Incorporating Expert Opinion into a Stochastic Model for the Chain-ladder Technique

R.J.Verrall
Faculty of Actuarial Science and Statistics
Cass Business School
City University

P.D.England
EMB Consultancy LLP

Abstract
To date, much effort has been directed towards the development of stochastic models that are analogous to traditional deterministic methods. In practice, however, the traditional models are often altered to incorporate expert opinion. This paper considers the use of Bayesian models to allow practitioners to apply their judgement to the development factors in the chain-ladder technique. The implementation uses MCMC methods within winBUGS. In this way, it is possible to use stochastic models to obtain predictive distributions of reserves in a much wider range of situations.

Keywords
Bayesian Statistics, Chain-ladder, Claims Reserving, Risk

Address for correspondence
Prof. R.J.Verrall,
Faculty of Actuarial Science and Statistics,
Cass Business School,
City University,
106 Bunhill Row,
London. EC1Y 8TZ
email: r.j.verrall@city.ac.uk

1. Introduction

In recent years, significant advances have been made in the development of stochastic models for claims reserving in general insurance. This has resulted in a greater understanding of the stochastic nature of forecasts of outstanding claims, and a range of models have been developed for obtaining prediction errors, and predictive distributions, for outstanding claims. England and Verrall (2002) provides a summary of the developments.

To date, much effort has been directed towards the development of stochastic models that are analogous to traditional deterministic methods. In practice, however, the traditional models are often altered to incorporate expert opinion. One aspect of this, which is considered in this paper, is when the actuary intervenes to change the parameter estimates used to forecast outstanding claims from the values actually estimated from the data. This is an important consideration, and it is desirable to be able to reproduce this process within a stochastic framework. The advantage of incorporating this practical approach within a stochastic framework is that additional quantities can also be calculated: prediction errors, prediction intervals, predictive distributions, and so on.

This paper considers the use of Bayesian models to allow the practitioner to intervene in the estimation of the development factors. In practice, the intervention usually involves changing the value used for a development factor for a particular row, or only using a portion of the data to estimate the development factors used to forecast outstanding claims. This approach is often taken in practice if there is evidence that the settlement pattern has changed, with the result that it would not be appropriate to use the same development factor for each row. We anticipate that a practitioner would be able to extend the specific cases considered in this paper to cover other situations, which, although not covered here, would also be useful and could be formulated as Bayesian models. The specific cases we consider are:

- the intervention in a development factor in a particular row, and
- the choice of how many years of data to use in the estimation.

A natural approach to use is a Bayesian method, and this paper shows how this can be done, and examines the effects of prior information on the results. This paper can be read as a sequel to England and Verrall (2002): it develops further the Bayesian models introduced in that paper and shows how the stochastic approach can include expert judgement.

A number of papers on Bayesian methods in claims reserving have appeared in recent years. Ntzoufras and Dellaportas (2002) consider a number of models in a Bayesian framework. De Alba (2002) formulates the chain-ladder technique as a conditional multinomial model, and thereby obtains a full predictive distribution of outstanding claims. De Alba uses vague prior distributions, and hence no prior knowledge is assumed about any of the parameters. In this paper, we look in more detail at the Bayesian models in order to consider how proper prior distributions may be used to incorporate prior opinion. England and Verrall (2002), Section 8, used a similar approach to obtain a full predictive distribution for the outstanding liabilities using the chain-ladder technique, again without assuming any prior knowledge. These papers showed how Bayesian methods naturally provide predictive distributions.

A different, although related, approach is bootstrapping. Bootstrapping also enables prediction errors and predictive distributions to be obtained fairly straightforwardly (using just a spreadsheet for simple models), as shown in England and Verrall (1999) and England (2002). This has resulted in bootstrapping becoming one of the most popular methods for stochastic reserving. We believe that Bayesian methods have even greater potential, because of the potential for introducing knowledge about the parameters obtained using additional information.

In a related paper, Verrall (2004) showed how the Bornhuetter-Ferguson technique can be reformulated as a Bayesian model. The similarity with this paper is that prior knowledge is incorporated into the reserve estimates and predictive distributions. However, the Bornhuetter-Ferguson technique (and the related stochastic model derived in Verrall, 2004) assumes that there is prior knowledge about the expected ultimate claims in each row. Thus, a prior distribution is specified for the parameters associated with each row – the accident, or underwriting years. Here we consider incorporating prior knowledge about the development factors – the parameters associated with each column, or development period.

Thus, this paper shows how proper (or "informative") Bayesian prior distributions can be used to combine knowledge from the business about the run-off patterns into reserve estimates and distributions using the chain-ladder technique. Of course, it is easy to obtain reserve estimates even when the estimates of the link ratios (development factors) using past data are not used. However, this suffers from the same deficiency as the chain-ladder technique: only point estimates are available, and no information concerning variability of the forecasts.

The paper is set out as follows. Section 2 summarises the stochastic models for the chain-ladder technique that are used in subsequent sections. Section 3 derives the recursive negative binomial model in a Bayesian context. Section 4 describes the extension of the Bayesian approach so that it is possible to intervene in the development factors, and section 5 contains the implementation. The computer code used in the implementation is supplied in the appendix. Section 6 concludes.

2. Stochastic Models for the Chain-ladder Technique

This section gives a brief summary of stochastic models that are related to the chain-ladder technique. A number of the models, with various positivity constraints, give the same reserve estimates as the chain-ladder technique.

Although the methods described in this paper can be applied to more general shapes of claims data, it simplifies the notation if we assume that we have a conventional triangle of data. Thus, without loss of generality, we assume that the data consist of a triangle of incremental claims:

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

The cumulative claims are defined by:

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

and the development factors of the chain-ladder technique are denoted by $\left\{\lambda_j : j = 2,\ldots,n\right\}$. The estimates of the development factor from the standard chain-ladder technique are

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

These are then applied to the latest cumulative claims in each row ($D_{i,n-i+1}$) to produce forecasts of future values of cumulative claims:

$$\hat{D}_{i,n-i+2} = D_{i,n-i+1}\hat{\lambda}_{n-i+2}$$

$$\hat{D}_{i,k} = \hat{D}_{i,k-1}\hat{\lambda}_k, \quad k = n-i+1, n-i+2,\ldots,n.$$

Previous papers have explored the connections between the chain-ladder technique and various stochastic models. Mack (1993) takes a non-parametric approach and specifies only the first 2 moments for the cumulative claims, and in that model the mean and variance of D_{ij} are $\lambda_j D_{i,j-1}$ and $\sigma_j^2 D_{i,j-1}$, respectively. Estimates of all the parameters are derived, and the properties of the model are examined. The advantages of this approach are that the parameter estimates and prediction errors can be obtained just using a spreadsheet, without having recourse to a statistical package or any complex programming. Also, the estimates can be obtained for most data, whether there are negative incremental claims or not. One disadvantage is that a predictive distribution is not available, since the distribution of the data has not been fully specified. Also, there are separate parameters in the variance that must also be estimated, separately from the estimation of the development factors.

