the two top and two bottom rows removed from the indicated matrix. Then $\alpha^*_{n+2} = (\mu_{3,n-1}, \ldots, \mu_{n2})'$. This process is repeated, successively defining $\alpha^*_{n+3} = (\mu_{4,n-1}, \ldots, \mu_{n3})'$ through to $\alpha^*_{2n-1} = \mu_{n,n-1}$. Defining

$$\alpha_{n+1} = ({\alpha^*_{n+1}}', \ldots, {\alpha^*_{2n+1}}')' + (h_{n-1}\epsilon_{2,n-1}, \ldots, h_1\epsilon_{n1}, \ldots, h_{n-1}\epsilon_{n,n-1})'$$

The forecast is thus achieved with a single iteration of the Kalman filter. The single iteration computes the standard deviations and correlations of the future δ_{ij}. The development correlation model is dealt with in the same way yielding

$$\alpha_{n+1} = (\mu_{n-1}, \ldots, \mu_1 + h_1\theta\epsilon_{n0}, \mu_{n-1}, \ldots, \mu_2, \mu_{n-1}, \ldots, \mu_3, \ldots, \mu_{n-1})'$$

$$+ (h_{n-1}\epsilon_{2,n-1}, \ldots, h_1\epsilon_{n1}, h_{n-1}\epsilon_{3,n-1}, \ldots, h_{n-1}\epsilon_{n,n-1}) \ .$$

The above approach produces $\hat{\alpha}_{n+1}$, containing forecasts of all unobserved entries of the runoff triangle. Also produced are the variances and covariances of all the errors in the forecasts, denoted $\text{cov}(\alpha_{n+1}-\hat{\alpha}_{n+1})$. Forecasts of sums and associated mean square errors are then computed as $J\hat{\alpha}_{n+1}$ and $J\text{cov}(\hat{\alpha}_{n+1})J'$, respectively, where J is an appropriately patterned matrix of zeros and ones.

The above forecasting formulas simplify in the case of the basic model. The growth rate $g_i = \ln(c_{i,n-1}/c_{i,n-i})$ is forecast with

$$\begin{pmatrix} \hat{g}_2 \\ \vdots \\ \hat{g}_n \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \iddots & \vdots \\ 0 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \hat{\mu}_{n-1} \\ \vdots \\ \hat{\mu}_1 \end{pmatrix} , \tag{A-3}$$

where the $\hat{\mu}_j$ are the basic model estimates of the μ_j. The vector on the left is the vector of expected growths in logarithms of rows $i = 2, \ldots, n$ in the runoff triangle through to development year $n-1$, conditioning on the observed data. If the μ_j were known exactly, then the prediction errors $y_{i,n-1} - \hat{y}_{i,n-1}$, $i = 2, \ldots, n$ would be uncorrelated. However the estimates $\hat{\mu}_j$ have sampling error. The prediction errors are linear combinations of the $\hat{\mu}_j$, and hence they are correlated with covariance matrix

$$\sigma^2 \begin{pmatrix} h_{n-1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_1^2 + \cdots + h_{n-1}^2 \end{pmatrix} + \sigma^2 \begin{pmatrix} 0 & \cdots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \frac{h_1^2}{n-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_{n-1}^2 \end{pmatrix} \begin{pmatrix} 0 & \cdots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} . \tag{A-4}$$

This is the conditional covariance matrix of the logarithms of unobserved entries in column $n-1$ of the runoff triangle, conditioning on the available data and assuming the variances σ_j^2 are known. The first term is the conditional covariance matrix if the μ_j are known. The second term is addition due to the uncertainty in the μ_j and induces correlation in the forecast errors corresponding to different accident years.

C.4 Simulation of loss distributions

Practical interest centres on the distribution of (2). The distribution is not analytically tractable, since it involves sums of correlated lognormally distributed random variables. However it is easy to simulate from the distribution given log normality and the conditional means and covariances as displayed above. The simulation generates normal random vectors with the required conditional mean and covariance matrix, each component of which is exponentiated and then summed.

