BAYESIAN ANALYSIS OF STOCK-RECRUITMENT MODELS IN FISHERIES

MARK CROCKER

Bayesian Analysis of Stock-Recruitment Models in Fisheries

by

© Mark Crocker

A practicum submitted to the

School of Graduate Studies

in partial fulfillment of the

requirements for the degree of

Master of Applied Statistics

Department of Mathematics and Statistics

Memorial University of Newfoundland

December, 2005

St. John's Newfoundland

Your file Votre référence
ISBN: 978-0-494-19354-9
Our file Notre référence
ISBN: 978-0-494-19354-9

Abstract

Bayesian approaches to the analysis of population dynamics are becoming more common in some areas of fisheries research. The search for a stock-recruitment relationship is considered a central problem of fish population dynamics and has been explored in many studies. In this study, posterior distributions are estimated for both the Ricker and Beverton-Holt stock-recruitment curves, using the Normal, Log-normal and Poisson distributions, for both the Baltic Areas 22-24 and NAFO subdivision 3Ps cod stocks. Prior distributions were changed and the sensitivity of the results were investigated. The posterior estimates are sensitive to our choice of prior distribution, which can have an important practical effect on the study of stock-recruitment issues in a number of different fisheries. The Gibbs sampler was used to overcome difficulties in calculating the posterior distributions.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to express my appreciation to my supervisors, Dr. Gary Sneddon and Dr. David Schneider. A special thanks to Dr. Sneddon for all his time, patience and encouragement, it was an honor to have worked with him. Thank you to Dr. Schneider for his helpful comments and suggestions. Also, the comments and time given by my readers, Dr. Hong Wang and Dr. Michael Dowd, have greatly improved and clarified this work. Most importantly, I would like to acknowledge and thank my loving wife, Corina, who has been very supportive and encouraging of all my endeavors.

Chapter 1

Introduction

Mathematical models are commonly used to describe relationships among fisheries variables (Hilborn and Walters, 1992; Quinn and Deriso, 1999). To relate a model to data observed in a fishery, an appropriate method is required to estimate parameters in the model (Chen and Fournier, 1999). In general, there are two statistical approaches that can be used for parameter estimation: frequentist and Bayesian approaches. The statistical problem is similar for both approaches: each is used to make statistical inferences about parameters in the model (Berger, 1985; Box and Tiao, 1992).

Bayesian approaches to the analysis of population dynamics are becoming more common in some areas of fisheries research (Liermann and Hilborn, 1997; Meyer and Millar, 1999; Harley and Myers, 2001; Chen and Holtby, 2002).

The stock and recruitment problem may be considered as the search for the relationship between spawning stock size and the subsequent recruitment in numbers to a year class. This is a central problem of fish population dynamics, since it represents nature's regulation of stock size, whether or not the populations are being exploited.

Understanding the dynamics of recruitment requires full understanding of the dynamics of spawning stock biomass (Ricker 1975). Knowledge of the stock-recruitment (SR) relationship is commonly obtained through quantitative modeling. SR models are mathematical functions that describe relationships between spawning stock abundance and its subsequent recruitment (Jiao *et al.*, 2004a). Many SR models have been developed. Two commonly used SR models are the Ricker (1954) and the Beverton-Holt (1957).

The SR relationship of cod stocks has been explored in many studies (Myers *et al.*, 1995; Myers and Barrowman, 1996). Hilborn and Walters (1992) noted that the estimation of the SR relationship is perhaps the most difficult work in fisheries stock assessment. For many fish stocks, the SR relationships are not clear (Ricker, 1975; Hilborn and Walters, 1992). The shape of the recruitment curve is often hard to determine because of small sample sizes and high variability in recruitment.

Large variations in recruitment have been observed for many fish species (Myers and Barrowman, 1996). The variations in recruitment tend to increase with the spawning stock biomass (Myers *et al.*, 1995), which has led to wide adoption of the assumption that recruitment at a given level of spawning stock biomass follows a lognormal distribution (Hilborn and Walters, 1992). The lognormal distribution has been used as an alternative to the normal distribution (Quinn and Deriso, 1999). The normal error distribution assumption is no longer widely used in SR analysis, although it tends to be more realistic if the survival of individuals during their early life stages is density-independent and constant (Shelton, 1992).

Practical difficulties, induced by the complexity (nonlinearity and high dimensionality) of biologically meaningful models, have limited the implementation of the Bayesian approach until recently. Recent advances in computing power and algorithms for solving complex integrals has contributed to the widespread use of the Bayesian approach. One approach to computing these complex integrals is Markov Chain Monte Carlo (MCMC) methods, which attempt to simulate values from the posterior distribution. These methods include traditional non-iterative methods such as importance sampling (Geweke, 1989) and the simple rejection sampling. More powerful iterative Monte Carlo methods such as data augmentation (Tanner and Wong, 1987), the Metropolis-Hastings algorithm (Hastings, 1970) and the Gibbs sampler are available. By using the above methods, difficult calculations are avoided and are replaced with a sequence of easier calculations.

In Chapter 2, we discuss the rationale behind the Bayesian approach to statistical analysis. Monte Carlo Markov Chains, in particular the Gibbs sampler, are introduced as methods for sampling from posterior distributions. Background is given on the Ricker and Beverton-Holt SR models and an outline of previous Bayes' estimation in fisheries research is given. Chapter 3 introduces the Baltic areas 22-24 data and gives detailed results of the analysis. Chapter 4 introduces the NAFO subdivision 3Ps data and gives detailed results of the analysis. In both Chapters 3 and 4, Bayesian and frequentist results are discussed and compared. One of our main goals is to assess how sensitive the estimates are to the choice of prior distribution. Finally, in Chapter 5, the results of the analysis are discussed and some suggestions are given for future research in this area.

Chapter 2

Background

2.1 Introduction

In this chapter, we start with an introduction to the concepts behind Bayesian analysis and the estimation of posterior distributions through MCMC methods. Next, the Ricker and Beverton-Holt stock-recruitment curves are introduced. Sequential population analysis is discussed and finally reference is made to the varying applications that have used Bayesian analysis in fisheries research.

2.2 Bayesian Modeling

The Bayesian approach to statistics is fundamentally different from the classical approach to statistics. In the classical approach the parameters $\theta = (\theta_1, ..., \theta_M)$ are considered to be unknown, but fixed quantities. A random sample $\mathbf{y} = (y_1, ..., y_n)$ is

drawn from a population and based on the observations in the sample, information about θ is obtained.

In the Bayesian approach, the parameters θ are considered to be random quantities, whose variation can be described by probability distributions $\pi(\theta)$ called prior distributions. A sample is taken from the population and the prior distributions are updated using the information contained in the sample. The updated prior distribution is called the posterior distribution or conditional distribution of θ given **y**. The posterior distribution for θ is given by the following equation:

$$\pi(\theta \mid y)=\frac{\pi(\theta) f(y \mid \theta)}{f(y)} \qquad (2.1)$$

where $f(y \mid \theta)$ is the probability distribution and $f(y)=\int \pi(\theta) f(y \mid \theta) d\theta$ is the marginal distribution of the data y . This formula is known as *Bayes' Theorem* (Carlin and Louis, 1998). Once the posterior distribution of a parameter has been obtained, it can be used to calculate estimates of the parameter.

There are two types of prior distributions: informative and non-informative (or reference) priors. Informative prior distributions summarize the evidence about the parameters, from different sources, and may have a considerable impact on the results. Box and Tiao (1992) define a non-informative prior as one that provides little information that is relative to the experiment. (An overview of the selection of prior distributions by formal rules is given by Kass and Wasserman (1996) and guidelines

on prior specification in fisheries stock assessment are given in Punt and Hilborn (1997).)

Even for models of moderate complexity, the integrals to find $f(y)$ in (2.1) are not solvable in closed form and must be solved numerically. This can be computationally difficult and has been a major impediment towards the widespread use of Bayesian methods. One approach to computing these complex integrals is Markov Chain Monte Carlo (MCMC) methods, which attempt to simulate values from the posterior distribution. These methods include traditional non-iterative methods such as importance sampling (Geweke, 1989) and the simple rejection sampling. More powerful iterative Monte Carlo methods such as data augmentation (Tanner and Wong, 1987), the Metropolis-Hastings algorithm (Hastings, 1970) and the Gibbs sampler (Gelfand and Smith, 1990) produce a Markov chain which represents a (correlated) sample from the joint posterior distribution. By using the above methods, difficult calculations are avoided and are replaced with a sequence of easier calculations.

The Gibbs sampler is a MCMC technique for generating random variables from a distribution indirectly, without having to calculate the density (Casella and George, 1992). The key to the Gibbs sampler is that only univariate conditional distributions are considered. This type of distribution is much easier to simulate than complex joint distributions since they often have simple, known forms, e.g. Normal, Poisson,

Gamma. Thus, n random variables are simulated sequentially from the n univariate conditional distributions rather than a single n-dimensional vector in a single pass from the full joint distribution (Walsh, 2004). This is discussed in more detail below.

Suppose we are given a joint density $f(x, y_1, ..., y_p)$ and we are interested in obtaining characteristics of the marginal density

$$f(x) = \int \cdots \int f(x, y_1, ..., y_p) dy_1 ... dy_p \,. \tag{2.2}$$

This integration would be quite difficult to solve either analytically or numerically. The Gibbs sampler allows for the effective generation of a sample $x_1, \ldots, x_m \sim f(x)$ without computing or approximating $f(x)$ directly.

We illustrate the Gibbs sampler by considering a bivariate random variable *(x, y)*. The Gibbs sampler generates a sample from $f(x)$ by sampling from the conditional distributions $f(x \mid y)$ and $f(y \mid x)$, which are often known statistical models. This is achieved by generating a 'Gibbs sequence' of random variables:

$$Y_0', X_0', Y_1', X_1', \ldots, Y_k', X_k'. \tag{2.3}$$

The Gibbs sampler starts with an initial value for $Y_0' = y_0'$ and obtains x_0 by generating a random variable from the conditional distribution $f(x \mid y = y_0)$. The Gibbs sampler then uses x_0 to generate the value of y_1 by generating a random

variable from the conditional distribution $f(y \mid x = x_0)$. The remainder of the 'Gibbs sequence' is obtained iteratively by alternately generating values from

$$X_j' \sim f(x \mid Y_j' = y_j')$$

$$Y_{j+1}' \sim f(y \mid X_j' = x_j'). \qquad (2.4)$$

Under reasonably general conditions, the distribution of X_k' converges to $f(x)$ as $k \to \infty$ (Casella and George, 1992). Repeating the above process k times generates a 'Gibbs sequence' of length k, where a subset of points (x_i, y_i) for $1 \le i \le m < k$ are the simulated sample from the full joint distribution.

Many software packages are available, free of charge on the Internet, to perform Bayesian analysis. One such package is BUGS, a near-acronym for Bayesian inference Using Gibbs Sampling. This software was developed at the Medical Research Council (MRC) Biostatistics Unit at the University of Cambridge and was initially described by Gilks, Thomas and Spiegelhalter (1994). (BUGS is available from the website http://www.mrc-bsu.cam.ac.uk/work/bugs and comes with complete documentation and two volumes of examples.)

As an illustration of the Gibbs sampler used in the BUGS software, Example 1 from Casella and George (1992) was reproduced using BUGS. The joint distribution of x=0,1,...,n and $0 \le y \le 1$ is given by:

$$f(x,y) \propto \binom{n}{x} y^{x+\alpha-1}(1-y)^{n-x+\beta-1}.$$

The conditional distributions are as follows:

$$f(x \mid y) \sim Binomial(n,y)$$

$$f(y \mid x) \sim Beta(x+\alpha, n-x+\beta).$$

The marginal distribution $f(x)$ can be obtained analytically and is:

$$f(x) = \binom{n}{x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(\alpha+\beta+n)},$$

the beta-binomial distribution. Characteristics of $f(x)$ can be directly obtained from the above formula either analytically or by generating a sample from the marginal distribution without fusing with the conditional distributions. Figure 2.1 shows two samples $x_1,\ldots,x_m$ of size m = 500 from the beta-binomial distribution given above with n = 16, α =2 and β =4. The two histograms are quite similar, illustrating that the Gibbs sampler method implemented in BUGS is indeed generating variables from the appropriate marginal distribution.