Renshaw and Verrall (1998) used an approach based specifically on generalised linear models (McCullagh and Nelder, 1989) and examined the over-dispersed Poisson model for incremental claims. “Over-dispersion” is used to add flexibility to the models used, by adding a parameter to the variance. We use it in the context of the Poisson and negative binomial distributions, and it is defined so that the random variable, *Y*, is defined by $Y = \varphi X$, where *X* has a Poisson or negative binomial distribution. Thus, for the over-dispersed Poisson distribution, *Y* has mean and variance $\varphi\mu$ and $\varphi^2\mu$, where $\mu = E\left[X\right]$. The over-dispersed Poisson distribution may be parameterised in a number of different ways. Renshaw and Verrall (1998) used a log link function and an additive predictor, so that

$\log(E\left[C_{ij}\right]) = c + \alpha_i + \beta_j$ with the restriction that $\alpha_1 = \beta_1 = 0$.

An alternative parameterisation uses an identity link and a multiplicative structure:

$$C_{ij} \mid x, y, \varphi \sim \text{ independent over-dispersed Poisson, with mean } x_i y_j \text{, and } \sum_{k=1}^{n} y_k = 1. \quad (2.1)$$

Here $x = \{x_1, x_2, \ldots, x_n\}$ is a parameter vector relating to the rows (accident years) and $x_i = E\left[D_{in}\right]$, expected ultimate cumulative claims (up to the latest development year so far observed, *n*) for the *i*th accident year. The parameter vector $y = \{y_1, y_2, \ldots, y_n\}$ relates to the columns (development years) of the run-off triangle, and y_j can be interpreted as the proportion of ultimate claims which emerge in development year *j*. When maximum likelihood estimation is used, or Bayesian estimation with non-informative prior distributions, this model gives the same reserve estimates as the chain-ladder technique (as long as the row and column sums of incremental claims are positive). However, the connection with the chain-ladder technique is not immediately apparent from the formulation of the model. For this reason, we prefer to use the negative binomial model. This was developed by Verrall (2000), building on the over-dispersed Poisson model. Verrall (2000) showed that the same predictive distribution can be obtained from a negative binomial model (also with the inclusion of an over-dispersion parameter). This model is the basis for the Bayesian approach in this paper, and is derived in the following section.

3. The Negative Binomial Model for the Claims Triangle

The negative binomial model is a recursive approach, and hence we need to consider the data recursively, and write the likelihood accordingly. The data is received in the following order:

$$C_{11}, \begin{pmatrix} C_{1,2} \\ C_{2,1} \end{pmatrix}, \begin{pmatrix} C_{13} \\ C_{22} \\ C_{31} \end{pmatrix}, \ldots$$

The data vector at time *t* can be defined by

$$\underline{C}_t = \begin{pmatrix} C_{1,t} \\ C_{2,t-1} \\ \vdots \\ C_{t,1} \end{pmatrix}.$$

The likelihood can be written in a recursive form as:

$$f\left(\underline{C}_1,\underline{C}_2,\ldots,\underline{C}_n \mid x,y,\varphi\right)$$
$$= f\left(\underline{C}_1 \mid x,y,\varphi\right) f\left(\underline{C}_2 \mid \underline{C}_1,x,y,\varphi\right)\ldots f\left(\underline{C}_n \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{n-1},x,y,\varphi\right)$$

In this paper, we use a Bayesian approach, and obtain the posterior distribution of the parameters using Bayes theorem:

$$f\left(x,y \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_n,\varphi\right) \propto f\left(\underline{C}_1,\underline{C}_2,\ldots,\underline{C}_n \mid x,y,\varphi\right) f\left(x,y\right)$$

Note that the over-dispersion parameter, φ, is treated as a nuisance parameter and a plug-in estimate is used. A full Bayesian approach would give this parameter a prior distribution and include it in the prior-posterior analysis. The approach taken here is simpler and is similar to the approach which is often used in a classical analysis.

By considering the likelihood in recursive form, we can also find the posterior distribution of the parameters recursively as follows:

$$f\left(x,y \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_t,\varphi\right) \propto f\left(\underline{C}_t \mid x,y,\varphi\right) f\left(x,y \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1}\right)$$

The aim of claims reserving is to derive the one-step-ahead predictive distribution of future claims:

$$f\left(\underline{C}_t \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1},\varphi\right). \qquad (3.1)$$

It is straightforward to extend this to *n*-steps-ahead prediction. Leaving aside the estimation of the dispersion parameter, φ, this predictive distribution can be derived for the over-dispersed Poisson model by first deriving the posterior distribution of the parameters, and then integrating these out:

$$f\left(\underline{C}_t \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1},\varphi\right) = \iint f\left(\underline{C}_t \mid x,y,\varphi\right) f\left(x,y \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1},\varphi\right) dxdy \qquad (3.2)$$

This predictive distribution can be written as

$$f\left(\underline{C}_t \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1},\varphi\right) =$$
$$= \int\left(\int f\left(\underline{C}_t \mid x,y,\varphi\right) f\left(x \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1},y,\varphi\right) dx\right) f\left(y \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1},\varphi\right) dy \qquad (3.3)$$

We consider first $\int f\left(\underline{C}_t \mid x,y,\varphi\right) f\left(x \mid \underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1},y,\varphi\right) dx$. This is the distribution of $\underline{C}_t$, using the posterior distribution for the row parameters, *x*, but conditioning throughout on the column parameters. Note that the distribution of the parameters, $x_1,x_2,\ldots,x_n$ is obtained conditional on *y* and φ, using only the data in the appropriate row. Also, there is one data point from each row in $\underline{C}_t$, and hence we consider

$\int f\left(C_{i,t-i+1} \mid x, y, \varphi\right) f\left(x_i \mid \underline{C}_1, \underline{C}_2, \ldots, \underline{C}_{t-1}, y, \varphi\right) dx_i$. This distribution was also considered in detail by Verrall (2000). The following theorem derives the form of this distribution when non-informative prior distributions are used for the row parameters.