D Conditional models and Mack's model

The basic model (3) is a conditional model in the sense that for each i and given the past data

$$\ln c_{ij} \sim (\mu_j + \ln c_{i,j-1} \ , \ \sigma_j^2) \ . \tag{A-5}$$

If the distribution is normal then (A-5) implies c_{ij} has conditional mean and conditional coefficient of variation (Aitchison and Brown 1957, p. 87)

$$c_{i,j-1}\mathrm{e}^{\mu_j+\sigma_j^2/2} \ , \qquad \sqrt{\mathrm{e}^{\sigma_j^2} - 1} \ ,$$

respectively. Scaling $c_{i,j-1}$ thus changes the conditional distribution of c_{ij} by scaling the mean by the same amount but leaving the coefficient of variation the same.

Model (A-5) is related to a conditional the model proposed by Mack (1993) to justify the chain ladder method of reserving. With this model each row of the runoff triangle is independent and, given the past data,

$$c_{ij} \sim \left(m_j c_{i,j-1} \ , \ \alpha_j^2 c_{i,j-1}\right) \ . \tag{A-6}$$

For this setup the conditional coefficient of variation of c_{ij} is

$$\frac{\alpha_j}{m_j\sqrt{c_{i,j-1}}} \ ,$$

which decreases with $c_{i,j-1}$. Thus an upwards scaling the data leads to a decrease in the conditional coefficient of variation.

Mack (1993) uses a curious argument to justify (A-6). It is inferred from the assumed optimality of commonly used chain ladder estimates. This compares to the standard approach where an optimal estimate is derived from an assumed model. Mack's argument is as follows. With the chain ladder method the m_j are estimated as

$$\frac{\sum_{i=1}^{n-j} c_{ij}}{\sum_{i=1}^{n-j} c_{i,j-1}} = \sum_{i=1}^{n-j} w_i \frac{c_{ij}}{c_{i,j-1}} , \qquad j = 1, \ldots, n-1 , \tag{A-7}$$

where w_i is proportional to $c_{i,j-1}$ and the weights w_i add to 1. Estimator (A-7) is unbiased for m_j assuming (A-6) since the ratios $c_{ij}/c_{i,j-1}$ have expectation m_j for all $c_{i,j-1}$. Additionally it is conditionally minimum variance since the ratios $c_{ij}/c_{i,j-1}$ are uncorrelated for different i and have variance inversely proportional to $c_{i,j-1}$.

Further curious features of (A-6) and its uses by Mack (1993) are:

- The minimum variance property of the chain ladder estimate (A-7) is, for each j, conditional on the values $c_{i,j-1}$, $i = 1, \ldots, n-j$. The minimum conditional variances are

$$\sum_{i=1}^{n-j} w_i^2 \frac{\nu_j^2}{c_{i,j-1}} = \frac{\nu_j^2}{\sum_{i=1}^{n-j} c_{i,j-1}} , \qquad j = 1, \ldots, n-1 .$$

  Thus with (A-6), the chain ladder estimates (A-7) are not conventional minimum variance estimates.

- Under (A-6), the estimator $\hat{c}_i \equiv c_{i,n-i}\hat{f}_{n-i+1}\cdots\hat{f}_{n-1}$ is shown by Mack (1993) to be unbiased for $c_i \equiv c_{i,n-1}$ in the sense that they both have the same unconditional expectation. Thus the reserve estimator $\hat{r}_i \equiv \hat{c}_{i,n-1} - c_{i,n-i}$ is unbiased for the uncertain reserve $r_i \equiv c_{i,n-1} - c_{i,n-i}$. However both $\hat{c}_{i,n-1}$ and $\hat{r}_i$ are conditionally biased, conditioning on the data. This is a drawback and avoided with (A-5) and methods of the previous sections.

