Model risk in claims reserving within Tweedie's compound Poisson models

Gareth W. Peters[1,2] Pavel V. Shevchenko[2,*] Mario V. Wüthrich[3]

[1] *UNSW Mathematics and Statistics Department, Sydney, 2052, Australia*

[2] *CSIRO Mathematical and Information Sciences, Sydney, Locked Bag 17, North Ryde, NSW, 1670, Australia*

[3] *ETH Zurich, Department of Mathematics, CH-8092 Zurich, Switzerland*

[*] *presenting/corresponding author, e-mail: Pavel.Shevchenko@csiro.au*

Abstract

In this paper we examine the claims reserving problem using Tweedie's compound Poisson model. We develop maximum likelihood and Bayesian Markov chain Monte Carlo simulation approaches to fit the model and then compare estimated models under different scenarios. The key point we demonstrate relates to comparison of reserving quantities with and without model uncertainty incorporated into the prediction. We consider both the model selection problem and the model averaging solutions for the predicted reserves. As a part of this process we also consider the sub problem of variable selection to obtain a parsimonious representation of the model being fitted.

Keywords: Claims reserving, model risk, Tweedie's compound Poisson model, Bayesian analysis, model selection, model averaging, Markov chain Monte Carlo.

Model risk in claims reserving within Tweedie's compound Poisson models

Dr Pavel V. Shevchenko

Principal Research Scientist, Team leader

CSIRO Division of Mathematical and Information Sciences

Quantitative Risk Management group, Sydney, Australia

E-mail: Pavel.Shevchenko@csiro.au   www.cmis.csiro.au/Pavel.Shevchenko

Gareth Peters (UNSW/CSIRO), Pavel Shevchenko (CSIRO), and Mario Wüthrich (ETH). ***Model uncertainty in claims reserving within Tweedie's compound Poisson models,*** **preprint 2008.**

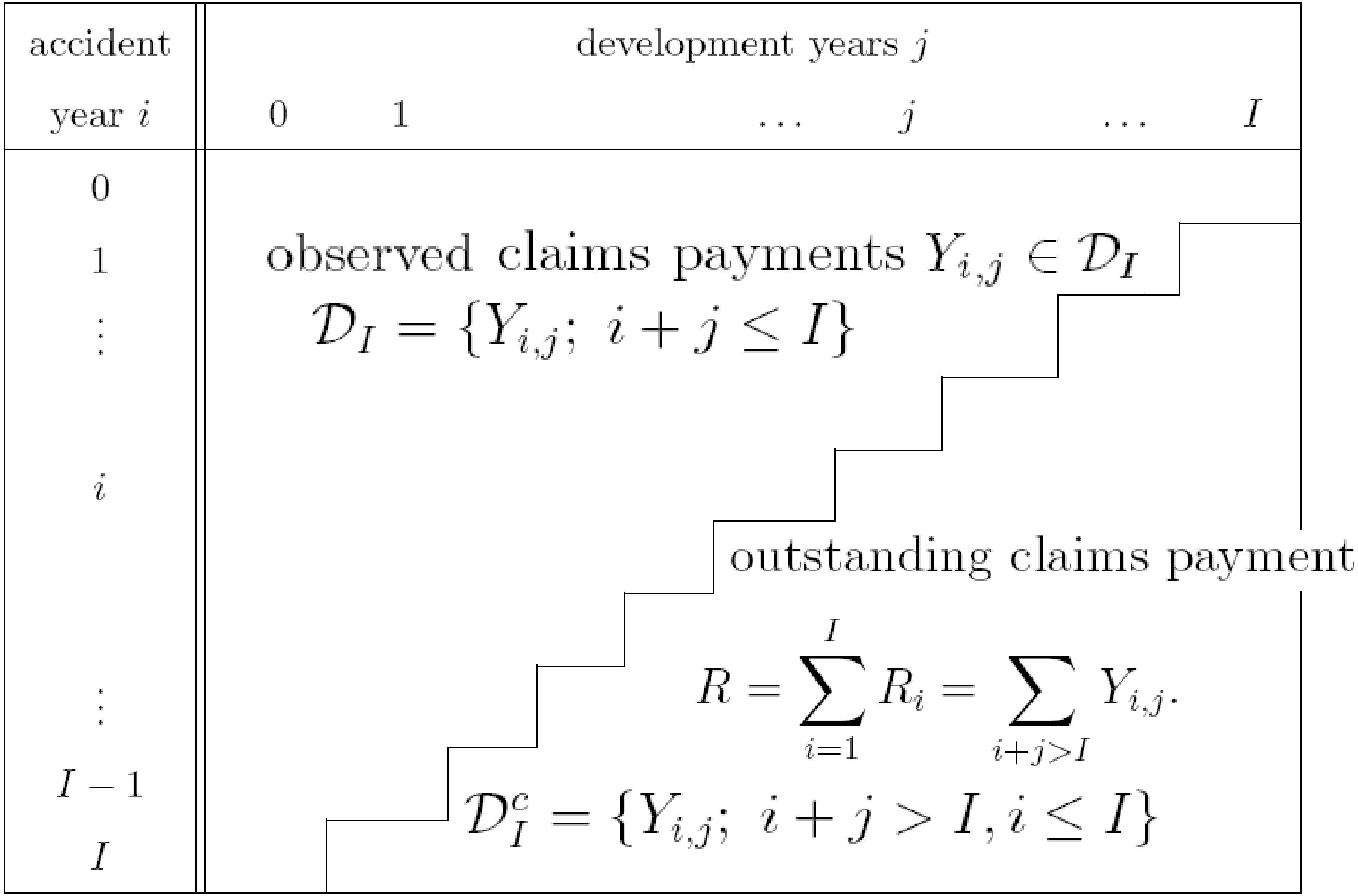
Department of Actuarial Studies, Macquarie University
8 October 2008

Commonwealth Scientific and Industrial Research Organization of Australia (**CSIRO**) – national research agency formed in 1926. Approx 6500 staff (Divisions: Industrial Physics, Minerals, Mathematical&Information Sciences, Marine and Atmospheric Research, etc.) www.csiro.au

Division of Mathematical and Information Sciences (**CMIS**) (over 100 researchers): Decision Technology, Biotechnology and Health Informatics, Environmental Informatics www.cmis.csiro.au

Quantitative Risk Management (**QRM**) group (approx. 20 staff): financial risk, infrastructure, environment risk, security, air-transport). Financial Risk – operational risk, credit risk, market risk, option pricing, insurance – validation, consulting, model and software development, www.cmis.csiro.au/QRM

Claims Reserving (non-life insurance), solvency requirements, claims development triangle



Year	0	1	2	3	4	5	6	7	8	9
0	594.6975	372.1236	89.5717	20.7760	20.6704	6.2124	6.5813	1.4850	1.1130	1.5813
1	634.6756	324.6406	72.3222	15.1797	6.7824	3.6603	5.2752	1.1186	1.1646	
2	626.9090	297.6223	84.7053	26.2768	15.2703	6.5444	5.3545	0.8924		
3	586.3015	268.3224	72.2532	19.0653	13.2976	8.8340	4.3329			
4	577.8885	274.5229	65.3894	27.3395	23.0288	10.5224				
5	618.4793	282.8338	57.2765	24.4899	10.4957					
6	560.0184	289.3207	56.3114	22.5517						
7	528.8066	244.0103	52.8043							
8	529.0793	235.7936								
9	567.5568									

Table 2: Data - annual claims Y_{ij} for each accident year i and development year j

Content

- Tweedie's compound Poisson family to model annual claims
- Process Uncertainty, Parameter Estimation Error, Model uncertainty
- Variable selection
- Maximum likelihood and Bayesian estimation
- MCMC (random walk Metropolis-Hastings within Gibbs)
- Analysis/Conclusions

$\widehat{R}$ - predictor for R and estimator for $E[R|\mathcal{D}_I]$

$$R = \sum_{i=1}^{I} R_i = \sum_{i+j>I} Y_{i,j} \qquad E[R|\mathcal{D}_I] = \sum_{i=1}^{I} E[R_i|\mathcal{D}_I]$$

$$\mathrm{msep}_{R|\mathcal{D}_I}\left(\widehat{R}\right) = E\left[\left(R - \widehat{R}\right)^2 \middle| \mathcal{D}_I\right]$$

Mean Square Error of Prediction

$$\mathrm{msep}_{R|\mathcal{D}_I}\left(\widehat{R}\right) = \mathrm{Var}(R|\mathcal{D}_I) + \left(E[R|\mathcal{D}_I] - \widehat{R}\right)^2$$

$$= \text{process variance} + \text{estimation error}$$

$\widehat{R} = E[R|\mathcal{D}_I]$ *"best estimate" of reserve*

Bayesian context – variance decomposition

$$\mathrm{Var}(R|\mathcal{D}_I) = E[\mathrm{Var}(R|\boldsymbol{\theta},\mathcal{D}_I)|\mathcal{D}_I] + \mathrm{Var}(E[R|\boldsymbol{\theta},\mathcal{D}_I]|\mathcal{D}_I)$$

$$= \text{average process variance} + \text{parameter estimation error.}$$

$\boldsymbol{\theta}$ *is model parameter vector modelled as random variable*

Tweedie's compound Poisson model

$$Y_{i,j} \text{ are independent for } i,j \in \{0,\ldots,I\}$$

$$Y_{i,j} = 1_{\{N_{i,j}>0\}} \sum_{k=1}^{N_{i,j}} X_{i,j}^{(k)}$$

$N_{i,j}$ and $X_{i,j}^{(k)}$ are independent for all k

$N_{i,j}$ is Poisson distributed with parameter $\lambda_{i,j}$

$X_{i,j}^{(k)}$ are independent gamma severities with mean $\tau_{i,j} > 0$ and shape parameter $\gamma > 0$