Theorem

If non-informative gamma prior distributions are used for the row parameters, $\int f\left(C_{i,t-i+1} \mid x, y, \varphi\right) f\left(x_i \mid \underline{C}_1, \underline{C}_2, \ldots, \underline{C}_{t-1}, y, \varphi\right) dx_i$ is an over-dispersed negative binomial distribution with mean and variance

$$\left(\lambda_{t-i+1} - 1\right) D_{i,t-i} \qquad \text{and} \quad \varphi \lambda_{t-i+1}\left(\lambda_{t-i+1} - 1\right) D_{i,t-i}, \quad \text{respectively,}$$

where $\lambda_j = \dfrac{\sum_{k=1}^{j} y_k}{\sum_{k=1}^{j-1} y_k}$, $j = 2, 3, \ldots$

Proof

Note that if Y has an over-dispersed Poisson distribution, with mean and variance $\varphi\mu$ and $\varphi^2\mu$, then $X = \dfrac{Y}{\varphi}$ has a Poisson distribution with mean $\mu = E\left[X\right]$. Hence $f\left(\dfrac{C_{i,t-i+1}}{\varphi} \mid x, y, \varphi\right)$ is a Poisson distribution with mean $\dfrac{x_i y_{t-i+1}}{\varphi}$.

Consider first $f\left(x_i \mid \underline{C}_1, \underline{C}_2, \ldots, \underline{C}_{t-1}, y, \varphi\right) \propto f\left(C_{i,1}, C_{i,2}, \ldots, C_{i,t-i} \mid x_i, y, \varphi\right) f\left(x_i\right)$

$$\propto \left(\frac{x_i y_1}{\varphi}\right)^{\frac{C_{i,1}}{\varphi}} \left(\frac{x_i y_2}{\varphi}\right)^{\frac{C_{i,2}}{\varphi}} \cdots \left(\frac{x_i y_{t-i}}{\varphi}\right)^{\frac{C_{i,t-i}}{\varphi}} \exp\left(-\frac{x_i\left(y_1 + y_2 + \ldots + y_{t-i}\right)}{\varphi}\right) x_i^{-1}$$

$$\propto x_i^{\frac{D_{i,t-i}}{\varphi} - 1} \exp\left(-\frac{x_i \sum_{k=1}^{t-i} y_k}{\varphi}\right).$$

Hence $x_i \mid \underline{C}_1, \underline{C}_2, \ldots, \underline{C}_{t-1}, y, \varphi$ has a gamma distribution, with parameters $\dfrac{D_{i,t-i}}{\varphi}$ and $\dfrac{\sum_{k=1}^{t-i} y_k}{\varphi}$.

Hence $\int f\left(\frac{C_{i,t-i+1}}{\varphi} \mid x, y, \varphi\right) f\left(x \mid \underline{C}_1, \underline{C}_2, \ldots, \underline{C}_{t-1}, y, \varphi\right) dx$

$$= \int_0^\infty \frac{1}{\Gamma\left(\frac{C_{i,t-i+1}}{\varphi}+1\right)} \left(\frac{x_i y_{t-i+1}}{\varphi}\right)^{C_{i,t-i+1}/\varphi} \exp\left(-\frac{x_i y_{t-i+1}}{\varphi}\right) \frac{\left(\frac{\sum_{k=1}^{t-i} y_k}{\varphi}\right)^{\frac{D_{i,t-i}}{\varphi}}}{\Gamma\left(\frac{D_{i,t-i}}{\varphi}\right)} x_i^{\frac{D_{i,t-i}}{\varphi}-1} \exp\left(-\frac{x_i \sum_{k=1}^{t-i} y_k}{\varphi}\right) dx_i$$

$$= \frac{\left(\frac{y_{t-i+1}}{\varphi}\right)^{C_{i,t-i+1}/\varphi} \left(\frac{\sum_{k=1}^{t-i} y_k}{\varphi}\right)^{\frac{D_{i,t-i}}{\varphi}}}{\Gamma\left(\frac{C_{i,t-i+1}}{\varphi}+1\right)\Gamma\left(\frac{D_{i,t-i}}{\varphi}\right)} \int_0^\infty x_i^{\frac{D_{i,t-i+1}}{\varphi}-1} \exp\left(-\frac{x_i \sum_{k=1}^{t-i+1} y_k}{\varphi}\right) dx_i$$

$$= \frac{\left(\frac{y_{t-i+1}}{\varphi}\right)^{C_{i,t-i+1}/\varphi} \left(\frac{\sum_{k=1}^{t-i} y_k}{\varphi}\right)^{\frac{D_{i,t-i}}{\varphi}}}{\Gamma\left(\frac{C_{i,t-i+1}}{\varphi}+1\right)\Gamma\left(\frac{D_{i,t-i}}{\varphi}\right)} \frac{\Gamma\left(\frac{D_{i,t-i+1}}{\varphi}\right)}{\left(\frac{\sum_{k=1}^{t-i+1} y_k}{\varphi}\right)^{\frac{D_{i,t-i+1}}{\varphi}}}$$

$$= \frac{\Gamma\left(\frac{D_{i,t-i+1}}{\varphi}\right)}{\Gamma\left(\frac{C_{i,t-i+1}}{\varphi}+1\right)\Gamma\left(\frac{D_{i,t-i}}{\varphi}\right)} \left(\frac{y_{t-i+1}}{\sum_{k=1}^{t-i+1} y_k}\right)^{C_{i,t-i+1}/\varphi} \left(\frac{\sum_{k=1}^{t-i} y_k}{\sum_{k=1}^{t-i+1} y_k}\right)^{\frac{D_{i,t-i}}{\varphi}}.$$

This is a negative binomial distribution, with parameters $\frac{D_{i,t-i}}{\varphi}$ and $\frac{\sum_{k=1}^{t-i} y_k}{\sum_{k=1}^{t-i+1} y_k}$. Noting that $\lambda_j = \frac{\sum_{k=1}^{j} y_k}{\sum_{k=1}^{j-1} y_k}$, the parameters of this distribution are $\frac{D_{i,t-i}}{\varphi}$ and $\frac{1}{\lambda_{t-i+1}}$, and the mean and variance are

$$\frac{D_{i,t-i}}{\varphi}\frac{\left(1-\frac{1}{\lambda_{t-i+1}}\right)}{\frac{1}{\lambda_{t-i+1}}}=\frac{D_{i,t-i}}{\varphi}\left(\lambda_{t-i+1}-1\right)$$

and $$\frac{D_{i,t-i}}{\varphi}\frac{\left(1-\frac{1}{\lambda_{t-i+1}}\right)}{\left(\frac{1}{\lambda_{t-i+1}}\right)^2}=\frac{D_{i,t-i}}{\varphi}\lambda_{t-i+1}\left(\lambda_{t-i+1}-1\right)$$ respectively.