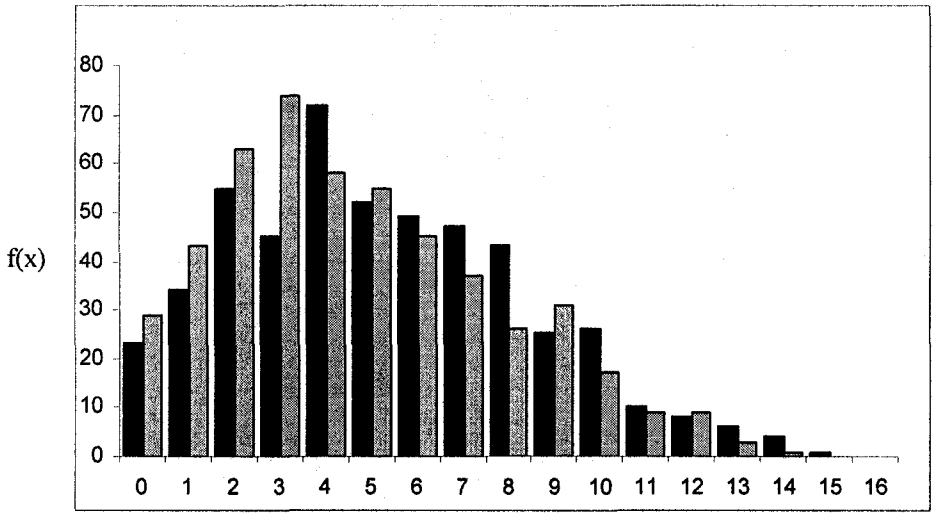


Figure 2.1 Comparison of two histograms of 500 samples each. The black histogram was obtained using the Gibbs sampler from BUGS and the white histogram was obtained directly from the Beta-Binomial distribution with $n = 16$, $\alpha = 2$ and $\beta = 4$.

A key issue in the successful implementation of the Gibbs sampler or any other MCMC method is the number of runs until the chain converges (the length of the burn-in period). A poor choice of starting value and/or prior distribution can greatly increase the required burn-in period. Typically, the first 1000 to 5000 elements are discarded and convergence is assessed by inspecting the time series trace (a plot of the random variables being generated versus the number of iterations). These trace plots can be used to determine a minimum burn-in period. A chain is said to be poorly mixing if it stays in small regions of the parameter space for a long period of time, as opposed to a well mixing chain that seems to explore the space (Walsh, 2004). A poorly mixing chain can arise because the posterior distribution is

multimodal and the choice of starting values traps the chain near one of the modes. The trace plots can also be used to show evidence of poor mixing of the chain.

2.3 Stock-Recruitment Models

A significant amount of research has been conducted on studying recruitment of juvenile fish to a fishable population at various stages in life history. A major issue surrounding recruitment is whether recruitment to the population is mostly due to the size of the spawning stock or environmental conditions. Empirical relationships between spawning stock and recruitment show extreme annual variability (Larkin 1973, Ricker 1975, Getz and Swartzman 1981, Rothschild 1986). Bounds must be placed on recruitment to a population due to limiting factors such as abundance of food, spawning area, rearing area and cannibalism (Quinn and Deriso 1999).

The regenerative process of a population is important to the sustainability of that population. This cycle of regeneration of a fish population can be illustrated as:

Eggs → Larvae → Juveniles → Recruits → Spawners → Eggs → ...

(Quinn and Deriso, 1999).

The abundance at each of these stages can be assumed to be in proportion to the previous stage, in the simplest model. For example, recruitment is proportional to spawning stock, or

$$R = \alpha S . \qquad (2.5)$$

This relationship is called 'density independent' since the ratio of recruitment to spawning stock (R/S) is independent of the population density as measured by spawning stock. Here, α is called the productivity or density-independent parameter. Normally, this relationship is not realistic since R is able to increase without bound as a function of S.

Generally, density-dependent effects are present at some (or all) stages in the life history of a fish population, making the simple model, (2.5), incorrect. Mortality in the population is comprised of both density-independent and density-dependent effects. Density dependent loss can arise from predation, the principal source of mortality in young fish. The population of young fish before recruitment inhibits itself through competition for food (resulting in larval starvation) or space (such as overcrowding of eggs). Using this density dependent assumption the relationship between spawning stock and recruitment can be modeled using the following:

$$R = \frac{\alpha S}{1 + \beta S}, \qquad (2.6)$$

where α is the productivity parameter, representing the number of recruits per spawner at low numbers of spawners and β controls the level of density dependence. This model is known as the Beverton-Holt stock-recruitment curve and was first derived by Beverton and Holt (1957). This relationship is a strictly increasing function of spawners, which approaches the asymptote

$$R_p = \frac{\alpha}{\beta}, \tag{2.7}$$

which is the maximum recruitment.

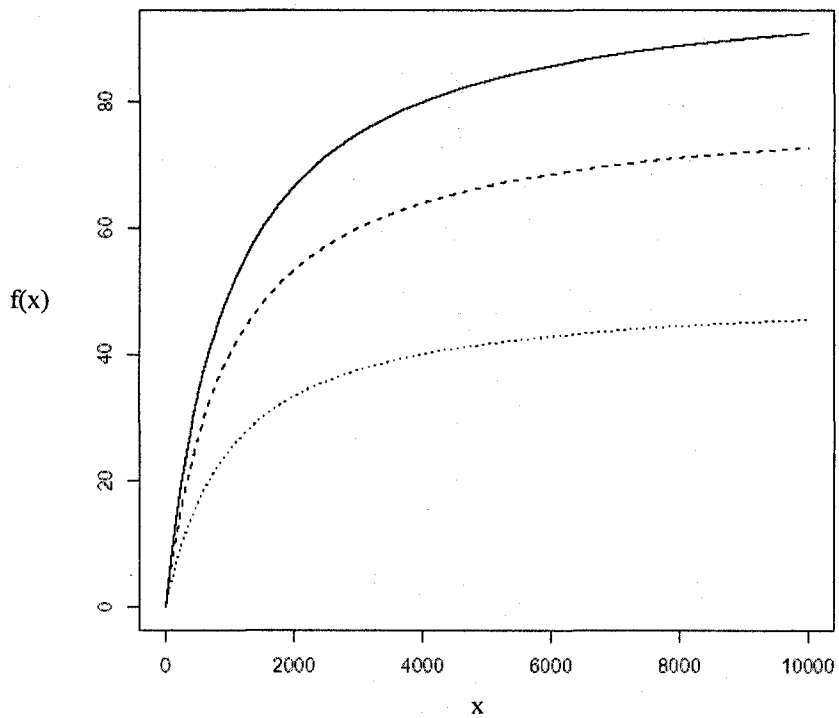


Figure 2.2 The Beverton-Holt stock-recruitment curve. Three curves are shown, corresponding to different values of the productivity parameter ($\alpha = 0.1$ is the solid line, $\alpha = 0.08$ is the dashed line and $\alpha = 0.05$ is the dotted line) and a constant value for $\beta = 0.001$.

Now, suppose that the spawning stock inhibits the population of young fish before recruitment (such as through cannibalism). Again, mortality in the population is made up of both density-independent and density-dependent effects. Using these assumptions, the relationship between spawning stock and recruitment can be represented by:

$$R = Se^{\alpha - \beta S}, \qquad (2.8)$$

where α is the productivity parameter, representing the number of recruits per spawner at low numbers of spawners and β controls the level of density dependence. This model is called the Ricker stock-recruitment curve and was first derived by Ricker (1954). This relationship is dome-shaped as a function of spawners, with a maximum at the point

$$(R_p, S_p) = \left(\frac{1}{\beta} e^{\alpha - 1}, \frac{1}{\beta}\right). \qquad (2.9)$$

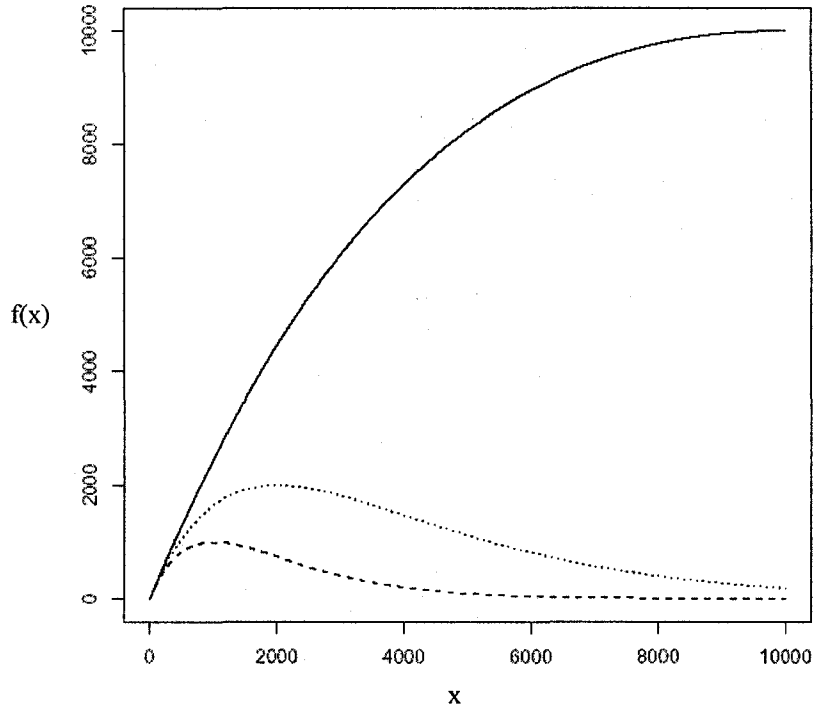


Figure 2.3 The Ricker stock-recruitment curve. Three curves are shown, corresponding to different values of the density dependence parameter ($\beta = 0.0001$ is the solid line, $\beta = 0.001$ is the dashed line and $\beta = 0.0005$ is the dotted line) and a constant value for $\alpha = 1$.

Many other stock-recruitment models exist. For example, the Ludwig-Walters model (Ludwig and Walters, 1989), the Cushing model (Cushing, 1973), the Deriso-Schnute model (Deriso, 1980; Schnute, 1985) and the Sheppard model (Sheppard, 1982). The focus of this analysis is on two commonly used stock-recruitment models, the Ricker and the Beverton-Holt as described above.

2.4 Sequential Population Analysis

For marine populations, population numbers and fishing mortality are usually estimated using sequential population analysis (SPA) of commercial catch-at-age data. SPA techniques include virtual population analysis (VPA), cohort analysis, and related methods that reconstruct population size from catch-at-age data (Hilborn and Walters, 1992).

The commercial catch-at-age data are combined with estimates from research surveys and commercial catch rates to estimate the numbers at age in the final year and to reconstruct previous numbers at age under the assumption that commercial catch at age is known without error and that natural mortality at age is known and constant (Myers *et al.*, 2002).

A major source of uncertainty in the SPA estimates of recruitment and spawning stock biomass is that catches are assumed known without error. This is important when estimates of discarding and misreporting are not included in the catch-at-age data used in the SPA. These errors are clearly important for some of the Atlantic cod stocks during certain periods (Myers *et al.*, 1997) and will affect estimates of replacements each spawner can produce at low population densities. Other sources of error are due to the survey methodology and the distribution of the fish population.

Both data sets used in this analysis were estimated using SPA techniques.

2.5 Bayesian Analysis of Fisheries Problems

Bayesian analysis methods are gaining popularity in fisheries research in part because they lend themselves to input from disparate sources and in part due to recent advances in computing algorithms and power. Numerous recent papers are available dealing with topics ranging from state-space model implementation of the delay-difference model (Meyer and Millar, 1999) to estimating salmon escapement goals (Adkinson and Peterman, 1996), which utilize the Bayesian approach to data analysis.