Tweedie's compound Poisson: exponential dispersion family representation

$$P\left[Y_{i,j}=0\right]=P\left[N_{i,j}=0\right]=\exp\left\{-\phi_{i,j}^{-1}\kappa_p(\theta_{i,j})\right\}$$

$$f_{\theta_{i,j}}(y;\phi_{i,j},p)=c(y;\phi_{i,j},p)\exp\left\{\frac{y\ \theta_{i,j}-\kappa_p(\theta_{i,j})}{\phi_{i,j}}\right\}\ \text{ for } y>0$$

where $\quad \theta_{i,j}<0,\ \phi_{i,j}>0,\quad \kappa_p(\theta)\overset{def.}{=}\frac{1}{2-p}\left[(1-p)\theta\right]^{\gamma}$

$$c(y;\phi,p)=\sum_{r\geq 1}\left(\frac{(1/\phi)^{\gamma+1}y^{\gamma}}{(p-1)^{\gamma}(2-p)}\right)^{r}\frac{1}{r!\Gamma(r\gamma)y}$$

$$p=p(\gamma)=\frac{\gamma+2}{\gamma+1}\ \in\ (1,2)\quad \phi_{i,j}=\frac{\lambda_{i,j}^{1-p}\tau_{i,j}^{2-p}}{2-p}>0$$

$$\theta_{i,j}=\left(\frac{1}{1-p}\right)(\mu_{i,j})^{(1-p)}<0,\ \ \mu_{i,j}=\lambda_{i,j}\tau_{i,j}>0$$

Final representation: Tweedie's compound Poisson model

$$P\left[Y_{i,j}=0\right]=P\left[N_{i,j}=0\right]=\exp\left\{-\phi_{i,j}^{-1}\frac{\mu_{i,j}^{2-p}}{2-p}\right\}$$

$$f_{\mu_{i,j}}(y;\phi_{i,j},p)=c(y;\phi_{i,j},p)\exp\left\{\phi_{i,j}^{-1}\left[y\,\frac{\mu_{i,j}^{1-p}}{1-p}-\frac{\mu_{i,j}^{2-p}}{2-p}\right]\right\}\quad \text{for } y>0$$

$$E\left[Y_{i,j}\right]=\frac{\partial}{\partial\theta_{i,j}}\kappa_p(\theta_{i,j})=\kappa_p'(\theta_{i,j})=\left[(1-p)\theta_{i,j}\right]^{1/(1-p)}=\mu_{i,j}$$

$$\mathrm{Var}\left(Y_{i,j}\right)=\phi_{i,j}\kappa_p''(\theta_{i,j})=\phi_{i,j}\;\mu_{i,j}^{p}$$

$p\in(1,2)$, typically fixed by the modeller *Model Risk*

$p\rightarrow 1$, overdispersed Poisson model

$p\rightarrow 2$, gamma model

$p=0$, Gaussian density

$p=3$, inverse Gaussian

Parameter estimation

estimate $\mu_{i,j}$, p and $\phi_{i,j}$ using observations $\mathcal{D}_I$

Model Assumptions

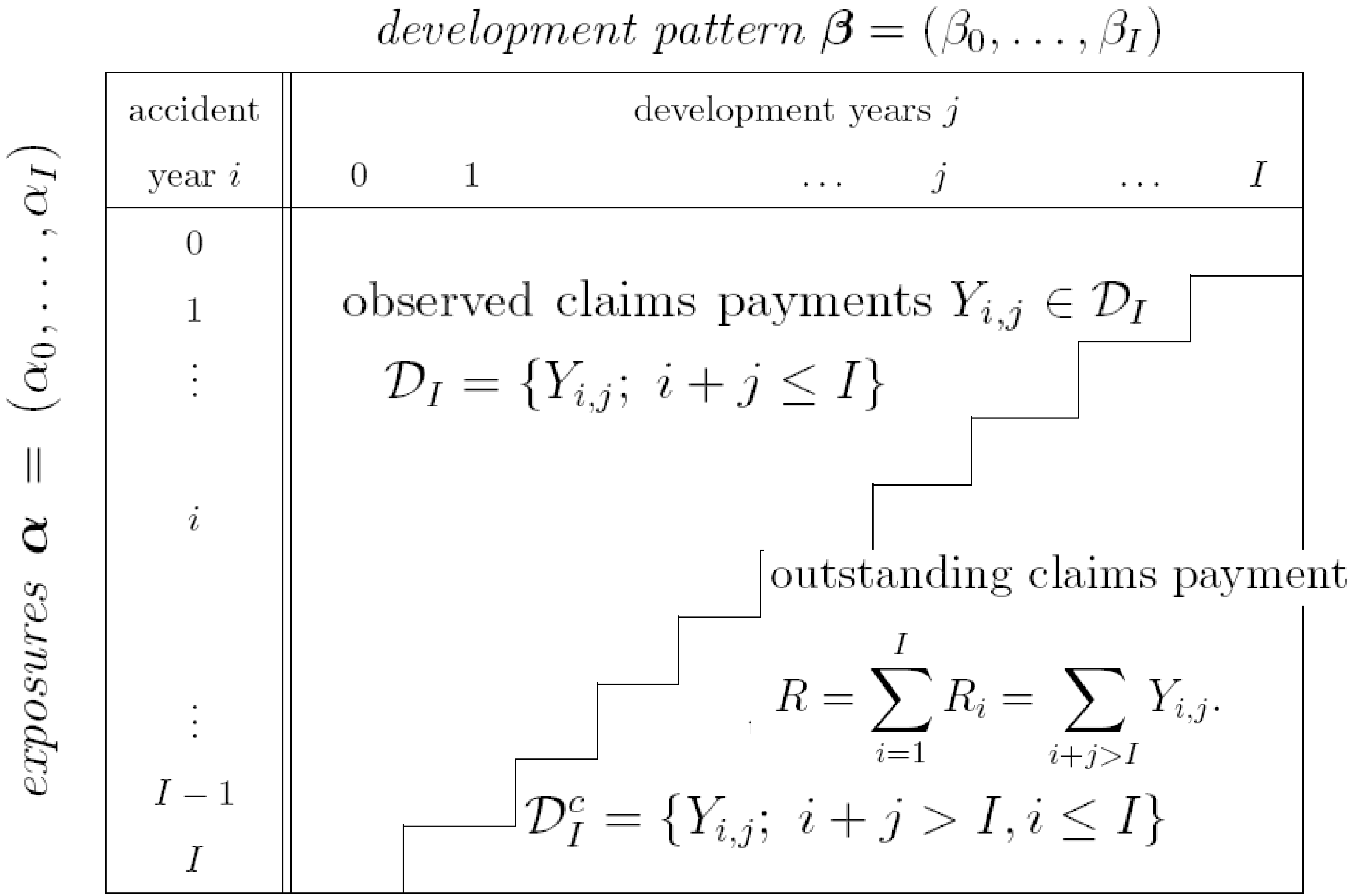
multiplicative model $\quad \mu_{i,j} = \alpha_i \; \beta_j$

$exposures \;\; \boldsymbol{\alpha} \;=\; (\alpha_0, \ldots, \alpha_I)$

$development\ pattern\ \boldsymbol{\beta} = (\beta_0, \ldots, \beta_I)$

$\phi_{i,j} = \phi \;\; and \;\; \alpha_i > 0,\ \beta_j > 0$

$\alpha_0 = 1$



development pattern $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_I)$
exposures $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_I)$
accident year i
development years j
0 1 … j … I
0
1
⋮
i
⋮
$I-1$
I
observed claims payments $Y_{i,j} \in \mathcal{D}_I$
$\mathcal{D}_I = \{Y_{i,j};\ i + j \leq I\}$
outstanding claims payment
$R = \sum_{i=1}^{I} R_i = \sum_{i+j>I} Y_{i,j}.$
$\mathcal{D}_I^c = \{Y_{i,j};\ i + j > I, i \leq I\}$

Likelihood function

$$\boldsymbol{\theta} = (p, \phi, \boldsymbol{\alpha}, \boldsymbol{\beta})$$

$$L_{\mathcal{D}_I}(\boldsymbol{\theta}) = \prod_{i+j \le I} c(Y_{i,j}; \phi, p) \exp\left\{\phi^{-1}\left[Y_{i,j}\frac{(\alpha_i\beta_j)^{1-p}}{1-p} - \frac{(\alpha_i\beta_j)^{2-p}}{2-p}\right]\right\}$$

$$c(y; \phi, p) = \sum_{r \ge 1}\left(\frac{(1/\phi)^{\gamma+1}y^{\gamma}}{(p-1)^{\gamma}(2-p)}\right)^r \frac{1}{r!\Gamma(r\gamma)y} = \frac{1}{y}\sum_{r \ge 1} W_r$$

$$\log W_r = r \log z - \log \Gamma(1+r) - \log \Gamma(\gamma r)$$

$$z = \frac{(1/\phi)^{\gamma+1}y^{\gamma}}{(p-1)^{\gamma}(2-p)}$$

$$\log W_r \approx r\left\{\log z + (1+\gamma) - \gamma \log \gamma - (1+\gamma)\log r\right\} - \log(2\pi) - \frac{1}{2}\log \gamma - \log r$$

Likelihood evaluation, Dunn and Smyth (2005)

$$\frac{\partial \log W_r}{\partial r} \approx \log z - \log r - \gamma \log(\gamma r)$$ W_r is unimodal in r

Solving $\partial W_r / \partial r = 0$ results in $R_0 = R_0(\phi, p) = \dfrac{y^{2-p}}{(2-p)\phi}$

find $R_L < R_0 < R_U$ such that

$$c(y;\phi,p) \approx \tilde{c}(y;\phi,p) = \frac{1}{y}\sum_{r=R_L}^{R_U} W_r$$

$$c(y;\phi,p) - \tilde{c}(y;\phi,p) < W_{R_L-1}\frac{1-q_L^{R_L-1}}{1-q_L} + W_{R_U+1}\frac{1}{1-q_U}$$

$$q_L = \exp\left(\frac{\partial \log W_r}{\partial r}\right)\Bigg|_{r=R_L-1}, \quad q_U = \exp\left(\frac{\partial \log W_r}{\partial r}\right)\Bigg|_{r=R_U+1}$$