Hence, $\int f\left(C_{i,t-i+1}\,|\,x,y,\varphi\right)f\left(x\,|\,\underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1},y,\varphi\right)dx$ is an over-dispersed negative binomial distribution with mean and variance

$$\left(\lambda_{t-i+1}-1\right)D_{i,t-i} \qquad \text{and} \quad \varphi\lambda_{t-i+1}\left(\lambda_{t-i+1}-1\right)D_{i,t-i}, \quad \text{respectively,}$$

which completes the proof of the theorem.

Notes:

1. In the proof of the theorem, we have derived the posterior distribution of $x_i\,|\,\underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1},y,\varphi$. It should be noted that only the data in row *i* is needed to derive this distribution, since the distribution of the data in the other rows does not depend on x_i.

2. The relationship between the parameters $y_1,y_2,\ldots,y_n$ and the parameters $\lambda_2,\lambda_3,\ldots,\lambda_n$ was explored in Verrall (1991). Note that there are the $n-1$ free parameters in each set, since $\sum_{k=1}^{n}y_k=1$. The relationship between the parameters is as follows:

$$\lambda_j=\frac{\sum_{k=1}^{j}y_k}{\sum_{k=1}^{j-1}y_k}$$

$$y_n=1-\frac{1}{\lambda_n}\,, \quad y_j=\left(1-\frac{1}{\lambda_j}\right)\frac{1}{\prod_{k=1}^{n-j}\lambda_{j+k}} \text{ for } j=2,\ldots,n-1 \text{ and } y_1=\frac{1}{\prod_{k=2}^{n}\lambda_k}$$

(which preserves the constraint $\sum_{k=1}^{n}y_k=1$).

3. The mean of the distribution of $\int f\left(C_{i,t-i+1}\,|\,x,y,\varphi\right)f\left(x\,|\,\underline{C}_1,\underline{C}_2,\ldots,\underline{C}_{t-1},y,\varphi\right)dx$ is clearly closely related to the chain-ladder technique. In fact, by adding the previous cumulative claims, an equivalent model for D_{ij} has an over-dispersed negative binomial distribution, with mean and variance $\lambda_j D_{i,j-1}$ and $\phi\lambda_j\left(\lambda_j-1\right)D_{i,j-1}$,

respectively. Here the connection with the chain-ladder technique is immediately apparent in the mean.

This theorem can be used with (3.3) in order to show that the predictive distribution of future claims can be obtained by supplying prior distributions for the column parameters. If non-informative, improper prior distributions are also used for the column parameters, then the forecasts will be the same as those from the chain-ladder technique. Again, the same positivity constraints apply as for the over-dispersed Poisson model. The necessity for the column sums to be positive is immediately apparent from the form of the variance: a negative sum would result in a development factor less than 1 ($\lambda_j < 1$), and hence a negative variance.

It is important to note that exactly the same predictive distribution can be obtained from either the Poisson or negative binomial models, as long as non-informative prior distributions are used. Verrall (2000) also argued that the model could be specified either for incremental or cumulative claims, with no difference in the results. The negative binomial model has the advantage that the form of the mean is exactly the same as that which naturally arises from the chain-ladder technique. This paper only considers proper prior distributions when modelling the column parameters, not the row parameters, which are always given non-informative prior distributions, and the result of the theorem, above, used: we always use the recursive negative binomial model. Giving the row parameters proper prior distributions is equivalent to the Bornhuetter-Ferguson method (see Verrall, 2004): in this paper, we are more concerned with the run-off shape.

The purpose of this paper is to show how expert opinion, from sources other than the specific data set under consideration, can be incorporated into the predictive distributions of the reserves for the models that, subject to the positivity constraints, give the same reserves as the chain-ladder technique. The next section specifies a Bayesian approach to the negative binomial model.

4. Bayesian Models

Section 3 has shown that by using non-informative prior distributions for the row parameters, we can then use the recursive negative binomial model for the data to estimate the column parameters. The column parameters can be defined in a number of ways, including those used in sections 2 and 3, $y_1, y_2, \ldots, y_n$ and $\lambda_2, \lambda_3, \ldots, \lambda_n$. Thus, it is possible to define prior distributions for either set of parameters. Since it is usual to think in terms of the development factors, in this section we define some prior distributions for $\lambda_2, \lambda_3, \ldots, \lambda_n$, and consider the predictive distribution of future claims using the equivalent of (3.3). In other words, we use the distribution of $\int f\left(C_{i,t-i+1} \mid x, y, \varphi\right) f\left(x \mid \underline{C}_1, \underline{C}_2, \ldots, \underline{C}_{t-1}, y, \varphi\right) dx$, which is an over-dispersed negative binomial distribution with mean $\left(\lambda_{t-i+1} - 1\right) D_{i,t-i}$ and variance $\varphi \lambda_{t-i+1}\left(\lambda_{t-i+1} - 1\right) D_{i,t-i}$, and apply appropriate prior distributions for $\lambda_2, \lambda_3, \ldots, \lambda_n$.