Meyer and Millar (1999) used a Bayesian approach to fit the delay-difference and surplus production models, since it could easily handle realistic distributional assumptions as well as the nonlinearities in the equations.

Liermann and Hilborn (1997) calculated a prior distribution for a depensation parameter that was used in the analysis of other similar fish stocks. Bayesian methods were used to incorporate the two sources of variability (measurement error and between stock variability) so that the distribution of the depensatory parameter could be isolated.

Chen and Holtby (2002) used a Bayesian framework to develop a regional stock recruitment meta-model that combined information from multiple fish stocks. Sample data was used to estimate the unknown regional parameters.

Millar (2002) addressed the problem of specifying default priors for several common fisheries models, the Ricker and Beverton-Holt stock recruitment curves, the von Bertalanfly growth curve, the Schaefer surplus production model and sequential population analysis.

Harley and Myers (2001) estimated catchability during research trawl surveys. The hierarchical Bayes model provided a more reliable estimator under a wide range of conditions.

Robb and Peterman (1998) developed a decision-making framework for management of a sockeye salmon fishery that explicitly accounts for uncertainties in the stock recruitment relationship, annual recruitment, run timing and catchability.

Adkinson and Peterman (1996) estimated optimal escapement goals for salmon, using both knowledge of the physical determinants of salmon productivity and stock recruitment data. A Bayesian approach allowed for the integration of information from diverse sources and provided a framework for decision-making that took into account the uncertainty reflected in that data.

Chen and Fournier (1999) evaluated the impact of outliers on the derivation of the posterior distribution in Bayesian analysis using a simple growth curve.

McAllister and Ianelli (1997) estimated population model parameters using catch-age data and indices of relative abundance for yellow fin sole in the Bering Sea. The example illustrates how catch-age data can markedly improve estimates via Bayesian methods.

Rivot et al. (2001) warned that in stock recruitment analysis, as in many applications, the Bayesian posterior inferences can be very sensitive to the choice of prior distribution. It has been suggested that a comprehensive sensitivity analysis be conducted for different specification of prior distribution to ensure a robust result (Chen and Holtby, 2002; Punt and Hilborn, 1997; Millar, 2002).

2.6 Research Objectives

Fisheries data have been frequently analyzed as if errors are normally, identically and independently distributed (Chen and Fournier, 1999). Jiao *et al.* (2004a, 2004b) introduce non-normal errors into the analysis of stock-recruitment data. In this study, we incorporate prior knowledge for both the Ricker and Beverton-Holt stock-recruitment curves, using the Normal, Lognormal and Poisson distributions.

Each model was fitted using the BUGS software package on cod (*Gadus morhua*) stock-recruitment data from the Baltic Areas 22-24 and the NAFO subdivision 3Ps. These two cod stocks were chosen because of the marked contrast in the environmental conditions, ranging from a highly brackish system with weak advective regime (Baltic) to a more oceanic system with stronger advective regime (3Ps). For each model, the prior distributions were changed and the sensitivity of the results were investigated. If the posterior estimates are sensitive to our choice of prior distribution, this has an important practical effect on the study of stock-recruitment issues in a number of different fisheries. The Deviance Information Criterion (DIC) (Spiegelhalter *et al*, 2002) was calculated for each model as a measure of model fit.

Chapter 3

Baltic Areas 22-24 Data

3.1 Background

The Baltic Areas 22-24 are located in the Baltic Sea. The Baltic Sea is a semi-enclosed sea bordered by nine countries as depicted in the center of the map below (Figure 3.1).

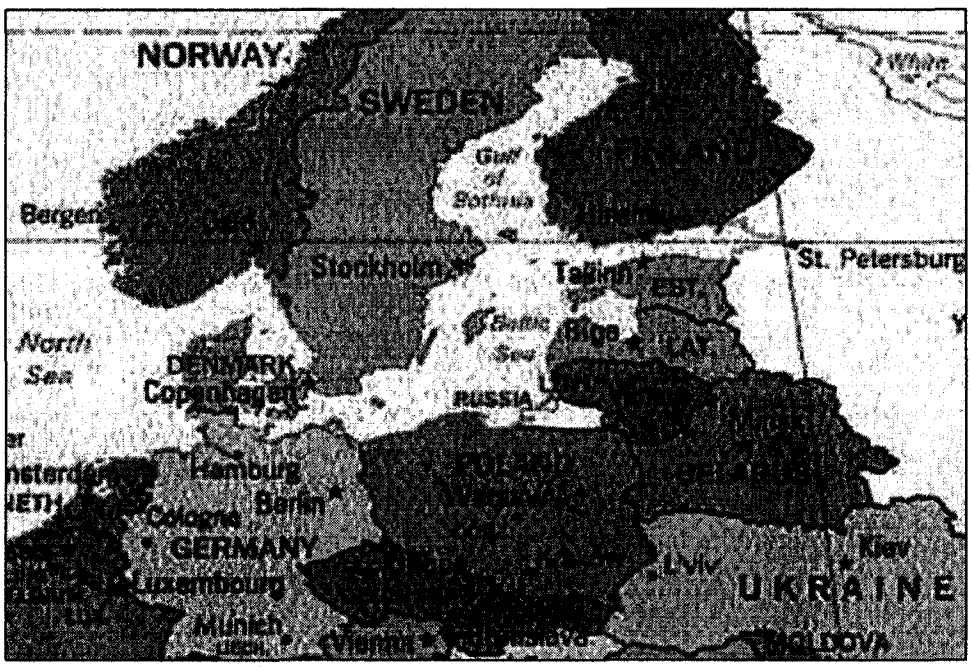


Figure 3.1 Map of the Baltic Sea

There are approximately 100 species of fish, many of commercial importance including cod (*Gadus morhua L.*) in the Baltic Sea. The cod are separated into two stocks: the western stock, located in ICES fishing areas 22-24 and the eastern stock, located in ICES areas 25-32 (http://oceanides.jrc.cec.eu.int/Baltic%20Sea%20Web%20Page/baltic_sea_environment.htm, May 15, 2005). The following figure shows the ICES fisheries areas in the Baltic Sea (Areas 22-24 are towards the bottom right).

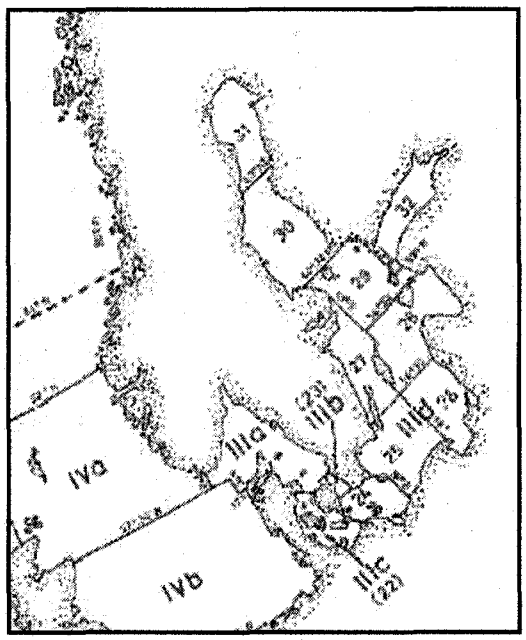


Figure 3.2 Map of the ICES fishing areas in the Baltic Sea

3.2 Exploratory Data Analysis

Analysis of the Baltic Areas 22-24 data began with a look at some exploratory plots (Figure 3.3). The first plot shows the relationship of recruitment to spawning stock biomass. For the most part, recruitment increases as spawning stock increases. The second plot shows the autocorrelation plot for recruitment. The 95% confidence limits about zero are included on the plot. Autocorrelation in recruitment is weak and hence was not of concern.

The third plot is of recruitment vs. time. This plot shows a very definite decrease in the counts of recruits over time. The fourth plot is of spawning stock biomass vs.

time. Spawning stock biomass is constant for the first 15 years before it decreases dramatically.

3.3 Ricker Model

3.3.1 Initial Results

The Ricker model (2.8) was fitted in BUGS on the Stock-Recruitment data from the Baltic Areas 22-24. We considered modeling recruitment using both normal and lognormal distributions. Four combinations of recruitment and stock were initially considered, where we used either the original data or transformed data as follows: R & S, R & log S, log R & S, and log R & log S.

The following prior distributions were assigned: $\alpha \sim N(0,1)$, $\beta \sim N(0,100)$ and $\tau \sim Gamma(1,1)$, where $\tau = \frac{1}{\sqrt{\sigma}}$, which is commonly referred to as the precision. The Gamma distribution is of the following form:

$$f(x,\lambda,\theta) = \frac{x^{\lambda-1} e^{\frac{-x}{\theta}}}{\Gamma(\lambda)\theta^{\lambda}}. \qquad (3.1)$$

Each model was run for a burn-in period of 5,000 iterations, followed by an additional 10,000 iterations. The same initial starting values for the MCMC were used for each model (α=0.2, β=2.0, τ=1.0). The results are given in Table 3.1.

Table 3.1 Initial results for the Ricker S-R model using the prior distributions: $\alpha \sim N(0,1)$, $\beta \sim N(0,100)$ and $\tau \sim Gamma(1,1)$

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
R	S	-0.02199	2.497	0.00912	10.82
		(0.9899)	(1.894)	(0.00267)	(1.665)
R*	log S*	-0.1699	-0.02138	1.952	0.7411
		(0.194)	(0.01869)	(0.5883)	(0.1166)
log R	S	-0.02199	2.497	0.1883	2.382
		(0.9899)	(1.894)	(0.05503)	(0.3664)
log R*	log S*	-0.6377	0.0801	10.91	0.3131
		(0.3372)	(0.03263)	(3.236)	(0.0488)

* 10,000 iteration burn-in period with an additional 10,000 iterations (due to visual evidence of non-convergence)

For the normal case (R and S) (Figure 3.4), the posterior distribution of α was symmetric, with a mean of -0.2199 and a standard deviation of 0.9899. The posterior distribution of α was virtually unchanged from the prior distribution. The posterior distribution of β was highly skewed to the right with a mean of 2.497 and standard deviation of 1.894. This was a substantial departure from the prior distribution of N(0, 100). Autocorrelation was present for the first 7 lags for α and the first 10 lags for β. The posterior distribution of τ was slightly skewed to the right, with mean 0.00912 and standard deviation of 0.00267. The posterior distribution of σ was also slightly skewed to the right ($\sigma = \frac{1}{\sqrt{\tau}}$). The mean and standard deviation of σ were 10.82 and 1.665 respectively. Autocorrelation in τ and σ was weak and hence was not of concern.

For the second case (R and log S) (Figure 3.5), the posterior distribution of α was skewed to the left. Its mean and standard deviation were -0.1699 and 0.194 respectively. This departs from the prior distribution of N(0, 1). The posterior distribution of β was also skewed to the left, with a mean of -0.02138 and standard deviation of 0.01869. Severe autocorrelation was present for both α and β. The posterior distributions may not be correct since the chain has not mixed and thoroughly explored the parameter space. The posterior distributions for τ and σ were both skewed to the right. Autocorrelation in τ and σ was weak and hence was not of concern.

For the lognormal case (log R and S) (Figure 3.6), the posterior distribution of α changed little relative to the prior distribution. It was still centered close to 0 (-0.02199), with a standard deviation of approximately 1 (0.9899). The posterior distribution of β changed little from the normal case (R and S). Autocorrelation was present for the first 7 lags for α and the first 10 lags for β as before. The posterior distribution of τ was approximately symmetric with a mean of 0.1833 and standard deviation of 0.05503. The posterior distribution of σ was approximately symmetric with mean 2.382 and standard deviation of 0.3664. Autocorrelation in τ and σ was weak and hence was not of concern.

For the fourth case (log R and log S) (Figure 3.7), the posterior distributions of α and β appeared to be bimodal. Severe autocorrelation was present for α and β.

The posterior distributions may not be correct since the chain has not mixed and thoroughly explored the parameter space. This interpretation holds for all subsequent posterior distributions. The posterior distribution of τ was approximately symmetric with a mean of 10.91 and standard deviation of 3.236. The posterior distribution of σ was slightly skewed to the right. Autocorrelation in τ and σ was weak and was not of concern.