Likelihood evaluation numerically

find $R_L < R_0 < R_U$ such that

$$W_{R_L} \leqslant e^{-37} W_{R_0} \text{ (or } R_L = 1) \text{ and } W_{R_U} \leqslant e^{-37} W_{R_0}$$

to avoid numerical overflow problems

$$c(y;\phi,p) \approx \widetilde{c}(y;\phi,p) = \frac{1}{y} \sum_{r=R_L}^{R_U} W_r$$

$$\log \widetilde{c}(y;\phi,p) = -\log y + \log W_{R_0} + \log \left(\sum_{r=R_L}^{R_U} \exp\left(\log\left(W_R\right) - \log\left(W_{R_0}\right)\right) \right)$$

if p is close to 1, then likelihood may become multimodal
if p is close to 2, then number of terms to evaluate may become large.
For our dataset, we restrict p to be in the range [1.1; 1.95]

Maximum likelihood estimation

maximizing $L_{\mathcal{D}_I}(\boldsymbol{\theta})$ in $\boldsymbol{\theta} = (p, \phi, \boldsymbol{\alpha}, \boldsymbol{\beta})$ leads to

$$\widehat{\boldsymbol{\theta}}^{\text{MLE}} = (\widehat{p}^{\text{MLE}}, \widehat{\phi}^{\text{MLE}}, \widehat{\boldsymbol{\alpha}}^{\text{MLE}}, \widehat{\boldsymbol{\beta}}^{\text{MLE}})$$

typically MLE is done for fixed p (expert choice)

$$\widehat{R}^{\text{MLE}} = \sum_{i+j>I} \widehat{\alpha}_i^{\text{MLE}} \, \widehat{\beta}_j^{\text{MLE}} \quad \text{the best estimate reserves for } R$$

$$\frac{\partial \ln L_{\mathcal{D}_I}(\boldsymbol{\theta})}{\partial \beta_k} = \frac{\partial}{\partial \beta_k} \sum_{j=0}^{I} \sum_{i=0}^{I-j} \phi^{-1} \left(Y_{i,j} \frac{(\alpha_i \beta_j)^{1-p}}{1-p} - \frac{(\alpha_i \beta_j)^{2-p}}{2-p} \right)$$

$$= \sum_{i=0}^{I-k} \phi^{-1} \left(Y_{i,k} \alpha_i^{1-p} \beta_k^{-p} - \alpha_i^{2-p} \beta_k^{1-p} \right)$$

$$\beta_k = \frac{\sum_{i=0}^{I-k} Y_{i,k} \alpha_i^{1-p}}{\sum_{i=0}^{I-k} \alpha_i^{2-p}}, \quad k = 0, \ldots, I$$

Covariances between maximum likelihood estimators

Observed Information matrix and Gaussian approximation

$$(\boldsymbol{I})_{i,j} = -\left.\frac{\partial^2 \ln L_{\mathcal{D}_I}(\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}^{\mathrm{MLE}}}$$

$$\mathrm{cov}\left(\widehat{\theta}_i^{\mathrm{MLE}}, \widehat{\theta}_j^{\mathrm{MLE}}\right) = \left(\boldsymbol{I}^{-1}\right)_{i,j}$$

$$\widehat{\beta}_I^{\mathrm{MLE}} = Y_{0,I}$$

$$\mathrm{cov}(\widehat{\beta}_I^{\mathrm{MLE}}, \widehat{\theta}_i^{\mathrm{MLE}}) = 0, \quad \widehat{\theta}_i^{\mathrm{MLE}} \neq \widehat{\beta}_I^{\mathrm{MLE}}$$

Maximum likelihood: process and estimation errors

$$\widehat{R}^{\text{MLE}} = \sum_{i+j>I} \widehat{\alpha}_i^{\text{MLE}} \, \widehat{\beta}_j^{\text{MLE}}$$

$$\text{stdev}\left(\widehat{R}^{\text{MLE}}\right) = \sqrt{\text{Var}\left(\widehat{R}^{\text{MLE}}\right)} \quad \text{parameter estimation error}$$

$$\begin{aligned}\widehat{\text{Var}}\left(\widehat{R}^{\text{MLE}}\right) = & \sum_{i_1+j_1>I} \sum_{i_2+j_2>I} \widehat{\alpha}_{i_1}^{\text{MLE}} \widehat{\alpha}_{i_2}^{\text{MLE}} \text{cov}\left(\widehat{\beta}_{j_1}^{\text{MLE}}, \widehat{\beta}_{j_2}^{\text{MLE}}\right) \\ & + \sum_{i_1+j_1>I} \sum_{i_2+j_2>I} \widehat{\beta}_{j_1}^{\text{MLE}} \widehat{\beta}_{j_2}^{\text{MLE}} \text{cov}\left(\widehat{\alpha}_{i_1}^{\text{MLE}}, \widehat{\alpha}_{i_2}^{\text{MLE}}\right) \\ & + 2 \sum_{i_1+j_1>I} \sum_{i_2+j_2>I} \widehat{\alpha}_{i_1}^{\text{MLE}} \widehat{\beta}_{j_2}^{\text{MLE}} \text{cov}\left(\widehat{\alpha}_{i_2}^{\text{MLE}}, \widehat{\beta}_{j_1}^{\text{MLE}}\right)\end{aligned}$$

$$\widehat{\text{Var}}\left(R\right) = \sum_{i+j>I} \left(\widehat{\alpha}_i^{\text{MLE}} \, \widehat{\beta}_j^{\text{MLE}}\right)^{\widehat{p}^{\text{MLE}}} \widehat{\phi}^{\text{MLE}} \quad \text{process variance}$$

$$\begin{aligned}\widehat{\text{msep}}_{R|\mathcal{D}_I}\left(\widehat{R}^{\text{MLE}}\right) &= \widehat{\text{Var}}\left(R\right) + \widehat{\text{Var}}\left(\widehat{R}^{\text{MLE}}\right) \\ &= \text{MLE process variance} + \text{MLE estimation error}\end{aligned}$$

Bayesian inference

$\boldsymbol{\theta} = (p, \phi, \boldsymbol{\alpha}, \boldsymbol{\beta})$ are treated as random

$\pi(\boldsymbol{\theta} \mid \mathcal{D}_I) \propto L_{\mathcal{D}_I}(\boldsymbol{\theta})\ \pi(\boldsymbol{\theta})$ Markov Chain Monte Carlo (MCMC)

$$MAP: \qquad \hat{\boldsymbol{\theta}}^{MAP} = \arg\max_{\boldsymbol{\theta}} \left[\pi(\boldsymbol{\theta} \mid \mathcal{D}_I)\right],$$

$$MMSE: \qquad \hat{\boldsymbol{\theta}}^{MMSE} = E\left[\boldsymbol{\theta} \mid \mathcal{D}_I\right].$$

if the prior $\pi(\boldsymbol{\theta})$ is constant and the parameter range includes the MLE, then the MAP of the posterior is the same as the MLE.

Gaussian approximation for $\pi(\boldsymbol{\theta} \mid \mathcal{D}_I)$

$$\ln \pi(\boldsymbol{\theta} \mid \mathcal{D}_I) \approx \ln \pi(\hat{\boldsymbol{\theta}}^{MAP} \mid \mathcal{D}_I) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial\theta_i \partial\theta_j} \ln \pi(\boldsymbol{\theta} \mid \mathcal{D}_I)\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}^{MAP}} \left(\theta_i - \hat{\theta}_i^{MAP}\right)\left(\theta_j - \hat{\theta}_j^{MAP}\right)$$

Bayesian inference estimates

$$E\left[R\middle|\mathcal{D}_I\right] = \sum_{i+j>I} E\left[\alpha_i\beta_j\middle|\mathcal{D}_I\right] \qquad \widetilde{R} = E\left[R|\boldsymbol{\theta}\right] = \sum_{i+j>I} \alpha_i\beta_j$$

the best consistent estimate of reserves (ER)

$$\widehat{R}^{\mathrm{B}} = E\left[\widetilde{R}\,\middle|\,\mathcal{D}_I\right] = \sum_{i+j>I} E\left[\alpha_i\beta_j\middle|\mathcal{D}_I\right] = E\left[R\middle|\mathcal{D}_I\right]$$

$$\mathrm{msep}_{R|\mathcal{D}_I}\left(\widehat{R}^{\mathrm{B}}\right) = E\left[\left(R-\widehat{R}^{\mathrm{B}}\right)^2\middle|\mathcal{D}_I\right] = \mathrm{Var}\left(R\middle|\mathcal{D}_I\right)$$

$$\mathrm{Var}\left(R\middle|\mathcal{D}_I\right) = \mathrm{Var}\left(\sum_{i+j>I} Y_{i,j}\middle|\mathcal{D}_I\right)$$

$$= \sum_{i+j>I} E\left[(\alpha_i\beta_j)^p\,\phi\middle|\mathcal{D}_I\right] + \mathrm{Var}\left(\widetilde{R}\,\middle|\,\mathcal{D}_I\right)$$

= average process variance (PV) + parameter estimation error (EE)

Note, model error is incorporated via averaging over values of p

Random Walk Metropolis Hastings (RW-MH) within Gibbs

1. Initialize randomly or deterministically for $t = 0$ the parameter vector $\boldsymbol{\theta}^{t=0}$ to the maximum likelihood estimates.