In this way, Bayesian models are specified which allows the practitioner to intervene in the estimation of the development factors. In practice, intervention usually means simply changing a development factor for a particular row. Thus, although the same development parameters (and hence run-off pattern) is usually applied for all accident years, if there is some exogenous information that indicates that this is not appropriate, the practitioner may decide to apply a different development factor (or set of factors) in some, or all, rows. The reasons for intervening in this way could be that there is information that the settlement pattern has changed, for example. This would mean that it would not be appropriate to use the same development factor for each row. Thus, the first step is to generalise the negative binomial model so that incremental claims, C_{ij}, have an over-dispersed negative binomial distribution, with mean and variance

$$\left(\lambda_{i,j}-1\right)D_{i,j-1} \quad \text{and} \quad \varphi\lambda_{i,j}\left(\lambda_{i,j}-1\right)D_{i,j-1}, \quad \text{respectively.}$$

That is, the development factors depend on both origin year *i* and development year *j*. As it stands, this model is obviously over-parameterised but it is not intended that it should be used in this form without, for example, prior assumptions on some of the parameters. Also, the development factors could be set equal to each other for most rows, with only those where the development pattern is thought to have changed being given a separate value. For example, it could be postulated that

$$\lambda_{1,2}=\lambda_{2,2}=\lambda_{3,2}=\lambda_2$$

but $\lambda_{4,2}$ has a different value (with a prior distribution motivated by external information).
An equivalent formulation (which is more in line with the statistical packages) is C_{ij} has an over-dispersed negative binomial distribution, with parameters

$$D_{i,j-1} \quad \text{and} \quad p_{i,j}=\frac{1}{\lambda_{i,j}}.$$

Additionally, the over-dispersion parameter can be more easily dealt with by noting that

$$\frac{C_{ij}}{\varphi} \sim \text{negative binomial, parameters } \frac{D_{i,j-1}}{\varphi} \quad \text{and} \quad \frac{1}{\lambda_{i,j}}.$$

Thus, the model can be treated by either using the original data and a "quasi-likelihood" to incorporate the over-dispersion, or else the original data can be divided by the over-dispersion parameter and the exact likelihood used. The quasi-likelihood approach cannot be used for prediction purposes, and the possible shortcomings of the second method are discussed in section 8 of England and Verrall (2002).

The next stage is to define prior distributions for the parameters, $p_{i,j}$ or (equivalently) $\lambda_{i,j}$. It is possible to set some of these equal to each other (within each column), as above. To revert to the standard chain-ladder model, one would set

$$\lambda_{i,j} = \lambda_j \qquad \text{for} \qquad i = 1,2,\ldots,n-j+1;\ j = 2,3,\ldots,n$$

and define vague prior distributions for λ_j $\left(j = 2,3,\ldots,n\right)$. This was the approach taken in Section 8 of England and Verrall (2002) and is very similar to that taken by de Alba (2002). This can provide a very straightforward method to obtain prediction errors and predictive distributions.

However, we show here how it is also possible to use a proper Bayes prior to encompass the expert opinions about the development parameters. There is a number of ways in which this could be used, and we describe some possibilities here. It is expected that a practitioner would be able to extend these to cover situations which, although not specifically covered here, would also be useful. The cases considered here are:

(1) the intervention in a development factor in a particular row, and
(2) the choice of how many years of data to use in the estimation.

The reasons for intervening in these ways could be that there is information that the settlement pattern has changed, making it inappropriate to use the same development factor for each row.

For the first case, what may happen in practice is that a development factor in a particular row is simply changed. Thus, although the same development parameters (and hence run-off pattern) is usually applied for all accident years, if there is some exogenous information that indicates that this is not appropriate, the practitioner may decide to apply a different development factor (or set of factors) in some, or all, rows.

In the second case, it is common to look at, say, 3-year volume weighted averages in calculating the development factors, rather than using all the available data in the triangle. Bayesian methods make this particularly easy to do, and are flexible enough to allow many possibilities

We use a 10×10 triangle in the illustrations in the following section. For the first of the two cases described above, we suppose that there is information that implies that the second development factor (from column 2 to column 3) should be given the value 1.5, for rows 8, 9,and 10, and that there is no indication that the other parameters should be treated differently from the standard chain-ladder technique. An appropriate way to treat this would be to specify:

$$\lambda_{i,j} = \lambda_j \qquad \text{for} \qquad i = 1,2,\ldots,n-j+1;\ j = 1,3,\ldots,n$$
$$\lambda_{i,2} = \lambda_2 \qquad \text{for} \qquad i = 1,2,\ldots,7$$
$$\lambda_{8,2} = \lambda_{9,2} = \lambda_{10,2}$$

The means and variances of the prior distributions of the parameters are chosen to reflect the expert opinion:

$\lambda_{8,2}$ has a prior distribution with mean 1.5 and variance *W*, where *W* is set to reflect the strength of the prior information

$\lambda_j \; (j = 2, 3, \ldots, n)$ have prior distributions with large variances.

For the second case, we divide the data into two parts using the prior distributions. To do this, we set

$$\lambda_{i,j} = \lambda_j \text{ for } i = n-j-3, n-j-2, n-j-1, n-j, n-j+1$$
$$\lambda_{i,j} = \lambda_j^* \text{ for } i = 1, 2, \ldots, n-j-4$$

and give both λ_j and λ_j^* prior distributions with large variances so that they are estimated from the data. Adjustments to the specification are made in the later development years, where there are less than 5 rows. For these columns there is just one development parameter, λ_j.

The specific form of the prior distribution (gamma, log-normal, etc) is usually chosen so that the numerical procedures in winBUGS work as well as possible. It is not appropriate to devote a large part of this paper on matters related to the implementation of winBUGS: the reader is invited to consult the references given, such as Congdon (2001, 2003) or Skollnik (2003) which provide many more details.

These models are used as illustrations of the possibilities for incorporating expert knowledge about the development pattern, but it is (of course) possible to specify many other prior distributions. Before looking at the uses of the Bayesian models, we should discuss the nuisance parameter φ. In a full Bayesian analysis, we should also give this a prior distribution and estimate it along with the other parameters. However, for ease of implementation we instead use a plug-in estimate, in line with the approach usually taken in classical methods (in England and Verrall, 2002, for example). The value used is that obtained from the application of the over-dispersed Poisson model, estimating the row and column parameters using maximum likelihood estimation (it is possible to use S-Plus or excel for this). The value in each case is found by calculating the sum of the squares of the deviance residuals and dividing by the degrees of freedom.

5. Illustration

In order to implement the Bayesian models, we make use of the software package winBUGS (Spiegelhalter et al, 1996). This software package is freely available from http://www.mrc-bsu.cam.ac.uk/bugs. In the Appendix, we provide the programme for carrying out the Bayesian analysis for the models described in this paper. An excellent reference for Bayesian modelling using MCMC methods in the context of actuarial modelling is Skollnik (2001). The software uses a simulation approach to obtain the posterior distributions for the parameters and predictive distributions for future observations, which are often not easy to find analytically. In other words, the posterior distribution is not derived analytically; instead Monte Carlo simulation is used.

We now consider using a prior distribution to intervene in some of the parameters of the chain-ladder model, instead of using prior distributions with large variances which just reproduce the chain-ladder estimates. The data set we consider in this section is taken from Taylor and Ashe (1983), and has also been used in a number of previous papers on stochastic reserving. The incremental claims data is given in table 1, together with the chain-ladder results for comparison purposes.