The prior distributions were changed to $\alpha \sim N(10,1)$, $\beta \sim N(0,1)$ and $\tau \sim Gamma(1,0.2)$. Once again, four combinations were considered (R & S, R & log S, log R & S, and log R & log S). Each model was run for a burn-in period of 5,000 iterations, followed by an additional 10,000 iterations. The same initial starting values were used for each model (α=0.2, β=2.0, τ=1.0). The results are given in Table 3.2. The kernel density plots of the posterior distributions and autocorrelation plots have similar shapes to those in Figures 3.4 – 3.7.

Table 3.2 Results for the Ricker S-R model using the prior distributions: $\alpha \sim N(10,1)$, $\beta \sim N(0,1)$ and $\tau \sim Gamma(1,0.2)$

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
R	S	9.97	0.7952	1.76e-10	77910.0
		(0.9871)	(0.6105)	(1.125e-10)	(11980.0)
R	log S	9.059	0.02643	7.03 e-10	39060.0
		(0.9216)	(0.09046)	(2.348 e-10)	(6215.0)
log R	S	9.985	0.8034	0.009091	10.840
		(0.9846)	(0.6093)	(0.00266)	(1.667)
log R	log S	5.796	0.5617	0.1396	2.855
		(0.8195)	(0.08215)	(0.05799)	(0.6176)

For the normal case, R and S, the posterior distribution of α was symmetric, with a mean of 9.97 and a standard deviation of 0.9871. Again, the posterior distribution of α was virtually unchanged from the prior distribution. The posterior distribution of β was highly skewed to the right with a mean of 0.7952 and standard deviation of 0.6105, which was a departure from the prior distribution of N(0, 1). Autocorrelation was present for the first 7 lags for α and the first 10 lags for β. The posterior distribution of τ was slightly skewed to the right, but was approximately symmetric, with a mean and standard deviation close to 0 ($1.76e^{-10}$ and $1.125e^{-10}$ respectively). Since $\tau = \frac{1}{\sqrt{\sigma}}$, we find the posterior distribution of σ was also slightly skewed to the right with a very large mean (77190.0) and standard deviation (11980.0). Autocorrelation in τ and σ was weak and hence was not of concern.

For the second case, R and log S, the posterior distribution of α was symmetric, with a mean of 9.059 and a standard deviation of 0.9871, which was close to the prior distribution of N(10, 1). The posterior distribution for β was highly skewed to the right with a mean of 0.7952 and standard deviation of 0.6105, which was a departure from the prior distribution. Autocorrelation was present for the first 7 lags for α and the first 10 lags for β. The posterior distributions of τ and σ were highly skewed to the right. Autocorrelation in τ and σ was weak and hence was not of concern.

For the lognormal case, log R and S, the posterior distributions of α and β were very similar to those obtained in the normal case (R and S). Autocorrelation was still present for the first 7 lags for α and the first 10 lags for β. The posterior distribution of τ was approximately symmetric with a mean of 0.009091 and standard deviation of 0.00266. The posterior distribution of σ was approximately symmetric with mean 10.84 and standard deviation of 1.667. Autocorrelation in τ and σ was weak and hence was not of concern.

For the fourth case, log R and log S, the posterior distributions of α and β were different than their respective prior distributions. Severe autocorrelation was present for both α and β. The posterior distributions of τ and σ were skewed to the right. Autocorrelation was weak and hence was not a concern with these parameters. The log transformation of the spawning stock biomass was not appropriate for this application and thus was not investigated any further.

The prior distributions were changed to $\alpha \sim N(0,1000)$, $\beta \sim N(0,500)$ and $\tau \sim Gamma(1,0.2)$. Each model was run for a burn-in period of 5,000 iterations, followed by an additional 10,000 iterations. The same initial starting values were used for each model (α=0.2, β=2.0, τ=1.0). The results are given in Table 3.3. The kernel density and autocorrelation plots were of similar shape to those in Figures 3.4 – 3.7 respectively.

Table 3.3 Results for the Ricker S-R model using the prior distributions: $\alpha \sim N(0,1000)$, $\beta \sim N(0,500)$ and $\tau \sim Gamma(1,0.2)$

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
R	S	-0.1314	3.817	1.76 e-10	77910.0
		(5.553)	(2.888)	(1.125 e-10)	(11980.0)
log R	S	-0.04664	3.819	0.009096	10.84
		(5.528)	(2.901)	(0.002661)	(1.668)

For the normal case (R and S), the posterior distributions of τ and σ were very similar to those obtained in the previous example. The posterior distribution of α was symmetric, with a mean and standard deviation that was substantiality different from the prior distribution (-0.1314 and 5.553 respectively). The posterior distribution of β was highly skewed to the right, with mean 3.817 and standard deviation of 2.888. Again, autocorrelation was present for the first 7 lags for α and the first 10 lags for β.

For the lognormal case (log R and S), the posterior distributions of α and β were very similar to those obtained in the first case (R and S). Autocorrelation was still present for the first 7 lags for α and the first 10 lags for β. The posterior distributions of τ and σ were very similar to those obtained in the previous example. Autocorrelation in τ and σ was weak and hence was not of concern.

3.3.2 NLS Estimates for the Ricker Model

The Ricker S-R model was fitted to the Baltic Areas 22-24 data using non-linear least squares (NLS). The NLS procedure calculated maximum likelihood estimates (MLE). The likelihood function is maximized when the sum of squared residuals ,

$$\sum_{i=1}^{n}\left(R_i - S_i e^{\alpha-\beta S_i}\right)^2, \qquad (3.2)$$

is minimized. The solutions to this equation are obtained by an iterative procedure.

The estimates obtained from NLS were used to find non-informative prior distributions for α, β and τ in BUGS.

Four sets of initial values for α and β were chosen in order to assess the behavior of the NLS function. The first set of values was obtained by taking the maximum (R, S) pair (99329, 100362.78) and using the fact that the maximum of the Ricker curve is at the point $\left(\frac{1}{\beta}e^{\alpha-1}, \frac{1}{\beta}\right)$. The fourth set of values was obtained similarly except using the minimum (R, S) pair (22835, 32710.72).

1) $\alpha = 11.65$, $\beta = 0.00002900$
2) $\alpha = 0$, $\beta = 0$
3) $\alpha = 5$, $\beta = 0.001$
4) $\alpha = 1.908$, $\beta = 0.00005845$

The results for all four combinations of Stock and Recruitment are given in Table 3.4.

Table 3.4 NLS estimates for the Ricker S-R model

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. error)	$\hat{\beta}$ (std. error)	$\hat{\tau}$	$\hat{\sigma}$
R	S	0.7299 (0.5820)	7.277 e-07 (1.470 e-05)	7.677 e-10	36092.2
R	log S	-0.0893 (4.811)	-0.8516 (0.4563)	7.678 e-10	36090.1
log R	S	-6.8935 (0.09247)	3.2178 e-05 (2.591 e-06)	0.6273	1.2626
log R	log S	0.2550 (0.3422)	0.01947 (0.03297)	2.2154	0.6719

Similar results were obtained for all four sets of initial starting values indicating that this function was not sensitive to the choice of the starting values. The estimates for α, β, τ and σ obtained using NLS, were significantly different than those obtained previously in BUGS.

3.3.3 Results for the Normal and Lognormal Distributions

In order to replicate these estimates in BUGS, non-informative prior distributions need to be assigned to α, β and τ. The following prior distributions were selected for α, β and τ for the normal and lognormal cases by trial and error.

i) S & R: $\alpha \sim N(0,10000)$

$\beta \sim N(0,10^{-12})$

$\tau \sim Gamma(1,0.2)$

ii) S & log R: $\alpha \sim N(0,100)$

$\beta \sim N(0,10^{-8})$

$\tau \sim Gamma(1,0.2)$

The estimates obtained using these prior distributions were very similar to those obtained using NLS (Table 3.5). Severe autocorrelation was present in both cases (Figure 3.8 and Figure 3.9). Running each model for a longer burn-in period (100,000) did not correct this.

Table 3.5 Estimates of Ricker S-R Model using the prior distributions: $\alpha \sim N(0,10000)$, $\beta \sim N(0,10^{-12})$ and $\tau \sim Gamma(1,0.2)$ for the Normal case and $\alpha \sim N(0,100)$, $\beta \sim N(0,10^{-8})$ and $\tau \sim Gamma(1,0.2)$ for the lognormal case

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
R*	S*	0.7222 (0.5141)	1.333 e-6 (1.305 e-5)	8.443 e-10 (2.744 e-10)	35660.0 (5697.0)
log R	S	-6.902 (0.08246)	3.197 e-5 (2.239 e-6)	0.5324 (0.1599)	1.419 (0.2248)

* 15,000 iteration burn-in period with an additional 10,000 iterations, due to visual evidence of non-convergence

The deviance information criterion (DIC) can be used as a measure of model fit or adequacy that depends on both the prior distributions of the parameters and the data

that are observed. The DIC is defined as a Bayesian measure of fit, plus twice the effective number of parameters (Spiegelhalter *et al.*, 2002). The Bayesian measure of model fit is the posterior mean deviance, $D(\bar{\theta})$, defined as

$$D(\bar{\theta}) = -2\log\{p(y \mid \bar{\theta})\} + 2\log\{f(y)\},$$

where $\bar{\theta} = E(\theta \mid y)$ is the posterior mean of the parameters. The effective number of parameters, p_D, can be estimated by

$$p_D = E_{\theta|y}[-2\log\{p(y \mid \theta)\}] + 2\log\{p(y \mid \bar{\theta})\}.$$

Thus the DIC is defined as

$$DIC = D(\bar{\theta}) + 2p_D.$$

The DIC value for each of the above models is given below.

1) S & R: 527.817
2) S & log R: 76.693

The initial starting values for the MCMC were changed (α=10, β=10, τ=10) to see if the choice of the initial values had an impact on the estimates. Similar results were obtained for both models, but convergence of both chains was much slower. Each chain was initially run for a 5,000 iteration burn-in period followed by an additional 10,000 iterations. Then, the chain was rerun, adding an additional 10,000 iterations to the burn-in period each time, until the chain converged. The details and estimates for the lognormal case are given in Table 3.6. Similar results were obtained for the normal case. After the 45,000 iteration burn-in period, the estimates are very

close to those obtained in Table 3.5. The kernel density plots of the posterior distributions and the autocorrelation plots obtained had similar shapes to those in Figures 3.8 and 3.9.

Table 3.6 Illustration of convergence of Ricker S-R Model estimates using different starting values

Iteration	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
5,001 – 15,000	-2.12 (3.77)	0.005575 (0.006201)	0.04579 (0.05926)	8.28 (3.91)
15,001 – 25,000	-6.193 (0.004312)	5.471 e-5 (1.009 e-10)	0.1418 (0.0415)	2.745 (0.4225)
25,001 – 35,000	-6.232 (0.06416)	5.36 e-5 (3.269 e-7)	0.1491 (0.04479)	2.682 (0.4226)
35,001 – 45,000	-6.389 (0.1561)	4.745 e-5 (5.277 e-6)	0.2526 (0.1194)	2.172 (0.544)
45,001 – 55,000	-6.569 (0.03143)	4.184 e-5 (1.879 e-7)	0.3637 (0.1056)	1.713 (0.2609)

The prior distribution for τ was changed to investigate how sensitive the estimates are to this change. The following prior distributions were used. Details are given for the lognormal case only (Table 3.7). Similar results were obtained for the normal case. For all of the prior distribution specifications, the estimates changed very little, suggesting that the prior distribution for τ has little effect on the parameter estimates. A burn-in period of 5,000 iterations followed by an additional 10,000 iterations was used in each case. Plots of the density functions for each of these Gamma distributions is given in Figure 3.10. (The plot in the top left hand corner is of the Gamma(1, 1) distribution. The plot in the top right corner is of the Gamma(0.01, 100) distribution. The rest of the plots are ordered similarly.)

i) $\tau \sim Gamma(1,1)$

ii) $\tau \sim Gamma(0.01,100)$

iii) $\tau \sim Gamma(1000,0.001)$

iv) $\tau \sim Gamma(1000,1)$

v) $\tau \sim Gamma(5,0.2)$

vi) $\tau \sim Gamma(1,0.001)$

The kernel density plots of the posterior distributions and the autocorrelation plots were similar to those shown in Figure 3.9.