2. For $t = 1, ..., T$

a) Set $\boldsymbol{\theta}^t = \boldsymbol{\theta}^{t-1}$

b) For $i = 1, ..., 2I + 3$

Sample proposal θ_i^* from Gaussian truncated density

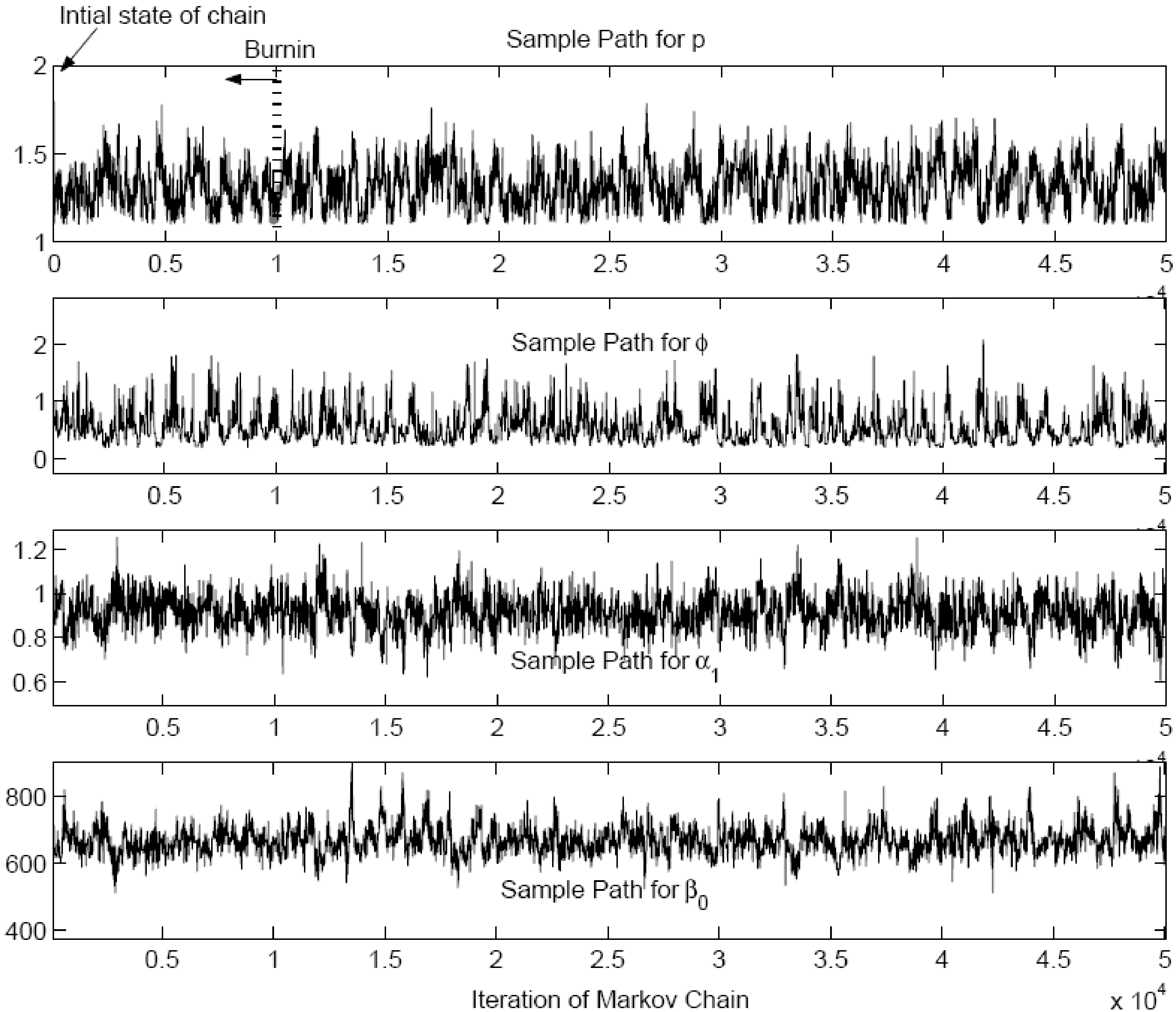
$$f_N^T\left(\theta_i^*; \theta_i^t, \sigma_{RWi}\right) = \frac{f_N\left(\theta_i^*; \theta_i^t, \sigma_{RWi}\right)}{F_N\left(b_i; \theta_i^t, \sigma_{RWi}\right) - F_N\left(a_i; \theta_i^t, \sigma_{RWi}\right)}$$

to obtain $\boldsymbol{\theta}^* = \left(\theta_1^t, ..., \theta_{i-1}^t, \theta_i^*, \theta_{i+1}^{t-1}, ...\right)$

Accept proposal with acceptance probability

$$\alpha\left(\boldsymbol{\theta}^t, \boldsymbol{\theta}^*\right) = \min\left\{1, \frac{\pi(\boldsymbol{\theta}^* \mid \mathcal{D}_I) f_N^T\left(\theta_i^t; \theta_i^*, \sigma_{RWi}\right)}{\pi(\boldsymbol{\theta}^t \mid \mathcal{D}_I) f_N^T\left(\theta_i^*; \theta_i^t, \sigma_{RWi}\right)}\right\}$$

***Note**: normalization constant in posterior is not needed; optimal acceptance rate is 0.234*



Intial state of chain
Burnin
Sample Path for p
Sample Path for ϕ
Sample Path for α1
Sample Path for β0
Iteration of Markov Chain
x 10^4

The prior domains

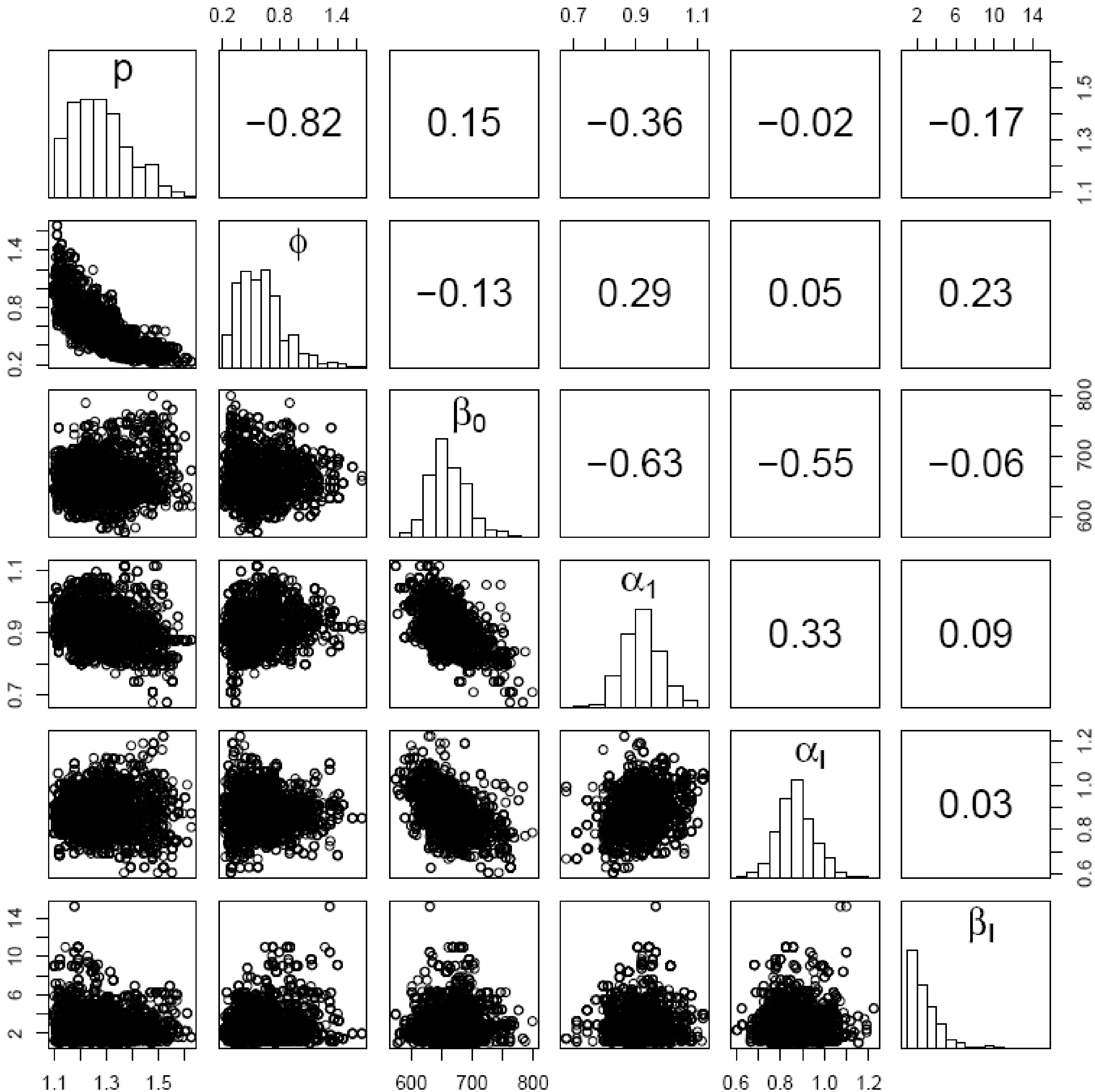
$p \in (1.1, 1.95)$, $\phi \in (0.01, 100)$, $\alpha_i \in (0.01, 100)$ and $\beta_j \in (0.01, 10^4)$

	MLE	MLE stdev	Bayesian posterior			σ_{RW}
			MMSE	stdev	$[Q_{0.05}; Q_{0.95}]$	
p	1.259	0.149	1.332 (0.007)	0.143 (0.004)	[1.127;1.590]	1.61
ϕ	0.351	0.201	0.533 (0.013)	0.289 (0.005)	[0.174;1.119]	1.94
α_1	0.918	0.056	0.901 (0.004)	0.074 (0.001)	[0.778;1.022]	0.842
α_2	0.946	0.051	0.946 (0.003)	0.073 (0.001)	[0.833;1.072]	0.907
α_3	0.861	0.048	0.861 (0.003)	0.068 (0.001)	[0.756;0.977]	0.849
α_4	0.891	0.049	0.902 (0.003)	0.072 (0.002)	[0.794;1.027]	0.893
α_5	0.879	0.051	0.876 (0.003)	0.070 (0.001)	[0.768;0.994]	0.932
α_6	0.842	0.048	0.843 (0.002)	0.069 (0.001)	[0.736;0.958]	0.751
α_7	0.762	0.046	0.762 (0.003)	0.066 (0.001)	[0.660;0.876]	0.888
α_8	0.763	0.047	0.765 (0.003)	0.067 (0.001)	[0.661;0.874]	0.897
α_9	0.848	0.059	0.856 (0.003)	0.090 (0.002)	[0.716;1.009]	1.276

Table 3: MLE and Bayesian estimators.