Table 1. Data from Taylor and Ashe (1983) with the chain-ladder estimates

357,848	766,940	610,542	482,940	527,326	574,398	146,342	139,950	227,229	67,948
352,118	884,021	933,894	1,183,289	445,745	320,996	527,804	266,172	425,046	
290,507	1,001,799	926,219	1,016,654	750,816	146,923	495,992	280,405		
310,608	1,108,250	776,189	1,562,400	272,482	352,053	206,286			
443,160	693,190	991,983	769,488	504,851	470,639				
396,132	937,085	847,498	805,037	705,960					
440,832	847,631	1,131,398	1,063,269						
359,480	1,061,648	1,443,370							
376,686	986,608								
344,014									

Chain-ladder development factors:

3.4906	1.7473	1.4574	1.1739	1.1038	1.0863	1.0539	1.0766	1.0177

Chain-ladder reserve estimates:

2	94,634
3	469,511
4	709,638
5	984,889
6	1,419,459
7	2,177,641
8	3,920,301
9	4,278,972
10	4,625,811
Overall	18,680,856

We consider 2 cases, as discussed in section 4. For the first case, we assume that there is information that implies that the second development factor (from column 2 to column 3) should be given the value 1.5, for rows 7, 8, 9,and 10, and that there is no indication that the other parameters should be treated differently from the standard chain-ladder technique. In order to implement this, the parameter for the second development factor for rows 7-10 is given a prior distribution with mean 1.5. We then look at two different choices for the prior variance for this parameter. Using a large variance means that the parameter is estimated separately from the other rows, but using the data without letting the prior mean influence it too greatly. We then use a standard deviation of 0.1 for the prior distribution, so that the prior mean has a greater influence. Since MCMC methods use a simulation approach, rather than deriving the posterior distributions analytically, there is some flexibility with the prior distributions that can be used. In particular, there is no need to use conjugate prior distributions or to restrict attention to distributions which make the derivation of posterior distributions tractable. In the examples given here, we define the prior distributions so that the development factors are automatically greater than one. Two ways to do this, both of which give reasonable results, are to define

$$\lambda_j = g_j + 1$$

and give g_j either a log-normal or a gamma distribution. For the first case illustrated here, we use gamma prior distributions for g_j with parameters 0.005 and 0.01, and 25 and 50. Hence, the mean of g_j is 0.5 and the mean of λ_j is 1.5, and the variance of g_j (and of λ_j) is either 50 or 0.01.

We consider first the estimate of the second development factor. The chain ladder estimate is 1.7473 and the individual development factors for the triangle are shown in table 2. The rows for the second development factor that are modelled separately are shown in italics. The estimate using the Bayesian models is 1.68 for rows 1-6. When a large variance is used for the prior distribution of the development factor for rows 7-10, the estimate using the Bayesian model is 1.971. With the smaller variance for this prior distribution, the estimate is 1.673, and has been drawn down towards the prior mean of 1.5. This clearly shows how the prior distributions can be used to influence the parameter estimates.

Table 2. Individual development factors

3.143	1.543	1.278	1.238	1.209	1.044	1.040	1.063	1.018
3.511	1.755	1.545	1.133	1.084	1.128	1.057	1.086	
4.448	1.717	1.458	1.232	1.037	1.120	1.061		
4.568	1.547	1.712	1.073	1.087	1.047			
2.564	1.873	1.362	1.174	1.138				
3.366	1.636	1.369	1.236					
2.923	*1.878*	1.439						
3.953	*2.016*							
3.619								

The effect on the reserve estimates is shown in table 3, which compares the reserves and prediction errors for the two cases outlined above with the results for the chain-ladder model. The chain-ladder figures are slightly different from those given in table 1 because winBUGS was used to produce the figures in table 3, and this is a simulation method.

Table 3. Reserves and prediction errors for the chain-ladder and Bayesian models

	Chain-ladder		Large variance		Small variance	
	Expected Reserve	Prediction Error (%)	Expected Reserve	Prediction Error (%)	Expected Reserve	Prediction Error (%)
Year 2	97,910	115%	96,890	117%	95,140	114%
Year 3	471,200	46%	478,100	46%	473,500	46%
Year 4	711,100	38%	720,000	36%	716,300	36%
Year 5	989,200	31%	1,001,000	31%	986,800	31%
Year 6	1,424,000	27%	1,435,000	26%	1,427,000	26%
Year 7	2,187,000	23%	2,200,000	22%	2,183,000	23%
Year 8	3,930,000	20%	3,960,000	20%	3,933,000	20%
Year 9	4,307,000	24%	5,026,000	26%	4,076,000	25%
Year 10	4,674,000	43%	5,348,000	44%	4,472,000	43%
Overall	18,790,000	16%	20,260,000	17%	18,360,000	16%

It can be seen that when a large variance is used for the second development factor in the last two rows, the result is that the reserves for those rows increase since the estimate of the second development factor for those rows is larger. When the strong prior information is used, the estimate is decreased towards the prior mean, and the reserves come back down again. It is interesting to note that, in this case, the intervention has not had a marked effect on the prediction errors (in percentage terms). Other prior distributions could have a greater effect on the percentage prediction error.

The second case we consider is when we use only the most recent data for the estimation of each development factor. To do this, we use separate prior distributions for the parameters in the last three diagonals, by defining

$$\lambda_{i,j} = \lambda_j^* \text{ for } j = 2,3,\ldots,7; i = 1,2,\ldots,7-j+1$$
$$\lambda_{i,j} = \lambda_j \text{ otherwise}$$

where λ_j^* and λ_j $(j = 2,3,\ldots,n)$ have log-normal distributions with large variances.

For the last 3 development factors, all the data is used because there is no more than 3 years for each. For the other development factors, only the 3 most recent years are used. The estimates of the development factors are shown in table 4. The estimates of the first development factor are not affected by the change in the model (the small differences could be due to simulation error or the changes elsewhere). For the other development factors, the estimates can be seen to be affected by the model assumptions.

Table 4. Development factors using 3 most recent years data separately

Incremental Factors									
3 Yr Wtd Ave	3.579	1.852	1.393	1.155	1.085	1.099	1.054	1.076	1.018
All rows	3.527	1.751	1.46	1.175	1.104	1.087	1.054	1.076	1.018
Cumulative Factors									
3 Yr Wtd Ave	14.681	4.102	2.215	1.590	1.377	1.269	1.155	1.095	1.018
All rows	14.678	4.162	2.377	1.628	1.385	1.255	1.155	1.095	1.018

The effect of using only the latest 3 years in the estimation of the development factors in the forecasting of outstanding claims can be seen in table 5.