Table 3.7 Effects of changing the prior distribution of τ

Prior Distribution of τ	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
$\tau \sim Gamma(1,1)$	-6.859 (0.06715)	3.32 e-5 (1.821 e-6)	0.6591 (0.1996)	1.276 (0.2035)
$\tau \sim Gamma(0.01,100)$	-6.894 (0.02669)	3.218 e-5 (7.495 e-7)	6.905 (0.652)	0.3818 (0.01814)
$\tau \sim Gamma(1000,0.001)$	-6.91 (0.08072)	3.175 e-5 (2.268 e-6)	0.9942 (0.03128)	1.003 (0.01579)
$\tau \sim Gamma(1000,1)$	-6.894 (0.009875)	3.216 e-5 (2.78 e-7)	59.62 (1.875)	0.1296 (0.002039)
$\tau \sim Gamma(5,0.2)$	-6.38 (0.04444)	4.844 e-5 (4.904 e-7)	0.2959 (0.07745)	1.887 (0.2561)
$\tau \sim Gamma(1,0.001)$	-0.9399 (3.547)	0.007156 (0.00619)	0.005716 (0.002225)	13.89 (2.464)

The prior distribution for β was changed in order to investigate the impact on the estimates and to show that the precision of β needs to be very large (1,000,000 for the normal case and 10,000 for the lognormal case). The prior distributions for α

and τ remained unchanged from the above models (Table 3.5). Each model was rerun using larger variances (smaller precision) for β as indicated below. Details are given for the lognormal case only. Similar results were obtained for the normal case.

i) $\beta \sim N(0,10^{-6})$

ii) $\beta \sim N(0,10^{-4})$

iii) $\beta \sim N(0,10^{-2})$

The results of these changes in the specification of the prior distribution for β are summarized in Table 3.8.

Table 3.8 Effects of changing the prior distribution for β

Prior Distribution of β	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
$\beta \sim N(0,10^{-6})$	-0.1101 (3.166)	0.02543 (0.01928)	0.009096 (0.002658)	10.84 (1.667)
$\beta \sim N(0,10^{-4})$	-0.08888 (3.12)	0.08073 (0.06067)	0.009095 (0.002659)	10.84 (1.667)
$\beta \sim N(0,10^{-2})$	-0.03443 (3.085)	0.257 (0.1927)	0.009092 (0.002657)	10.84 (1.667)

Changing the prior distribution of β has a strong effect on the resulting estimates. These results strongly suggest that the estimates are very sensitive to the choice of the prior distribution of β and that the precision of β needs to be large.

The prior distribution for α was changed in order to investigate the sensitivity and hence required precision in specifying α. The prior distributions for β and τ remained unchanged from the above model (Table 3.5). The model was rerun using the different variances for α as indicated below. Again, results are given for the lognormal case (Table 3.9). Similar results were obtained for the normal case.

i) $\alpha \sim N(0,0.01)$

ii) $\alpha \sim N(0,1)$

iii) $\alpha \sim N(0,10000)$

iv) $\alpha \sim N(0,10^6)$

v) $\alpha \sim N(0,10^8)$

vi) $\alpha \sim N(0,10^{10})$

Table 3.9. Effects of changing the prior distribution for α

α	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
$\alpha \sim N(0,0.01)$	-0.01206 (0.3084)	0.008244 (0.005965)	0.009107 (0.002664)	10.83 (1.665)
$\alpha \sim N(0,1)$	-0.03378 (0.9811)	0.0083 (0.005942)	0.009106 (0.002661)	10.83 (1.665)
$\alpha \sim N(0,10000)$	-6.908 (0.1028)	3.183 e-5 (2.88 e-6)	0.5284 (0.1604)	1.425 (0.2279)
$\alpha \sim N(0,10^6)$	-6.904 (0.106)	3.198 e-5 (3.008 e-6)	0.5272 (0.1617)	1.4727 (0.2283)
$\alpha \sim N(0,10^8)$	-6.873 (0.1556)	3.287 e-5 (4.57 e-6)	0.4968 (0.1617)	1.478 (0.2545)
$\alpha \sim N(0,10^{10})$	-254.3 (213.7)	0.005645 (0.007779)	0.009094 (0.002659)	10.84 (1.668)

When the variance of α was greater than or equal to 100, the estimates were not sensitive to the choice of prior distribution for α. If the variance of α was less than 100, the estimates were very sensitive to the choice of prior distribution for α.

3.3.4 ML Estimates for the Ricker Model

The Poisson distributions are a first approximation to counts of organisms, recognizing that aggregating organisms may well be fit by distributions allowing for overdispersion, such as negative binomial or gamma. In order to calculate estimates of the Ricker model, using the Poisson distribution, we need to maximize the Poisson log-likelihood. This is identical to minimizing the negative of the Poisson log-likelihood. This minimization was performed in R using the non-linear minimization function (NLM).

The Poisson log-likelihood for the Ricker model can be written in the following form.

$$l(\alpha,\beta)=\sum_i\left[-S_i e^{\alpha-\beta S_i}+\alpha R_i-\beta R_i S_i\right]. \qquad (3.3)$$

The minimization function was sensitive to the choice of starting values for α and β. The following pairs of starting values were investigated.

i) $\alpha=0,\ \beta=0$

ii) α=5, β=5

iii) α=10, β=10

iv) α=20, β=0.007

The following estimates were obtained (Table 3.10).

Table 3.10 Poisson ML Estimates

Starting Values (α, β)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)
(0, 0)	5.286 e-10	-1.796 e-5
(5, 5)	0.9507	7.147 e-6
(10, 10)	0.9679	-3.251 e-6
(20, 0.007)	1.135	1.164 e-5

The Poisson model would not run in BUGS since the counts of stock recruitment were so large (up to 147,000). Spawning stock biomass also has very large values (up to 50,000 tonnes). For means approaching 10, there is little practical difference between the Poisson and normal distribution. With a count of 147,000 per unit, the mean will either exceed 10, or the dispersion will be so huge as to preclude the Poisson distribution, where the variance is fixed as the mean value. Exploring the Poisson distribution at a different unit size is, however, interesting.

The original Stock-Recruitment data is estimated using VPA or SPA. The Stock-recruitment data is in tonnes and numbers of recruits, which can be expressed as well in thousands of tonnes and millions of recruits. Recruitment was rounded to the nearest integer to represent counts.

The negative of the Poisson log-likelihood was minimized again using this scaled-down data. The following estimates were obtained for α and β. Similar results were obtained for each pair of starting values. The estimates obtained (Table 3.11) were not sensitive to the starting values of α and β.

Table 3.11 Poisson ML Estimates for scaled-down Baltic data (Ricker model)

Starting Values (α, β)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)
(0, 0)	0.9066	0.005288
(1, 0.1)	0.9062	0.005278
(0.1, 1)	0.9066	0.005288
(10, 0.00001)	0.9066	0.005288

3.3.5 Results for the Poisson Distribution

In BUGS, Poisson models were fitted, using data at the ecosystem scale of millions of recruits and thousands of tonnes of spawning stock biomass. Several different prior distributions for α and β were investigated to see how sensitive the estimates were to the choice of prior distribution. A burn-in period of 5,000 iterations with an additional 10,000 iterations was needed for convergence in all cases. The following combinations of prior distributions were investigated.

i) $\alpha \sim N(0,10^6)$, $\beta \sim N(0,10^{-10})$

ii) $\alpha \sim N(0,10000)$, $\beta \sim N(0,10^{-6})$

iii) $\alpha \sim N(0,1),\ \beta \sim N(0,1)$

iv) $\alpha \sim \exp(0.1),\ \beta \sim Gamma(0.1,0.1)$

v) $\alpha \sim Gamma(1,1),\ \beta \sim \exp(1)$

(The form of the exponential distribution that was used was: $f(x,\lambda)=\lambda e^{-\lambda x}$.) (3.4)

The following results were obtained (Table 3.12). In all cases, the estimates of α and β changed very little, indicating that these estimates were not sensitive to the choice of prior distribution. These estimates were not sensitive to the initial starting values of α and β. Figure 3.11 shows the kernel density plots and the autocorrelation plots for α and β for the first case in Table 3.12. Plots with similar shapes as in Figure 3.11 were obtained for the other cases.

Table 3.12 Poisson estimates for Ricker S-R model

Prior Distributions	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	DIC
$\alpha \sim N(0,10^6)$ $\beta \sim N(0,10^{-10})$	0.8178 (0.08201)	0.002876 (0.002121)	561.473
$\alpha \sim N(0,10000)$ $\beta \sim N(0,10^{-6})$	0.9112 (0.1066)	0.005427 (0.002841)	561.702
$\alpha \sim N(0,1)$ $\beta \sim N(0,1)$	0.9011 (0.1076)	0.00516 (0.002862)	561.709
$\alpha \sim \exp(0.1)$ $\beta \sim Gamma(0.1,0.1)$	0.792 (0.1002)	0.00218 (0.002652)	562.587
$\alpha \sim Gamma(1,1)$ $\beta \sim \exp(1)$	0.9018 (0.09797)	0.005194 (0.002593)	561.354

Estimates obtained from R using the GLM function with an offset were very similar to those obtained by minimizing the negative Poisson log likelihood using the NLM function. However, the sign of beta is reversed.

Table 3.13 Poisson GLM Estimates obtained using an offset

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. error)	$\hat{\beta}$ (std. error)
R	S	0.9213 (0.1069)	-0.005705 (0.002845)

3.4 Beverton-Holt Model

3.4.1 NLS Estimates for the Beverton-Holt Model

The Beverton-Holt model (2.6) was fitted to the Baltic Areas 22-24 data set using non-linear least squares. Numerous combinations of starting values for α and β were tried and the following estimates were obtained (Table 3.14). These estimates were not sensitive to the choice of starting values. However, for the normal case, the model would not run unless β was 0 or very close to 0 ($-0.00001 \leq \beta \leq 0.001$).

Table 3.14 Beverton-Holt NLS estimates

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$	$\hat{\sigma}$
R	S	2.073 (1.231)	7.164 e-7 (1.544 e-5)	0.001600	24.9966
log R	S	1.33240 (0.5684)	0.02542 (0.05197)	36.6300	0.1652

Next, the NLS estimates were re-created in BUGS. Through trial and error, the following prior distributions were selected.

i) S & R: $\alpha \sim N(0,10000)$

$\beta \sim N(0,10^{-6})$

$\tau \sim Gamma(10^{7},1)$

ii) S & log R: $\alpha \sim N(0,10000)$

$\beta \sim N(0,10^{-6})$

$\tau \sim Gamma(200,1)$

A burn-in period of 30,000 iterations with an additional 10,000 iterations was needed for the lognormal case to converge and a burn-in period of 150,000 iterations with an additional 10,000 iterations was needed for the normal case to converge. Severe autocorrelation was present in both cases and was not corrected by running the chains for longer burn-in periods. The following estimates were obtained in BUGS

(Table 3.15). Kernel density plots of the posterior distributions and autocorrelation plots are given in Figures 3.12 and 3.13 respectively.