Table 3: MLE and Bayesian estimators.

	MLE	MLE stdev	Bayesian posterior			σ_{RW}
			MMSE	stdev	$[Q_{0.05}; Q_{0.95}]$	
β_0	669.1	27.7	672.7 (2.1)	39.7 (0.7)	[610.0;740.0]	296
β_1	329.0	14.4	331.1 (1.0)	20.6 (0.4)	[298.1;365.9]	190
β_2	77.43	4.38	78.06 (0.24)	6.10 (0.06)	[68.58;88.29]	75.4
β_3	24.59	1.96	24.95 (0.08)	2.64 (0.03)	[20.89;29.64]	40.9
β_4	16.28	1.55	16.65 (0.05)	2.09 (0.03)	[13.44;20.30]	40.6
β_5	7.773	1.028	8.068 (0.024)	1.356 (0.020)	[6.064;10.473]	26.0
β_6	5.776	0.937	6.115 (0.022)	1.261 (0.016)	[4.246;8.347]	24.1
β_7	1.219	0.396	1.494 (0.006)	0.609 (0.013)	[0.739;2.609]	13.1
β_8	1.188	0.476	1.622 (0.008)	0.802 (0.016)	[0.674;3.070]	15.1
β_9	1.581	0.790	2.439 (0.021)	1.496 (0.026)	[0.829;5.250]	32.1



p
ϕ
β0
α1
αl
βl
−0.82
0.15
−0.36
−0.02
−0.17
−0.13
0.29
0.05
0.23
−0.63
−0.55
−0.06
0.33
0.09
0.03
0.2 0.8 1.4
0.7 0.9 1.1
2 6 10 14
1.1 1.3 1.5
600 700 800
0.6 0.8 1.0 1.2

CSIRO

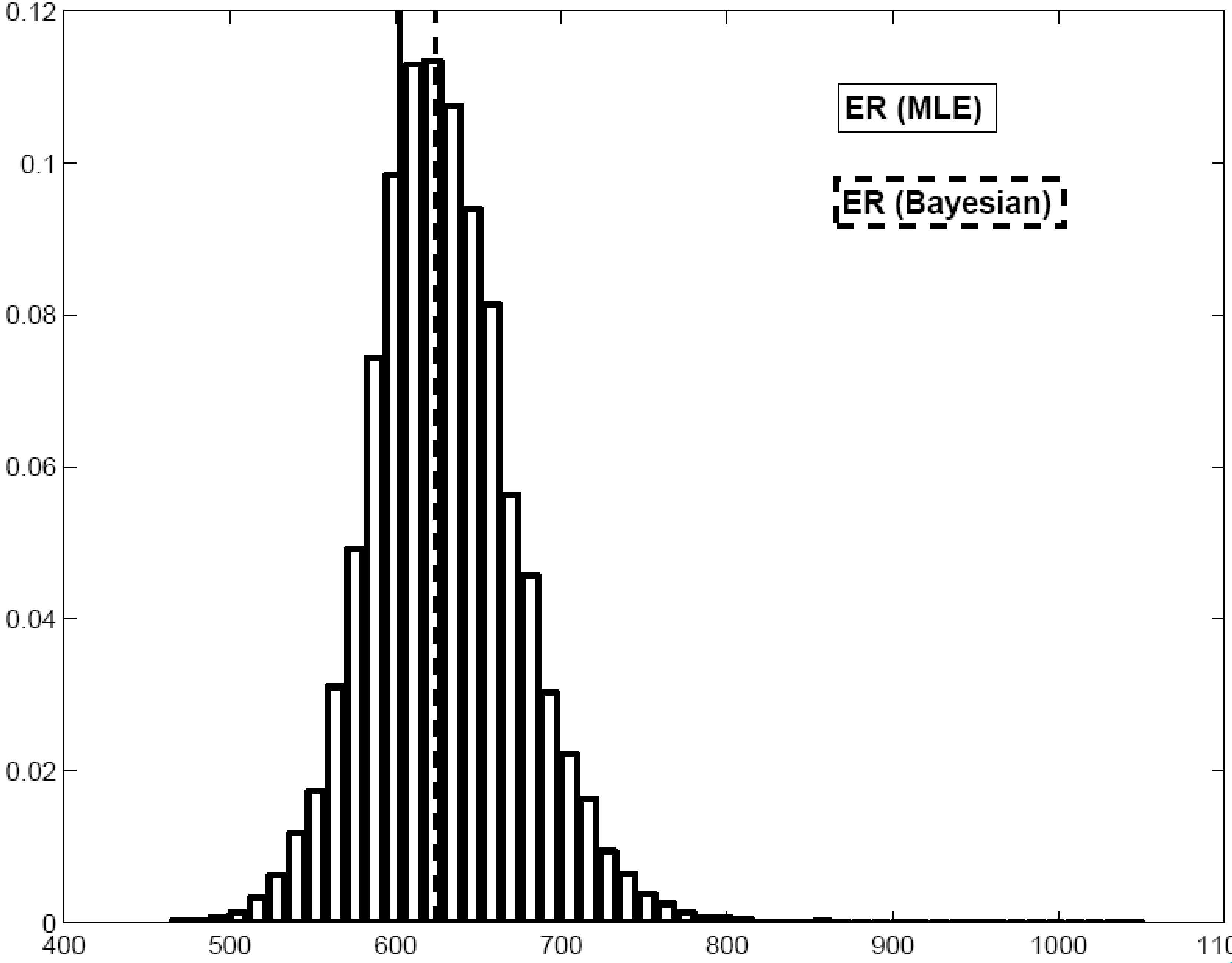


Figure 3: Predicted distribution of reserves, $\widetilde{R} = \sum_{i+j>I} \alpha_i \beta_j$.

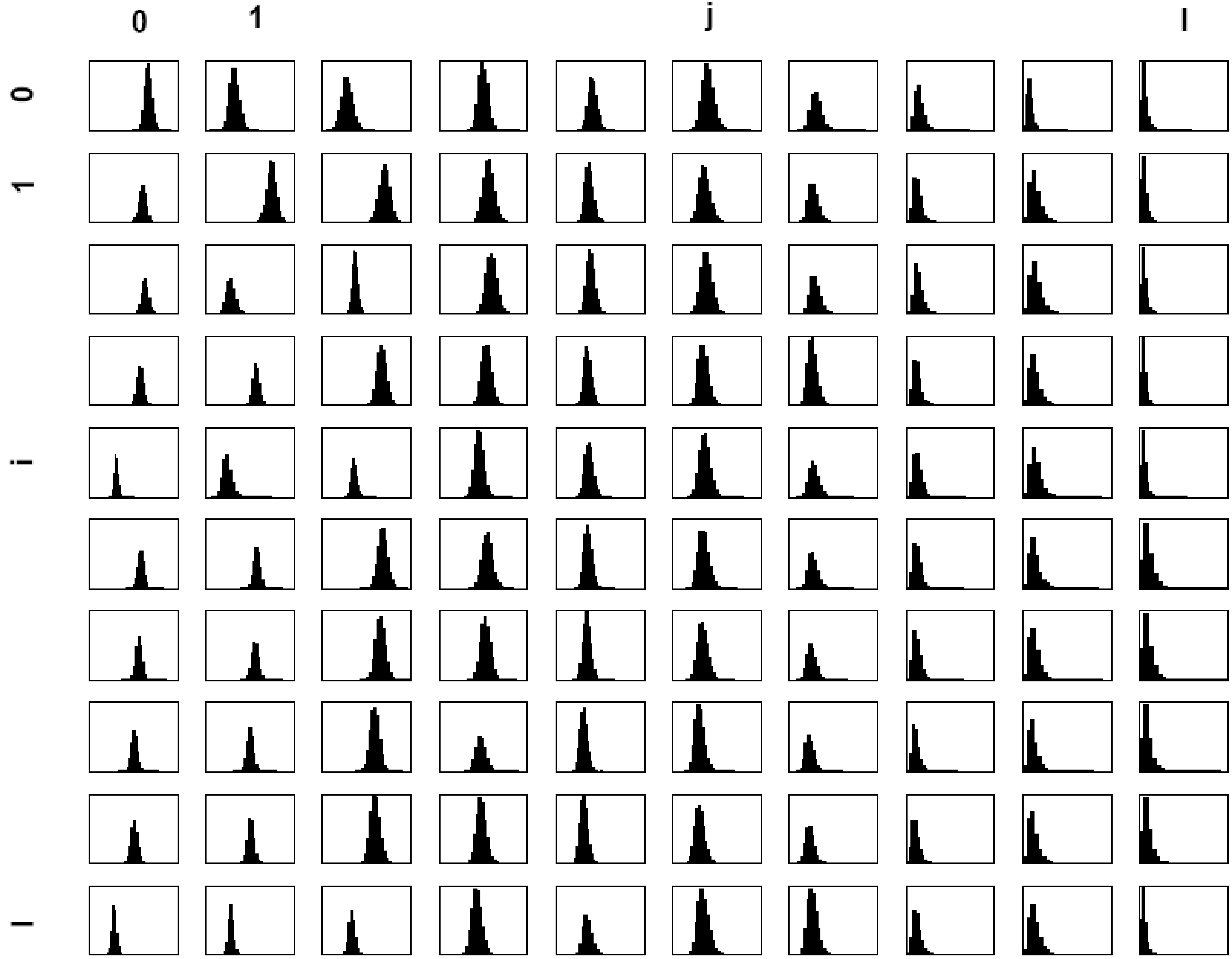


Figure 4: Posterior distributions for $\widetilde{R}_{i,j} = \alpha_i \beta_j$ estimated using MCMC.