- As a measure of the uncertainty associated with the reserve estimate Mack (1993) uses the conditional expectation of $(\hat{r}_i - r_i)^2$, conditioning on the data. This conditional mean square error, equals, using the usual argument, the conditional variance of $\hat{c}_i - c_i$ plus the squared bias $(\hat{c}_i - \tilde{c}_i)^2$ where $\tilde{c}_i \equiv c_{i,n-i}m_{n-i+1}\cdots m_{n-1}$ is the conditional expectation of c_i given the data. Expressions for the variance and bias under (A-6) are developed by Mack (1993). However the properties of their suggested empirical implementation are unclear.

- The conditional mean square error of $\hat{r}_i - r_i$ is likely to be a very imperfect characterization of the likely error of the reserve estimator especially if its bias is severe and the distribution is far from normal.

- Suppose all entries in a runoff triangle are held constant except for the entry in the bottom left corner, which is increased. Then with (A-6) and the formula proposed by Mack (1993), the mean and variance of forecast liabilities increase in such a way that the coefficient of variation decreases. Thus in percentage terms the uncertainty of the forecast is estimated to decrease whenever the latest accident year's observation increases.

- The assumption of independence between accident years is important in the development of Mack (1993), in particular in relation to the expression for mean square error of $\sum_i(\hat{r}_i - r_i)$. It is not clear how modifications, such as correlations between accident and development years can be accommodated and the effect of such important features on the properties of the methods or formulas.

- With (A-6) there is no explicit treatment of initial conditions which appear to be regarded as fixed. Inference is thus conditional on the possibly fortuitous outcomes in development year 0.

- Conventional minimum mean square error forecasting constructs predictors such that the unconditional error variance is minimum. In the standard settings the error does not depend on the past data and hence the unconditional and conditional error variances coincide.

References

Aitchison, J. and J. A. C. Brown (1957). *The Lognormal distribution, with special reference to its uses in economics*. Cambridge, England: Cambridge University Press.

Anderson, B. D. O. and J. B. Moore (1979). *Optimal filtering*. Englewood Cliffs: Prentice-Hall.

Barnett, G. and B. Zehnwirth (2000). Best estimates for reserves. *Proceedings of the Casualty Actuarial Society LXXXVII*(166,167).

Collins, S. and G. White (2001). APRA risk margin analysis. Institute of Actuaries Australia, XIII General Insurance Seminar.

De Jong, P. and B. Zehnwirth (1983). Claims reserving, state-space models and the Kalman filter. *Journal of the Institute of Actuaries 110*, 157–181.

England, P. and R. Verrall (2002). Stochastic claims reserving in general insurance. *Journal of the Institute of Actuaries 129*, 1–76.

Goovaerts, M. and R. Redant (1999). On the distribution of IBNR-reserves. *Insurance: Mathematics and Economics 25*, 1–10.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Hertig, J. (1985). A statistical approach to IBNR–reserves in marine reinsurance. *ASTIN Bulletin 15*, 171–183.

Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin 23*(2), 213–225.

Mack, T. (1994). Measuring the variability of chain ladder reserve estimates. *Proceedings of the Casualty Actuarial Society Spring Forum*, 101–182.

Murphy, D. (1994). Unbiased loss development factors. *Proceedings of the Casualty Actuarial Society 81*, 154–222.

Renshaw, A. E. and R. J. Verrall (1994). A stochastic model underlying the chain ladder technique. Proceedings XXV ASTIN Colloquium, Cannes.

Sanders, D. E. A. (1990). *Competition Presented at a London Market Actuaries dinner.*

Taylor, G. (2000). *Loss reserving. An actuarial perspective.* Boston: Kluwer.

Verrall, R. J. (1990). Bayes and empirical Bayes estimation for the chain ladder model. *ASTIN Bulletin 20*, 217–243.

Verrall, R. J. (1991). On the estimation of reserves from loglinear models. *Insurance: Mathematics and Economics 10*, 75–80.

Verrall, R. J. (1996). Claims reserving and generalized additive models. *Insurance: Mathematics and Economics 19*, 31–43.

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

Zehnwirth, B. (1985). *Interactive Claims Reserving Forecasting System.* St Kilda, Vic. 3183, Australia: Insureware P/L.