Table 3.15 Estimates of Beverton-Holt model using the prior distributions: $\alpha \sim N(0,10000)$, $\beta \sim N(0,10^{-6})$ and $\tau \sim Gamma(10^{7},1)$ for the Normal case and $\alpha \sim N(0,10000)$, $\beta \sim N(0,10^{-6})$ and $\tau \sim Gamma(200,1)$ for the lognormal case

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
R	S	2.169 (0.01037)	1.947 e-6 (1.33 e-7)	7.674 e-4 (2.479 e-7)	36.1 (0.00583)
log R	S	0.8903 (0.06827)	0.08156 (0.006251)	31.11 (2.115)	0.1796 (0.006116)

The DIC value for each model were:

1) S & R: 20000300.00

2) S & log R: 325.530

The prior distribution for α was changed in order to investigate the sensitivity and hence required precision in specifying α. Details are given for the lognormal case (Table 3.16). Similar results were obtained for the normal case. The prior distributions for β and τ remained unchanged from the original models (Table 3.15). Each model was rerun using different variances for α, as given below. In all cases, a burn-in period of 30,000 iterations, with an additional 10,000 iterations was needed for convergence.

i) $\alpha \sim N(0,0.1)$

ii) $\alpha \sim N(0,1)$

iii) $\alpha \sim N(0,100)$

iv) $\alpha \sim N(0,10^{10})$

Table 3.16 Effects of changing the prior distribution of α

Prior distribution For α	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
$\alpha \sim N(0,0.1)$	0.2497 (0.02377)	0.02285 (0.002177)	31.28 (2.127)	0.1796 (0.0061)
$\alpha \sim N(0,1)$	0.7939 (0.07261)	0.07273 (0.006651)	31.12 (2.116)	0.1796 (0.006114)
$\alpha \sim N(0,100)$	0.8564 (0.07898)	0.07846 (0.007233)	31.12 (2.115)	0.1796 (0.006115)
$\alpha \sim N(0,10^{10})$	0.8905 (0.06827)	0.08158 (0.006251)	31.11 (2.115)	0.1796 (0.006116)

The prior distribution for β was changed in order to investigate the sensitivity and hence the required precision in specifying β. Again, details are given for the lognormal case. Similar results were obtained for the Normal case. The prior distributions for α and τ remained unchanged from the original model (Table 3.15). Each model was rerun using different prior distributions for β as given below. In both cases, a burn-in period of 30,000 iterations, followed by an additional 10,000 iterations was needed for convergence.

i) $\beta \sim N(0, 0.01)$

ii) $\beta \sim N(0, 10)$

Table 3.17 Effects of changing the prior distribution of β

Prior distribution For β	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
$\beta \sim N(0, 0.01)$	9.012 (0.7197)	0.826 (0.06593)	31.05 (2.11)	0.1798 (0.006119)
$\beta \sim N(0, 10)$	16.03 (1.618)	1.469 (0.1482)	31.05 (2.11)	0.1798 (0.006119)

The prior distribution for τ was changed in order to investigate the sensitivity and hence the precision required in specifying τ. Details are given for the lognormal case. Similar results were obtained for the normal case. The prior distributions for α and β remained unchanged from the original model (Table 3.15). Each model was rerun using the prior distributions for τ as given below (Table 3.18). A burn-in period of 30,000 iterations, followed by an additional 10,000 iterations was needed for convergence.

Table 3.18 Effects of changing the prior distribution of τ

Prior distribution For τ	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
$\tau \sim Gamma(1,1)$	0.3544 (0.1483)	0.03245 (0.01358)	1.706 (0.5034)	0.7919 (0.1225)

3.4.2 ML Estimates for the Beverton-Holt Model

In order to calculate estimates of the Beverton-Holt model using the Poisson distribution, we need to maximize the Poisson log-likelihood. The equation for the Poisson log-likelihood for the Beverton-Holt model is

$$l(\alpha,\beta)=\sum_i\left[-\left(\frac{\alpha S_i}{1+\beta S_i}\right)+R_i\log(\alpha S_i)-R_i\log(1+\beta S_i)\right]. \qquad (3.4)$$

The following estimates were obtained for α and β. This NLM function does not run when the initial value of α is equal to 0.

Table 3.19 Poisson ML Estimates for scaled-down Baltic data

Starting Values (α,β)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)
(0.1, 1)	2.5985	0.007532
(1, 0)	2.5985	0.007532
(100, 1)	2.5999	0.007552
(1, 0.001)	2.5985	0.007532
(0.01, 10)	920.6031	13.4628

These estimates were not sensitive to the starting value of β. However, when the starting value of α was smaller than 0.1, the estimates were very sensitive to the starting value.

3.4.3 Estimates of the Poisson Distribution

The NLM estimates were replicated in BUGS. A burn-in period of 5,000 iterations with an additional 10,000 was required. Several different prior distributions for α and β were investigated to see the effect on the estimates. The following combinations of prior distributions were investigated.

i) $\alpha \sim N(0,10^8)$, $\beta \sim N(0,10^{-10})$

ii) $\alpha \sim N(0,1)$, $\beta \sim N(0,10^{-8})$

iii) $\alpha \sim N(0,1)$, $\beta \sim N(0,1)$

iv) $\alpha \sim \exp(1)$, $\beta \sim Gamma(1,1)$

v) $\alpha \sim Gamma(1,1)$, $\beta \sim \exp(1)$

The following results were obtained (Table 3.20). In all cases, the estimates of α and β changed very little, indicating that these estimates were not sensitive to the choice of prior distribution. Also, these estimates were not sensitive to the initial starting values of α and β. Figure 3.14 shows the kernel density plots of the posterior distributions and the autocorrelation plots for α and β for the N(0, 1) case. Similar shaped plots were obtained for the other cases.

Table 3.20 Poisson estimates for Beverton-Holt S-R model

Prior Distributions	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	DIC
$\alpha \sim N(0,10^{8})$ $\beta \sim N(0,10^{-10})$	2.283 (0.1828)	0.003247 (0.002326)	561.203
$\alpha \sim N(0,1)$ $\beta \sim N(0,10^{-8})$	2.219 (0.1629)	0.002489 (0.002072)	561.554
$\alpha \sim N(0,1)$ $\beta \sim N(0,1)$	2.437 (0.02893)	0.005468 (0.003879)	561.204
$\alpha \sim \exp(1)$ $\beta \sim Gamma(1,1)$	2.646 (0.3362)	0.008259 (0.004536)	560.830

Figure 3.3 Exploratory plots of the Baltic data

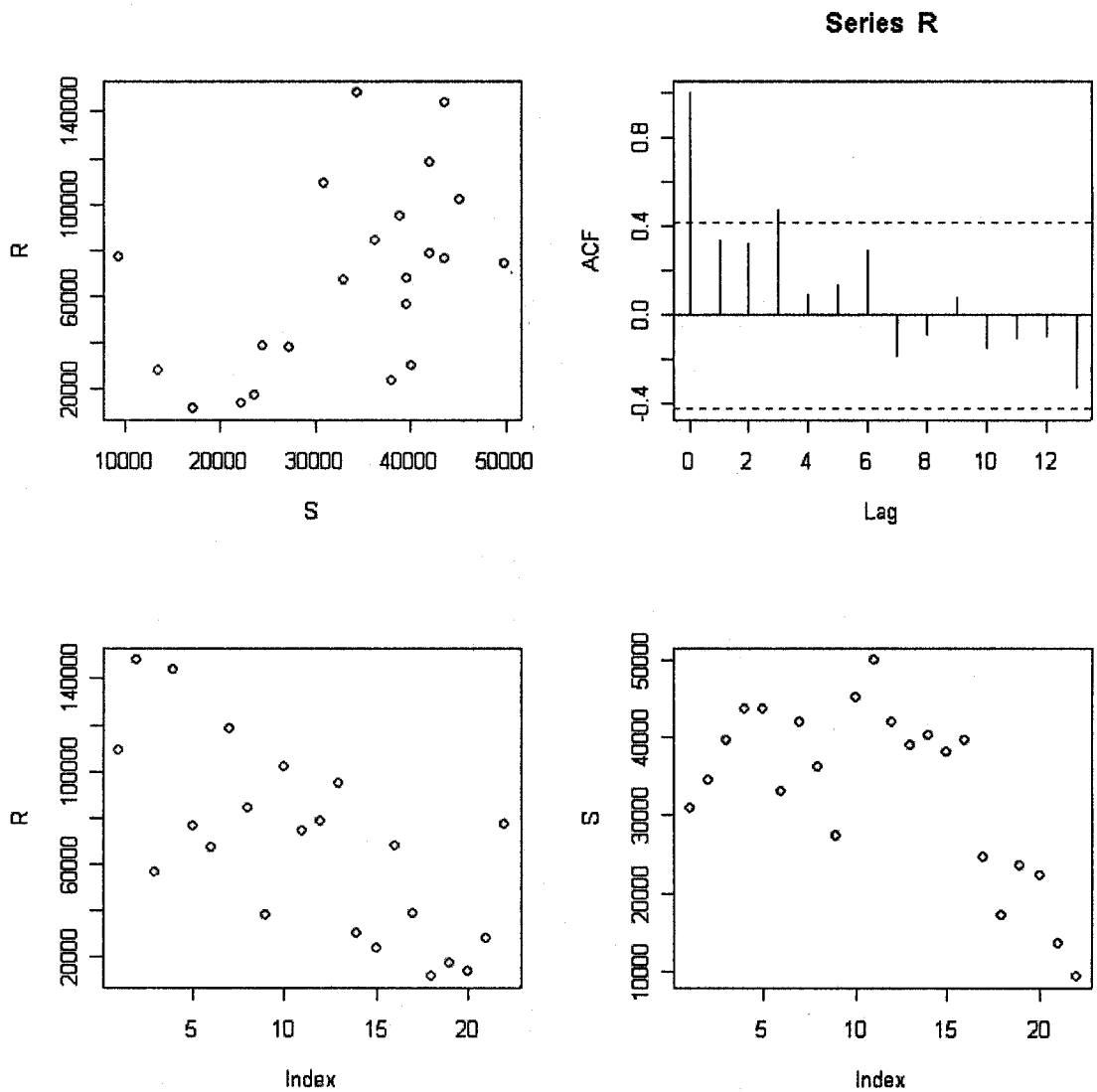


Figure 3.4 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the initial results in Table 3.1 (R &S)

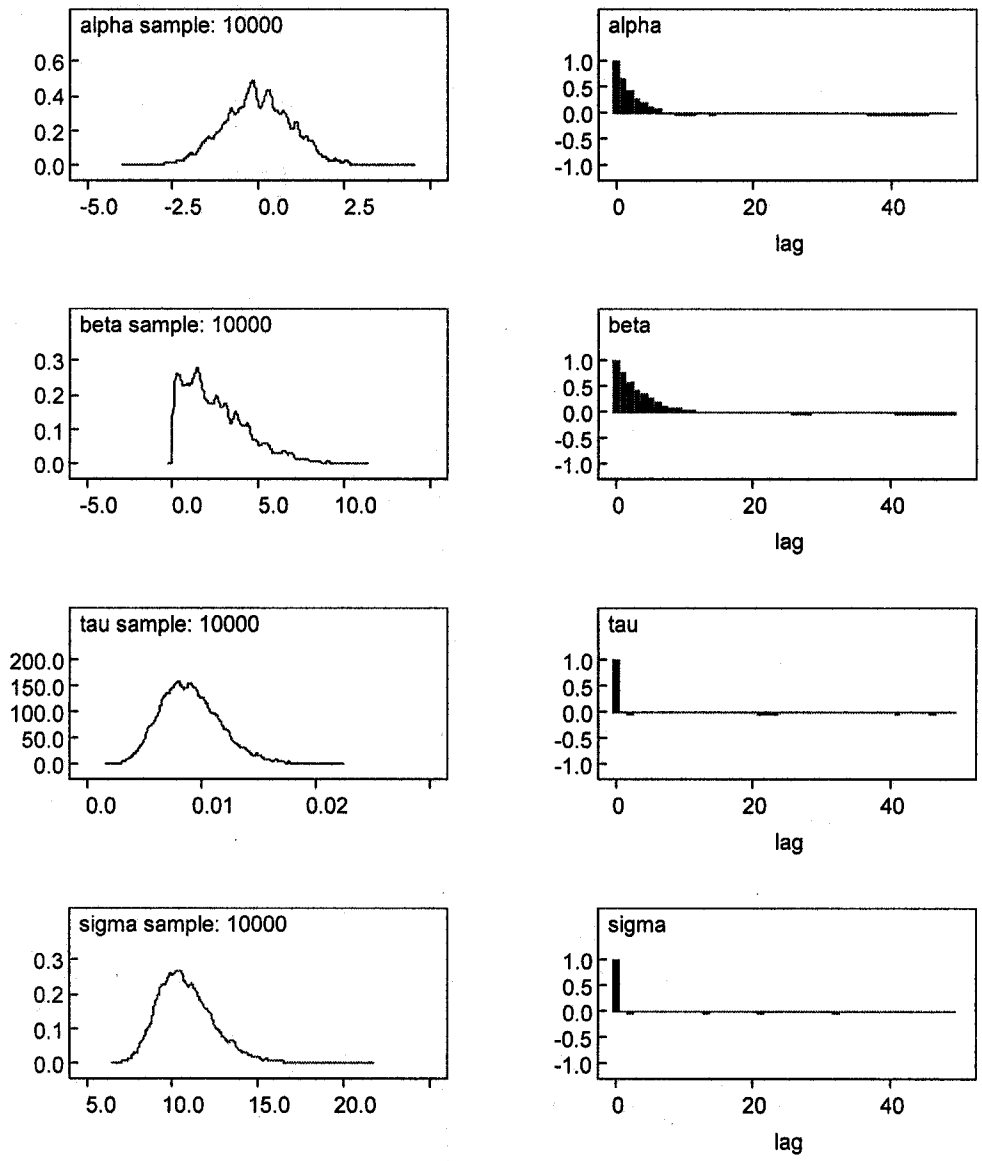


Figure 3.5 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the initial results in Table 3.1 (R & log S)