CSIRO

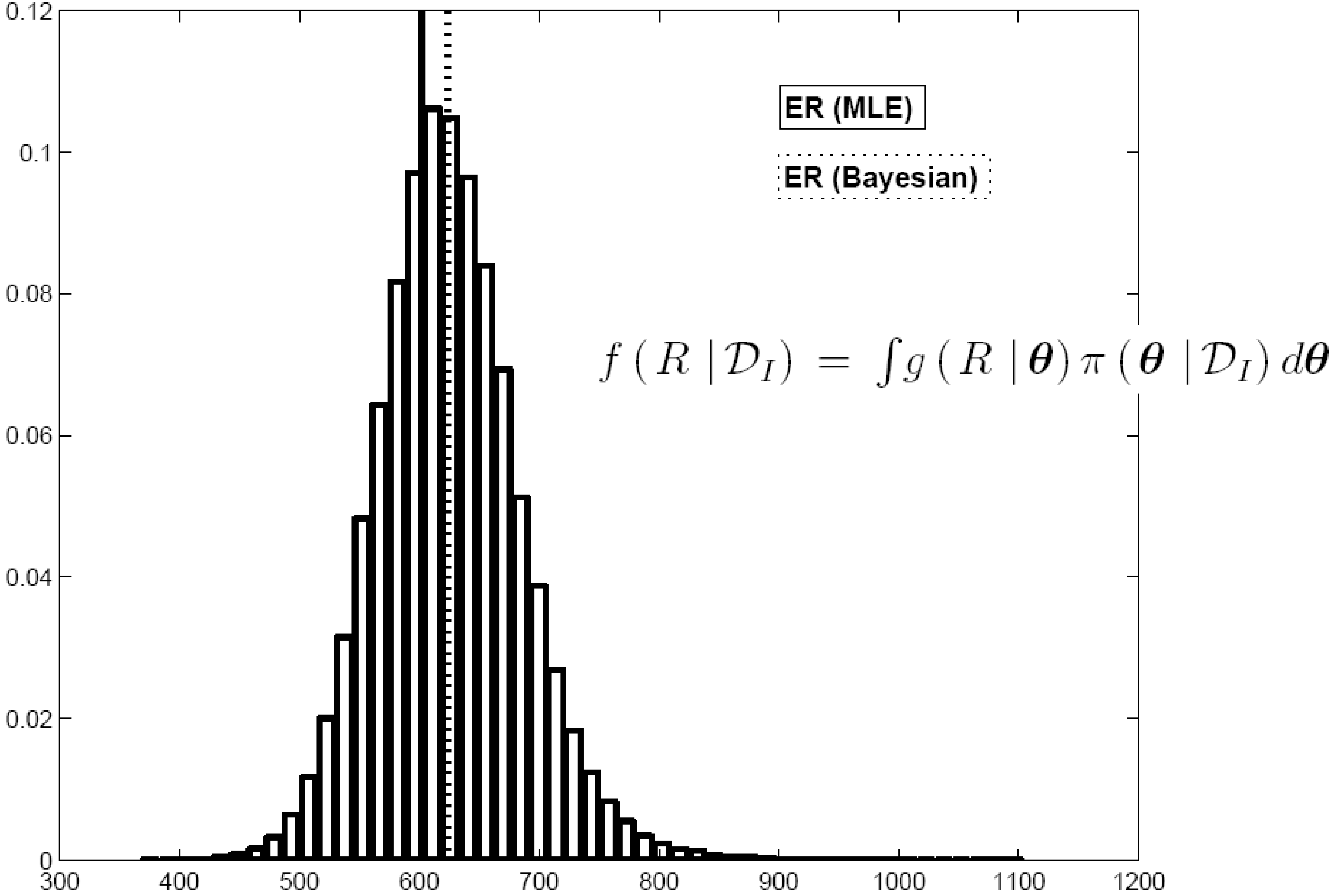
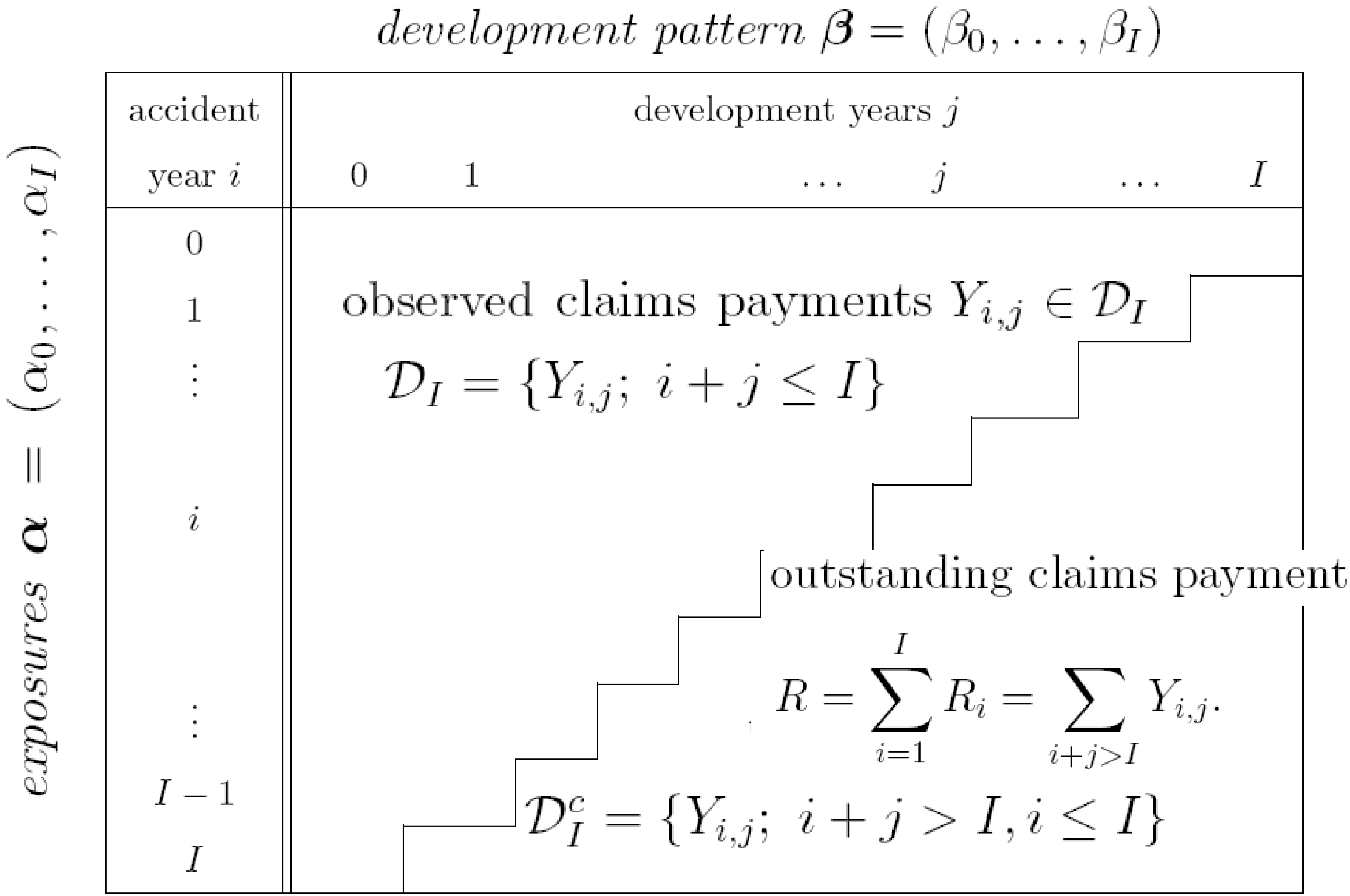


Figure 5: Distribution of total outstanding claims payment $R = \sum_{i+j>I} Y_{i,j}$, accounting for all process, estimation and model uncertainties.

Variable selection models



Variable selection models

- $M_0 : \boldsymbol{\theta}_{[0]} = \left(p, \phi, \widetilde{\alpha}_0 = \alpha_0, ..., \widetilde{\alpha}_I = \alpha_I, \widetilde{\beta}_0 = \beta_0, ..., \widetilde{\beta}_I = \beta_I\right)$ - **saturated model**.
- $M_1 : \boldsymbol{\theta}_{[1]} = \left(p, \phi, \widetilde{\beta}_0\right)$ with $\left(\widetilde{\beta}_0 = \beta_0 ... = \beta_I\right), (\alpha_0 = ... = \alpha_I = 1)$.
- $M_2 : \boldsymbol{\theta}_{[2]} = \left(p, \phi, \widetilde{\alpha}_1, \widetilde{\beta}_0, \widetilde{\beta}_1\right)$ with $(\alpha_0 = ... = \alpha_4 = 1)$, $(\widetilde{\alpha}_1 = \alpha_5 = ... = \alpha_I)$, $\left(\widetilde{\beta}_0 = \beta_0 = ... = \beta_4\right)$, $\left(\widetilde{\beta}_1 = \beta_5 = ... = \beta_I\right)$.
- $M_3 : \boldsymbol{\theta}_{[3]} = \left(p, \phi, \widetilde{\alpha}_1, \widetilde{\alpha}_2, \widetilde{\beta}_0, \widetilde{\beta}_1, \widetilde{\beta}_2\right)$ with $(\alpha_0 = \alpha_1 = 1)$, $(\widetilde{\alpha}_1 = \alpha_2 = ... = \alpha_5)$, $(\widetilde{\alpha}_2 = \alpha_6 = ... = \alpha_I)$, $\left(\widetilde{\beta}_0 = \beta_0 = \beta_1\right)$, $\left(\widetilde{\beta}_1 = \beta_2 = ... = \beta_5\right)$, $\left(\widetilde{\beta}_2 = \beta_6 = ... = \beta_I\right)$

Variable selection models

- $M_4 : \boldsymbol{\theta}_{[4]} = \left(p, \phi, \widetilde{\alpha}_1, \widetilde{\alpha}_2, \widetilde{\alpha}_3, \widetilde{\beta}_0, \widetilde{\beta}_1, \widetilde{\beta}_2, \widetilde{\beta}_3\right)$ with $(\alpha_0 = \alpha_1 = 1)$, $(\widetilde{\alpha}_1 = \alpha_2 = \alpha_3)$, $(\widetilde{\alpha}_2 = \alpha_4 = \alpha_5 = \alpha_6)$, $(\widetilde{\alpha}_3 = \alpha_7 = \alpha_8 = \alpha_I)$, $\left(\widetilde{\beta}_0 = \beta_0 = \beta_1\right)$, $\left(\widetilde{\beta}_1 = \beta_2 = \beta_3\right)$, $\left(\widetilde{\beta}_2 = \beta_4 = \beta_5 = \beta_6\right)$, $\left(\widetilde{\beta}_3 = \beta_7 = \beta_8 = \beta_I\right)$.