Table 5 Reserve estimates using 3 most recent years data

	Chain-ladder		Bayesian Model	
	Expected Reserve	Prediction Error (%)	Expected Reserve	Prediction Error (%)
Year 2	97,910	115%	96,910	121%
Year 3	471,200	46%	468,200	48%
Year 4	711,100	38%	708,100	38%
Year 5	989,200	31%	1,032,000	31%
Year 6	1,424,000	27%	1,382,000	28%
Year 7	2,187,000	23%	2,058,000	25%
Year 8	3,930,000	20%	3,481,000	23%
Year 9	4,307,000	24%	4,269,000	28%
Year 10	4,674,000	43%	4,682,000	50%
Overall	18,790,000	16%	18,180,000	19%

In this case, the effect on the reserves is not particularly strong, since there is little evidence from the data of a changing payment pattern. Even so, the movements in the expected reserves compared to the stochastic chain ladder model can be explained by comparing the cumulative development factors for the pure chain ladder model and the 3-year volume weighted average model at various stages of development. For example, the expected reserves are lower in years 6,7,8 and 9, but higher in year 10, which is consistent with a higher cumulative development factor at the first development period, but lower cumulative development factors at development periods 2,3,4 and 5.

The prediction errors have increased for most years, although the effect is not strong.

The importance of the Bayesian method is the ability to assess the uncertainty of the outcome using different prior assumptions in a way that is analogous to traditional models for claims reserving. These 2 cases illustrate the possibilities available, and both depart from the straightforward chain-ladder technique.

6. Conclusions

This paper has shown how expert opinion, separate from the reserving data, can be incorporated into the prediction intervals for a stochastic reserving model. The advantages of a stochastic approach are that statistics associated with the predictive distribution are also available, rather than just a point estimate. In fact, with Bayesian methods, it is possible to produce a full predictive distribution of all cash flows, rather than just the first two moments, which is essential in dynamic financial analysis (DFA).

Strictly, the use of the over-dispersed negative binomial distribution implies that the data consist of multiples of the dispersion parameter. In practice, this can be over-looked for the observed past data, although it is obviously more problematic when simulating future observations: the properties of the predictive distributions will be consistent with the assumptions made, but the distribution of a single cell in the run-off triangle will appear unrealistic. When forecasting in practice, it is more convenient to use an alternative forecast distribution, such as the Gamma distribution, parameterised such that the mean and variance equal their theoretical values. In making this pragmatic compromise, it could be argued that we are introducing a conceptual inconsistency.

Alternative modelling distributions could also be used, such as the normal distribution As shown in England and Verrall (2002), use of the normal distribution gives the same results for the mean and prediction error as Mack’s approach (Mack, 1993). When formulated as a Bayesian model, a full predictive distribution for Mack’s model can also be obtained, in addition to the prediction error. With Mack’s model, each development period has its own variance parameter. In the same way, the assumption used in this paper of a constant scale parameter for the over-dispersed negative binomial model can be relaxed to allow the scale parameter to vary by development period.

In this paper, we have concentrated on two common situations where expert opinion is used. However, the same approach could also be taken for other modelling methods, for example where curves are fitted to allow extrapolation for the estimation of tail factors. We acknowledge that methods based on the chain-ladder framework are commonly used, and by staying within that framework, we hope that actuaries will appreciate the suggestions made in this paper, and experiment with the programmes supplied.

References

Bornhuetter, R.L. and Ferguson, R.E. (1972) *The Actuary and IBNR* Proc. Cas. Act. Soc., LIX, pp 181-195

Congdon, P. (2001) *Bayesian Statistical Modelling*. Wiley.

Congdon, P. (2003) *Applied Bayesian Modelling*. Wiley.

De Alba, E. (2002) *Bayesian Estimation of Outstanding Claims Reserves* NAAJ, Vol. 6, No 4, pp 1-20.

England, P.D. (2002). Addendum to “Analytic and bootstrap estimates of prediction errors in claims reserving”. *Insurance: Mathematics and Economics* Vol. 31 pp 461-466.

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

England, P.D. and Verrall, R.J. (2000) *A Flexible Framework for Stochastic Claims Reserving*. Casualty Actuarial Society, Vol. 88.

England, P.D. and Verrall, R.J. (2002) *Stochastic Claims Reserving in General Insurance (with discussion)*. British Actuarial Journal, 8, pp 443-544

Ntzoufras, I. and Dellaportas, P. (2002) *Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty* NAAJ, Vol 6, No 1, pp 113-128

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

Mack, T. (2000) *Credible Claims Reserves: The Benktander Method.* ASTIN Bulletin, Vol. 30, pp 333-347

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

Scollnik, D.P.M. (2001). *Actuarial Modeling with MCMC and BUGS*. North American Actuarial Journal, **5** (2), 96-124.

Spiegelhalter, D.J., Thomas, A., Best, N.G. and Gilks, W.R. (1996) *BUGS 0.5: Bayesian Inference using Gibbs Sampling*, MRC Biostatistics Unit, Cambridge, UK.

Taylor, G.C. (2000) *Loss Reserving: An Actuarial Perspective*, Kluwer.

Taylor, G.C. and Ashe, F.R. (1983) *Second moments of estimates of outstanding claims*. Journal of Econometrics 23, pp37-61.

Verrall, R.J. (1991) *Chain Ladder and Maximum Likelihood*. Journal of the Institute of Actuaries, 1991, **118**, Part III, No.471, pp 489-499.

Verrall, R.J. (2000) *An Investigation into Stochastic Claims Reserving Models and the Chain-Ladder Technique*. Insurance: Mathematics and Economics, Volume 26, pp 91-99.