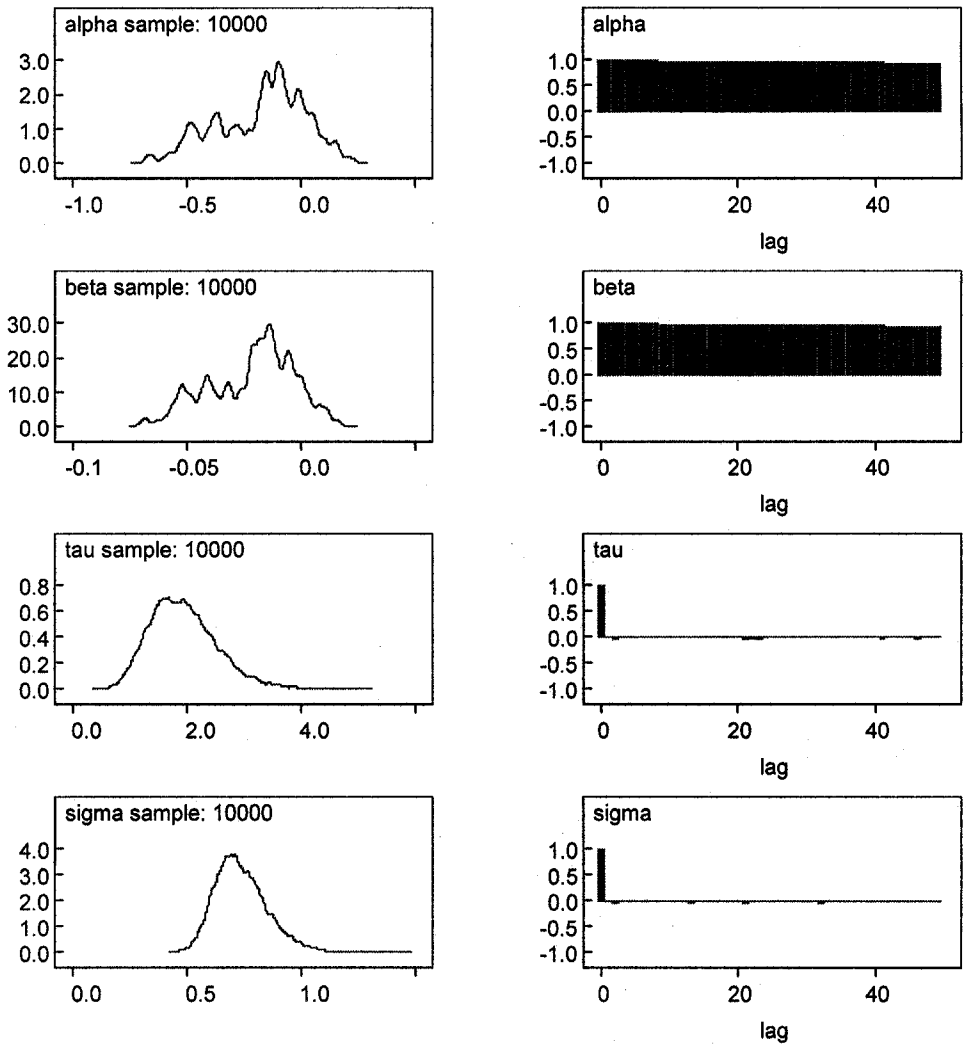


Figure 3.6 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the initial results in Table 3.1 (log R & S)

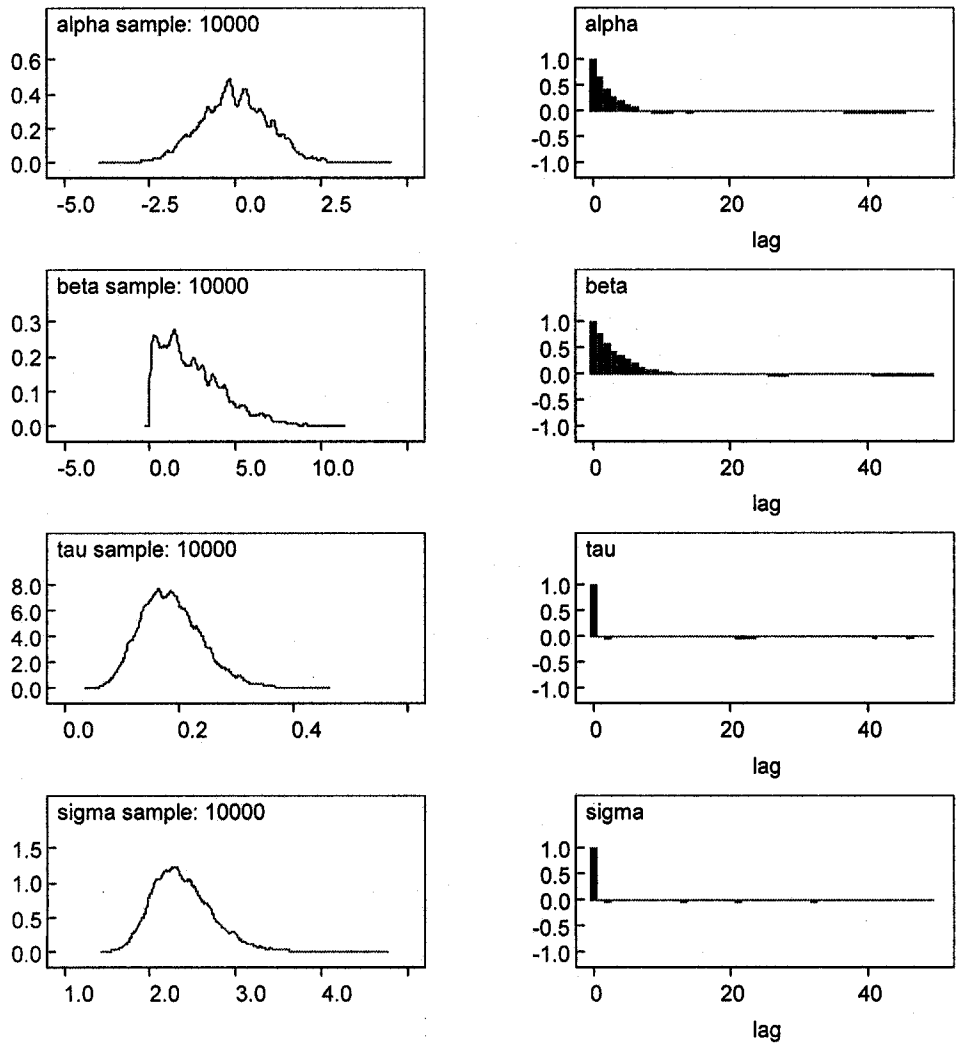


Figure 3.7 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the initial results in Table 3.1 (log R & log S)

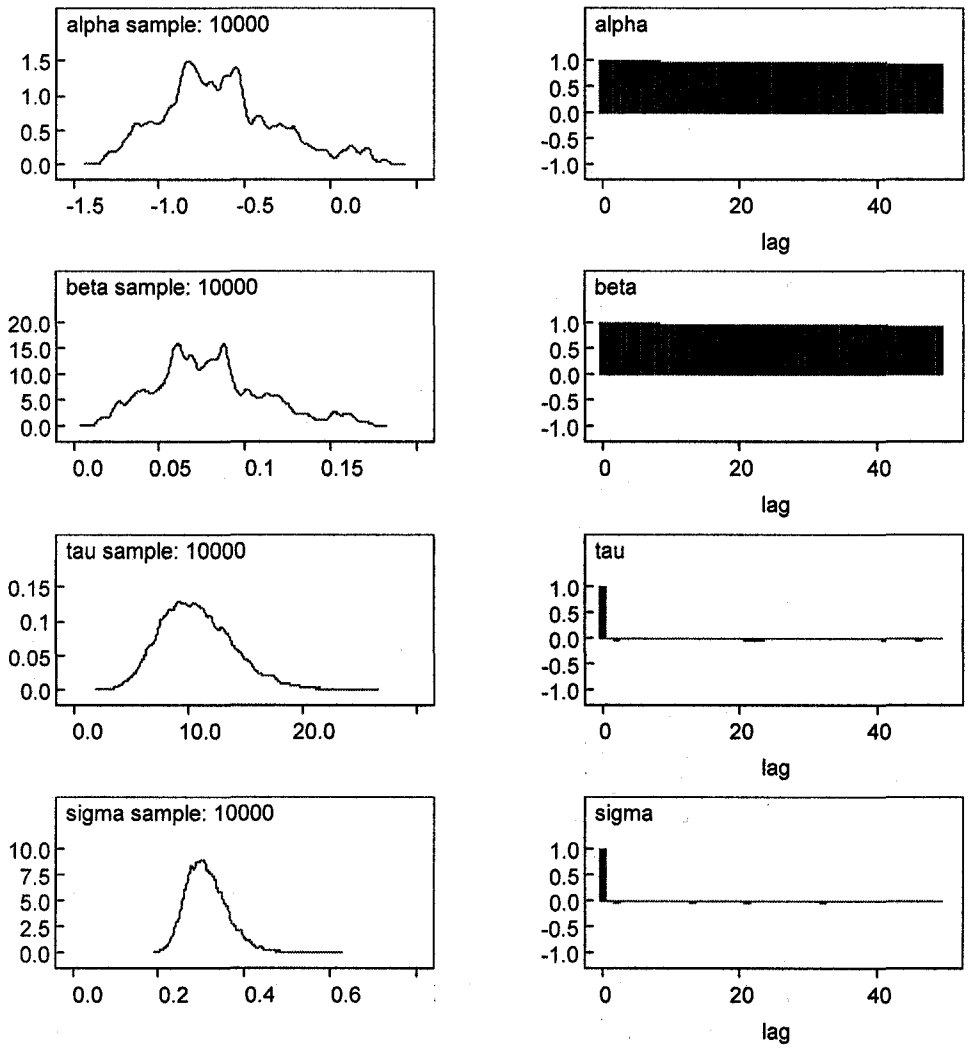


Figure 3.8 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the results in Table 3.5 (R &S)