- $M_5 : \boldsymbol{\theta}_{[5]} = \left(p, \phi, \widetilde{\alpha}_1, \widetilde{\alpha}_2, \widetilde{\alpha}_3, \widetilde{\alpha}_4, \widetilde{\beta}_0, \widetilde{\beta}_1, \widetilde{\beta}_2, \widetilde{\beta}_3, \widetilde{\beta}_4\right)$ with $(\alpha_0 = \alpha_1 = 1)$, $(\widetilde{\alpha}_1 = \alpha_2 = \alpha_3)$, $(\widetilde{\alpha}_2 = \alpha_4 = \alpha_5)$, $(\widetilde{\alpha}_3 = \alpha_6 = \alpha_7)$, $(\widetilde{\alpha}_4 = \alpha_8 = \alpha_I)$, $\left(\widetilde{\beta}_0 = \beta_0 = \beta_1\right)$, $\left(\widetilde{\beta}_1 = \beta_2 = \beta_3\right)$, $\left(\widetilde{\beta}_2 = \beta_4 = \beta_5\right)$, $\left(\widetilde{\beta}_3 = \beta_6 = \beta_7\right)$, $\left(\widetilde{\beta}_4 = \beta_8 = \beta_I\right)$.

- $M_6 : \boldsymbol{\theta}_{[6]} = (p, \phi, \alpha_0, \widetilde{\alpha}_1, \beta_0, \beta_1, ..., \beta_I)$ with $(\widetilde{\alpha}_1 = \alpha_1 = ... = \alpha_I)$.

Variable selection models

the joint posterior distribution $\pi(M_k, \boldsymbol{\theta}_{[k]} \mid \mathcal{D}_I)$, $\boldsymbol{\theta}_{[k]} = \left(\widetilde{\theta}_1, \widetilde{\theta}_2, ..., \widetilde{\theta}_{N_{[k]}}\right)$

a prior distribution $\pi(M_k)$ for the model

a prior for the parameters conditional on the model $\pi(\boldsymbol{\theta}_{[k]} \mid M_k)$

$$\pi(\boldsymbol{\theta}_{[m]} \mid M_k) = \prod_{i=1}^{N_{[m]}} \left[b_{\widetilde{\theta}_{i,[m]}} - a_{\widetilde{\theta}_{i,[m]}}\right]^{-1}$$

The modified version of Congdon's (2006), formula A.3

$$\pi(M_k \mid \mathcal{D}_I) = \int \pi(M_k, \boldsymbol{\theta}_{[k]} \mid \mathcal{D}_I) d\boldsymbol{\theta}_{[k]} = \int \pi(M_k \mid \boldsymbol{\theta}_{[k]}, \mathcal{D}_I)\pi(\boldsymbol{\theta}_{[k]} \mid \mathcal{D}_I) d\boldsymbol{\theta}_{[k]}$$

$$\approx \frac{1}{T - T_b} \sum_{j=T_b+1}^{T} \pi(M_k \mid \mathcal{D}_I, \boldsymbol{\theta}_{j,[k]})$$

$$= \frac{1}{T - T_b} \sum_{j=T_b+1}^{T} \frac{L_{\mathcal{D}_I}(M_k, \boldsymbol{\theta}_{j,[k]}) \prod_{k=0}^{K} \pi(\boldsymbol{\theta}_{j,[k]} \mid M_k)\pi(M_k)}{\sum_{m=0}^{K} L_{\mathcal{D}_I}(M_m, \boldsymbol{\theta}_{j,[m]}) \prod_{k=0}^{K} \pi(\boldsymbol{\theta}_{j,[k]} \mid M_m)\pi(M_m)}$$

$$= \frac{1}{T - T_b} \sum_{j=T_b+1}^{T} \frac{L_{\mathcal{D}_I}(M_k, \boldsymbol{\theta}_{j,[k]})}{\sum_{m=0}^{K} L_{\mathcal{D}_I}(M_m, \boldsymbol{\theta}_{j,[m]})}.$$

Variable selection models

	M_0	M_1	M_2	M_3	M_4	M_5	M_6
$\pi(M_k \mid D_I)$	0.71	4.19E-54	3.04E-43	1.03E-28	6.71E-20	2.17E-21	0.29
DIC	399	649	600	535	498	507	398
LHR $p - value$	1	2.76E-50	1.67E-40	3.53E-28	5.78E-21	3.03E-23	0.043

Table 4: Posterior model probabilities $\pi(M_k|D_I)$,
Deviance Information Criterion (DIC) and
Likelihood Ratio (LHR) p-values
for variable selection models $M_0, \ldots, M_6$

Claims reserves

	Model Averaging	Model Selection for p
Estimated Reserves	$ER = \widehat{R}^{\mathrm{B}} = E[\widetilde{R}\vert\mathcal{D}_I]$	$ER_p = E[\widetilde{R}\vert\mathcal{D}_I, p]$
Process Variance	$PV = E\left[\sum \phi\,(\alpha_i\beta_j)^p \vert\, \mathcal{D}_I\right]$	$PV_p = E\left[\sum \phi\,(\alpha_i\beta_j)^p \vert\, \mathcal{D}_I, p\right]$
Estimation Error	$EE = \mathrm{Var}(\widetilde{R}\vert\mathcal{D}_I)$	$EE_p = \mathrm{Var}(\widetilde{R}\vert\mathcal{D}_I, p)$

Table 5: Quantities used for analysis of the claims reserving problem

$$ER = E[ER_p|\mathcal{D}_I],$$

$$PV = E[PV_p|\mathcal{D}_I],$$

$$EE = E[EE_p|\mathcal{D}_I] + \mathrm{Var}(ER_p|\mathcal{D}_I)$$

	Model Averaging	
Statistic	Bayesian Estimate	MLE Estimate
ER	624.1 (0.7)	602.630
$\sqrt{PV}$	37.3 (0.2)	25.937
$\sqrt{EE}$	44.8 (0.5)	28.336
$\sqrt{MSEP}$	58.3(0.5)	38.414

Table 6: Model averaged estimates

	Model Averaging	
VaR_q	R	$\widetilde{R}$
$\text{VaR}_{75\%}$	659.8 (0.9)	650.6 (1.0)
$\text{VaR}_{90\%}$	698.4 (1.2)	680.4 (1.3)
$\text{VaR}_{95\%}$	724.0 (1.5)	701.7 (1.6)

Table 7: Bayesian model averaged estimates of Value at Risk

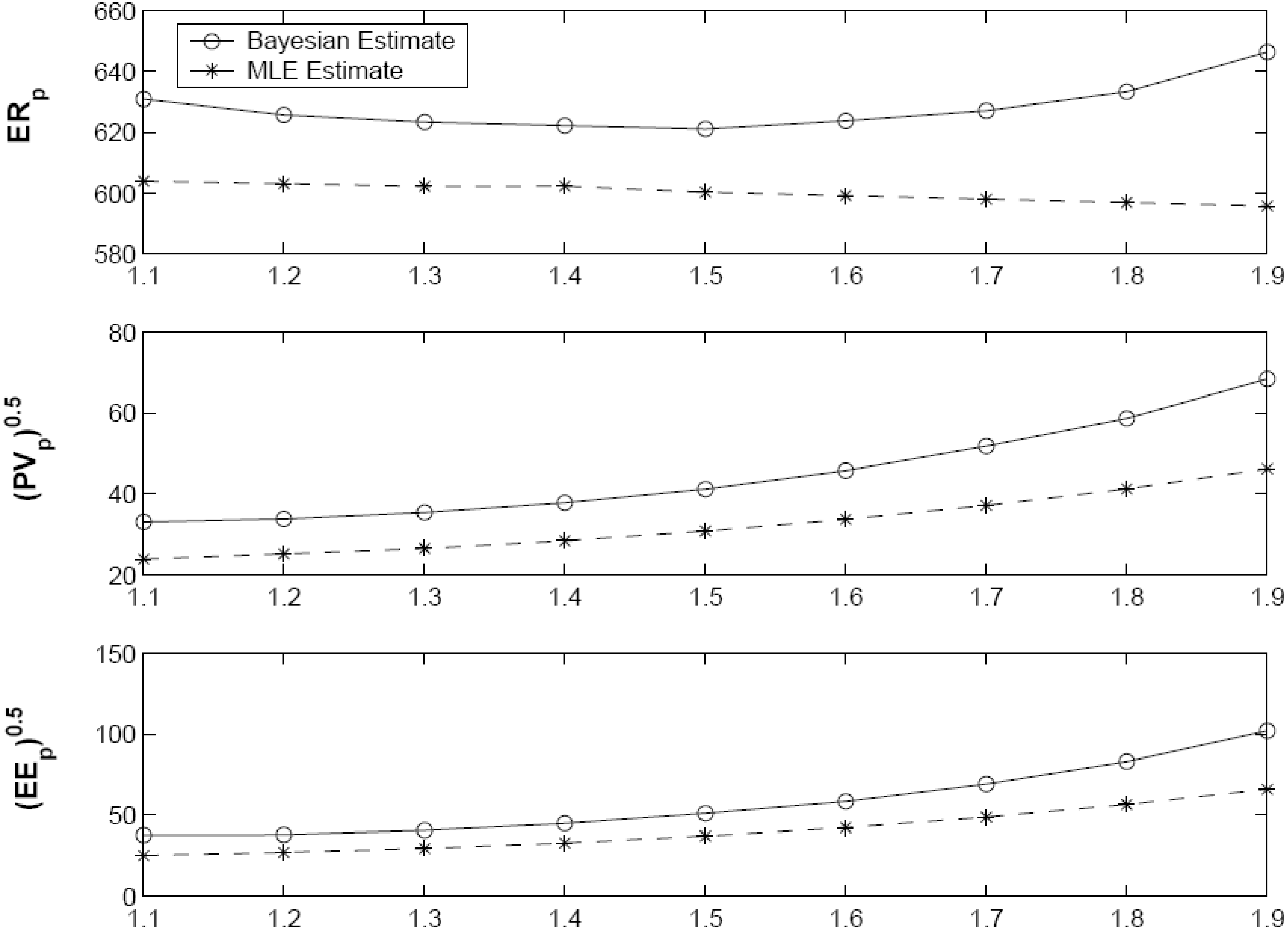


Figure 6: Estimates of quantities from Table 5 conditional on p.