Appendix

```
model
{
#Model for data:
	for( i in 1 : 45 ) {
		Z[i] <- Y[i]/(scale*1000)
		pC[i]<-D[i]/(scale*1000)
		C[i]<-Z[i]+pC[i]
	 zeros[i]<- 0
		zeros[i] ~ dpois(phi[i])
		phi[i]<-(loggam(Z[i]+1)+loggam(pC[i])-loggam(C[i])-
pC[i]*log(p1[row[i],col[i]])-Z[i]*log(1-p1[row[i],col[i]]))
		}

							DD[3]<-DD[2]+Y[47]
	for( i in 1 : 2 ) {DD[4+i]<-DD[4+i-1]+Y[49+i-1]}
	for( i in 1 : 3 ) {DD[7+i]<-DD[7+i-1]+Y[52+i-1]}
	for( i in 1 : 4 ) {DD[11+i]<-DD[11+i-1]+Y[56+i-1]}
	for( i in 1 : 5 ) {DD[16+i]<-DD[16+i-1]+Y[61+i-1]}
	for( i in 1 : 6 ) {DD[22+i]<-DD[22+i-1]+Y[67+i-1]}
	for( i in 1 : 7 ) {DD[29+i]<-DD[29+i-1]+Y[74+i-1]}
	for( i in 1 : 8 ) {DD[37+i]<-DD[37+i-1]+Y[82+i-1]}

#Model for future observations
	for( i in 46 : 90 ) {
		a1[i]<- max(0.01,(1-p1[row[i],col[i]])*DD[i-45]/(1000*scale))
			b1[i]<- p1[row[i],col[i]]/(1000*scale)
			Z[i]~dgamma(a1[i],b1[i])
			Y[i]<-Z[i]
					}




# Set up the parameters of the negative binomial model.
	for (k in 1:9) {
			p[k]<-1/lambda[k]
			lambda[k]<-exp(g[k])+1
			g[k]~dnorm(0.5,1.0E-6)
		}
# Choose one of the following (1,2 or 3) and delete the "#" at the start of each line
before running.

# 1. Vague Priors: Chain-ladder model
#	scale <- 52.8615
#	for (j in 1:9) {
#	for (i in 1:10) {p1[i,j]<-p[j]}
#			}
```

```
# 2. Intervention in second development factor.
#	scale <- 51.285
#	for (i in 1:10) {p1[i,1]<-p[1]}
#	for (i in 1:6) {p1[i,2]<-p[2]}
#	p1[7,2]<-p82
#	p1[8,2]<-p82
#	p1[9,2]<-p82
#	p1[10,2]<-p82
#	for (j in 3:9) {
#	for (i in 1:10) {p1[i,j]<-p[j]}
#			}
#	lambda82<-g82+1
#	p82<-1/lambda82
# Use one of the following 2 lines:
#	g82~dgamma(0.005,0.01) #This is a prior with a large variance
#	g82~dgamma(25,50) #This is a prior with a small variance

#3. Using latest 3 years for estimation of development factors.
#	scale <- 55.7366
#	for (j in 1:6) {
#	for (i in 1:(7-j)) {p1[i,j]<-op[j]}
#	for (i in (8-j):10) {p1[i,j]<-p[j]}
#	}
#	for (j in 7:9) {
#	for (i in 1:10) {p1[i,j]<-p[j]}
#		}
#	for (k in 1:6) {
#			op[k]<-1/olambda[k]
#			olambda[k]<-exp(og[k])+1
#			og[k]~dnorm(0.5,1.0E-6)
#					}

# Row totals and overall reserve
		R[1] <- 0
		R[2] <- Y[46]
		R[3] <- sum(Y[47:48])
		R[4] <- sum(Y[49:51])
		R[5] <- sum(Y[52:55])
		R[6] <- sum(Y[56:60])
		R[7] <- sum(Y[61:66])
		R[8] <- sum(Y[67:73])
		R[9] <- sum(Y[74:81])
		R[10] <- sum(Y[82:90])
		Total <- sum(R[2:10])

	}
```

```
# DATA
list(
row=c(1,1,1,1,1,1,1,1,1,
2,2,2,2,2,2,2,2,
3,3,3,3,3,3,3,4,4,
4,4,4,4,5,5,5,5,5,
6,6,6,6,7,7,7,8,
8,9,2,3,3,4,4,
4,5,5,5,5,6,6,6,6,6,
7,7,7,7,7,7,8,8,8,8,
8,8,8,9,9,9,9,9,9,9,
9,10,10,10,10,10,10,10,10,10),
col=c(1,2,3,4,5,6,7,8,9,
1,2,3,4,5,6,7,8,
1,2,3,4,5,6,7,1,2,3,
4,5,6,1,2,3,4,5,1,
2,3,4,1,2,3,1,
2,1,9,8,9,7,8,9,
6,7,8,9,5,6,7,8,9,4,
5,6,7,8,9,3,4,5,6,7,
8,9,2,3,4,5,6,7,8,9,
1,2,3,4,5,6,7,8,9),
Y=c(
766940,610542,482940,527326,574398,146342,139950,227229,67948,
884021,933894,1183289,445745,320996,527804,266172,425046,
1001799,926219,1016654,750816,146923,495992,280405,
1108250,776189,1562400,272482,352053,206286,
693190,991983,769488,504851,470639,
937085,847498,805037,705960,
847631,1131398,1063269,
1061648,1443370,
986608,
NA,
NA,NA,
NA,NA,NA,
NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA),
D=c(
357848,1124788,1735330,2218270,2745596,3319994,3466336,3606286,3833515,
352118,1236139,2170033,3353322,3799067,4120063,4647867,4914039,
290507,1292306,2218525,3235179,3985995,4132918,4628910,
310608,1418858,2195047,3757447,4029929,4381982,
443160,1136350,2128333,2897821,3402672,
396132,1333217,2180715,2985752,
440832,1288463,2419861,
359480,1421128,
```

```
376686,
NA,
NA,NA,
NA,NA,NA,
NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA),
DD=c(5339085,
4909315,NA,
4588268,NA,NA,
3873311,NA,NA,NA,
3691712,NA,NA,NA,NA,
3483130,NA,NA,NA,NA,NA,
2864498,NA,NA,NA,NA,NA,NA,
1363294,NA,NA,NA,NA,NA,NA,NA,
344014,NA,NA,NA,NA,NA,NA,NA,NA))
```

#INITIAL VALUES

This is what is used for 1.

For 2, replace the first line by

```
list(g=c(0,0,0,0,0,0,0,0,0),  g82=0.5,
```

For 3, replace the first line by

```
list(g=c(0,0,0,0,0,0,0,0,0), og=c(0,0,0,0,0,0),
```

```
list(g=c(0,0,0,0,0,0,0,0,0),
Z=c(NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,
NA,NA,NA,
NA,NA,
NA,
0,
0,0,
0,0,0,
0,0,0,0,
0,0,0,0,0,
0,0,0,0,0,0,
0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0))
```