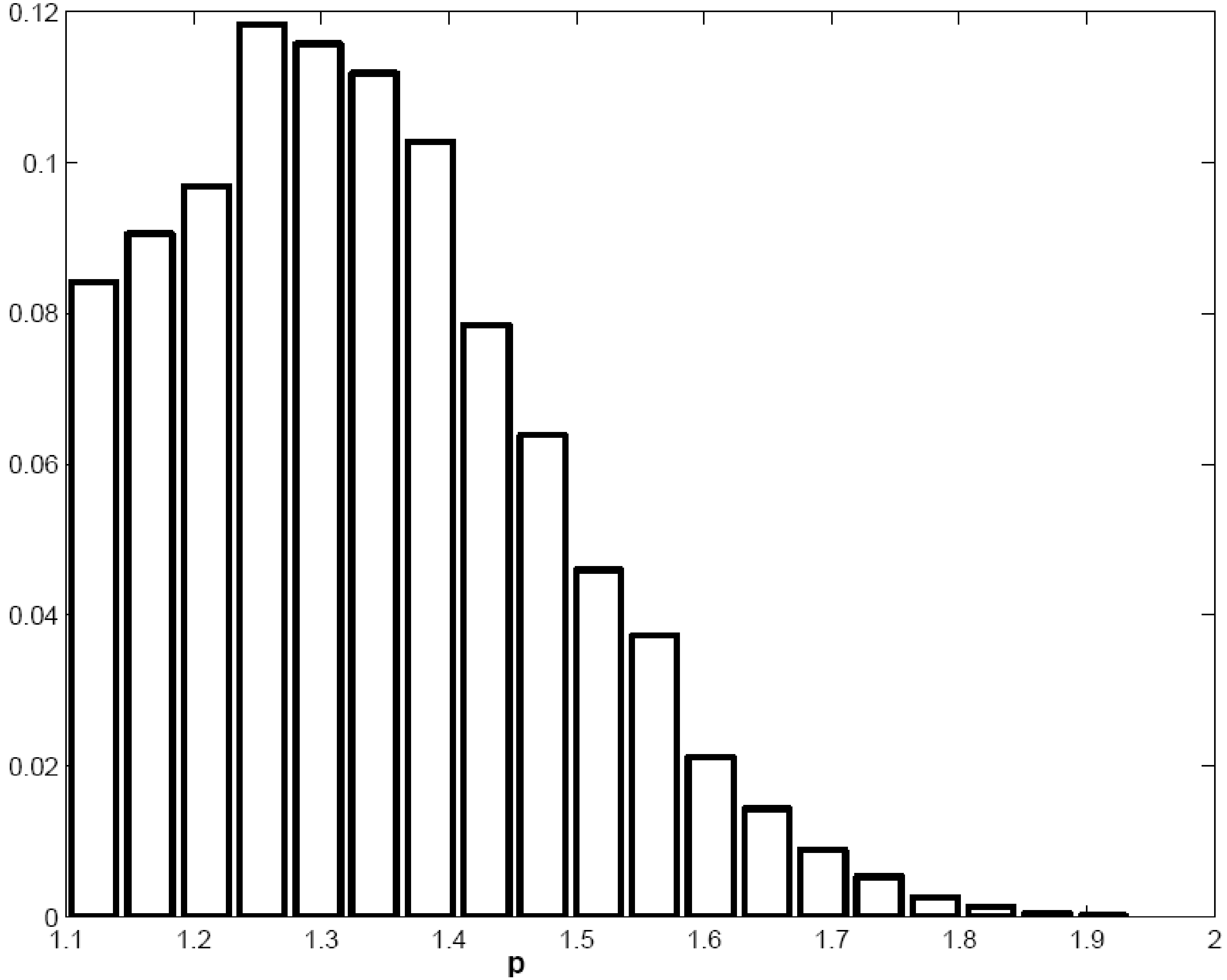


Figure 7: Posterior distribution of the model parameter p.

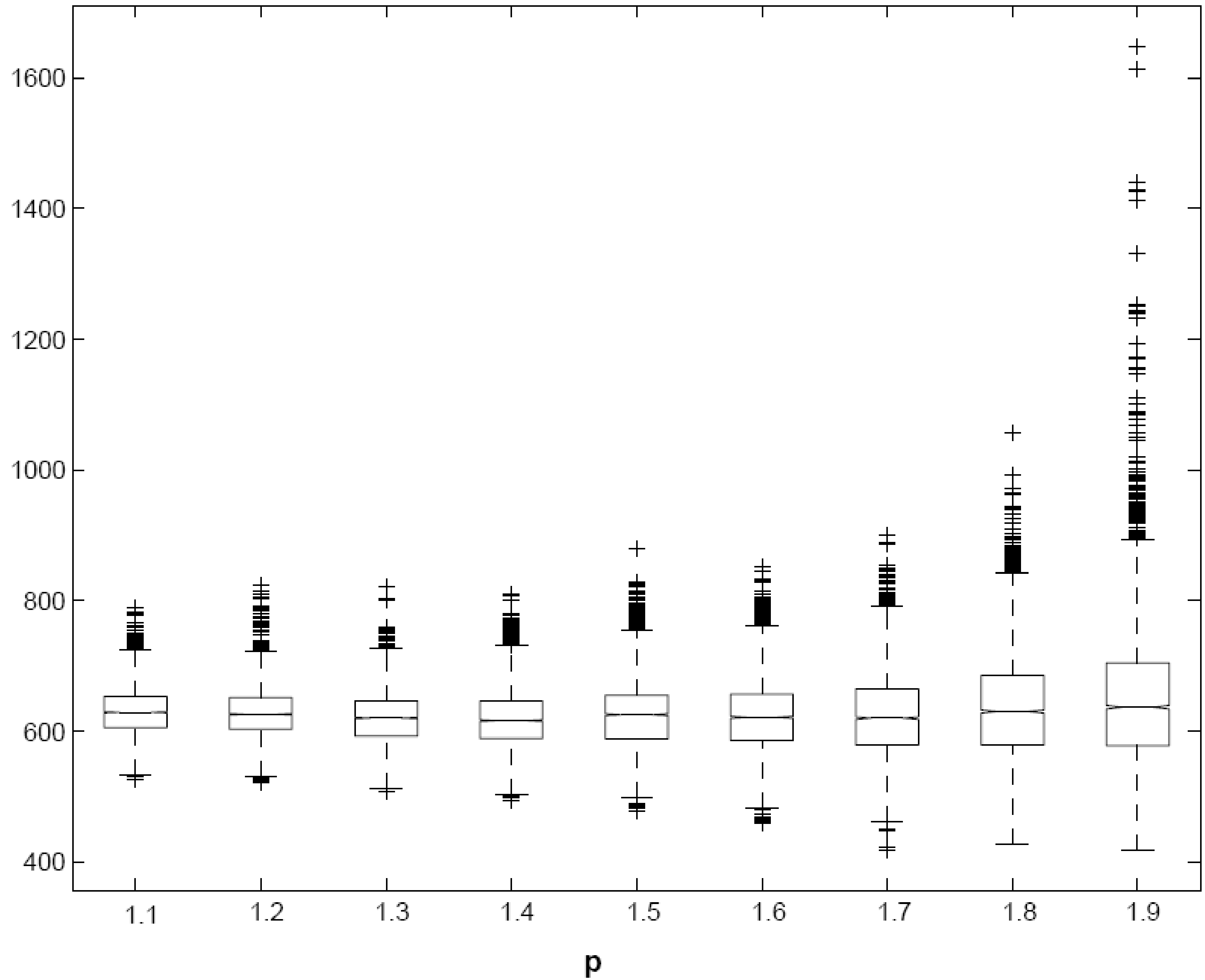


Figure 8: Predicted claim reserves $\widetilde{R}$ distributional summaries

Results: average over *p*

- Claims reserve MLE, $\widehat{R}^{\mathrm{MLE}}$, is less than Bayesian estimate $\widehat{R}^{\mathrm{B}}$
- $\sqrt{EE}$ and $\sqrt{PV}$ are of the same magnitude, 6-7% of total reserve
- MLEs for $\sqrt{EE}$ and $\sqrt{PV}$ are less than Bayesian estimates, $\approx 30\%$
- The difference between $\widehat{R}^{\mathrm{MLE}}$ and $\widehat{R}^{\mathrm{B}}$ is of the order of magnitude as $\sqrt{EE}$ and $\sqrt{PV}$

Results: conditioning on *p*

- MLE of ER_p is almost constant
- Bayesian estimates for ER_p changes as a function of p.
- Bayesian estimators for $\sqrt{PV_p}$ and $\sqrt{EE_p}$ increase as p increases. The MLEs for PV_p and EE_p are significantly less than Bayesian estimators.

Conclusions

- Development of a Bayesian model for claims reserving under Tweedie's compound Poisson model covering range between Poisson and Gamma models
- Quantification process, estimation and model uncertainties and variable selection using MCMC (random walk Metropolis-Hastings within Gibbs)
- MLEs are materially different from Bayesian estimators – posterior distributions are different from Gaussian.

Future work: variable selection problem – Reversible Jump MCMC; considering model parameter p outside the (1, 2) range; dynamic model.

Pavel Shevchenko
Principal Research Scientist
CSIRO Mathematical & Information Sciences

Phone: 02 9325 3218
Email: Pavel.Shevchenko@csiro.au
www.cmis.csiro.au/Pavel.Shevchenko

www.csiro.au

Thank you