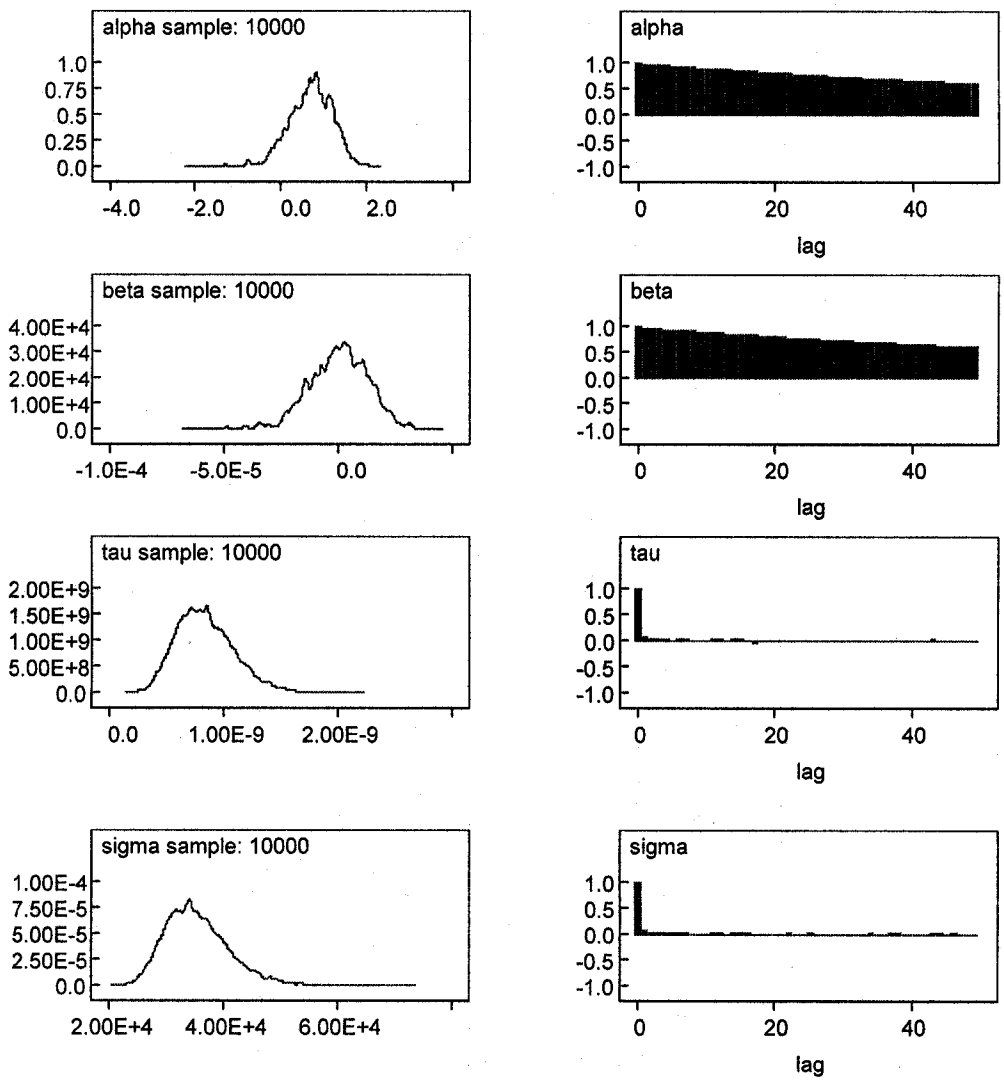


Figure 3.9 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the results in Table 3.5 (log R & S)

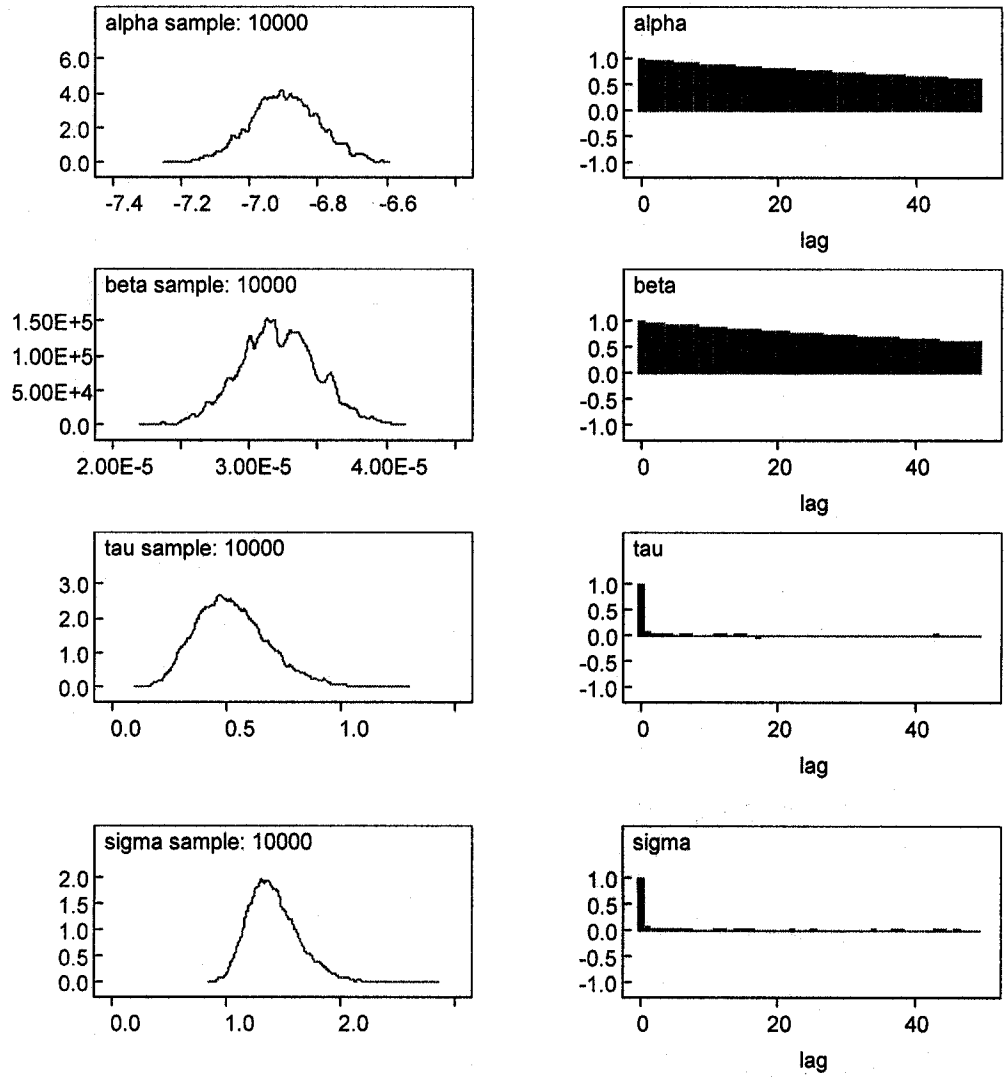


Figure 3.10 Gamma density function plots

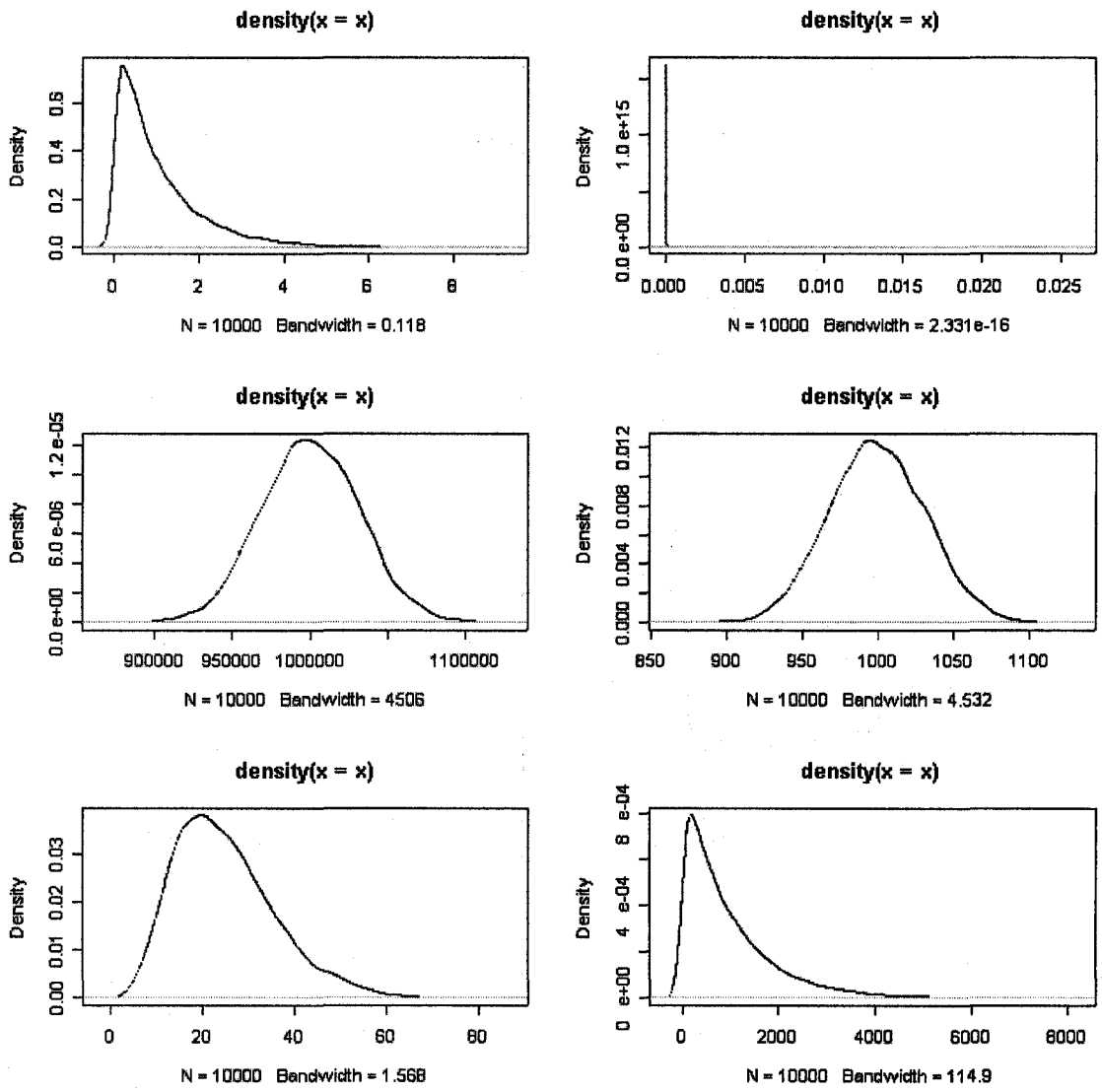


Figure 3.11 Kernel density and autocorrelation plots of the posterior distributions of α and β for the results in Table 3.12 (R is Poisson)

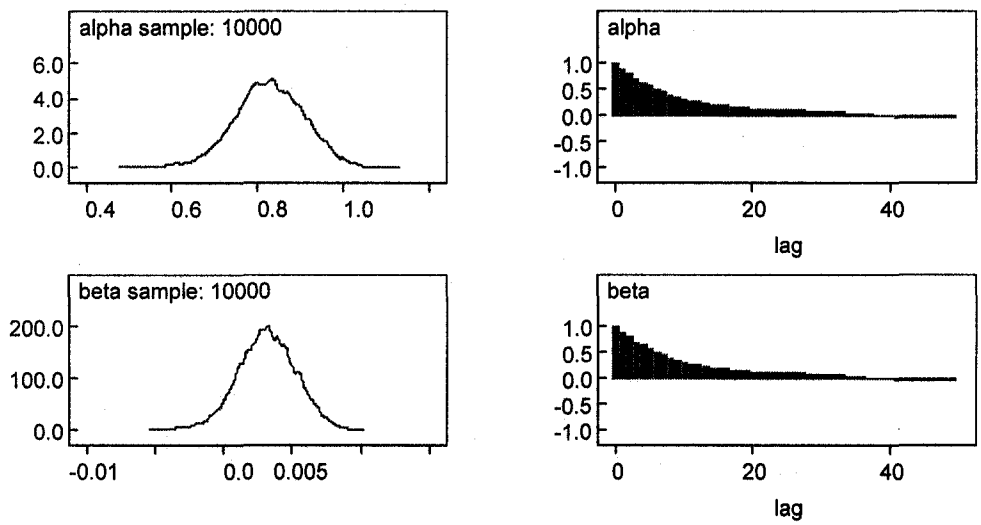


Figure 3.12 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the results in Table 3.14 (R &S)

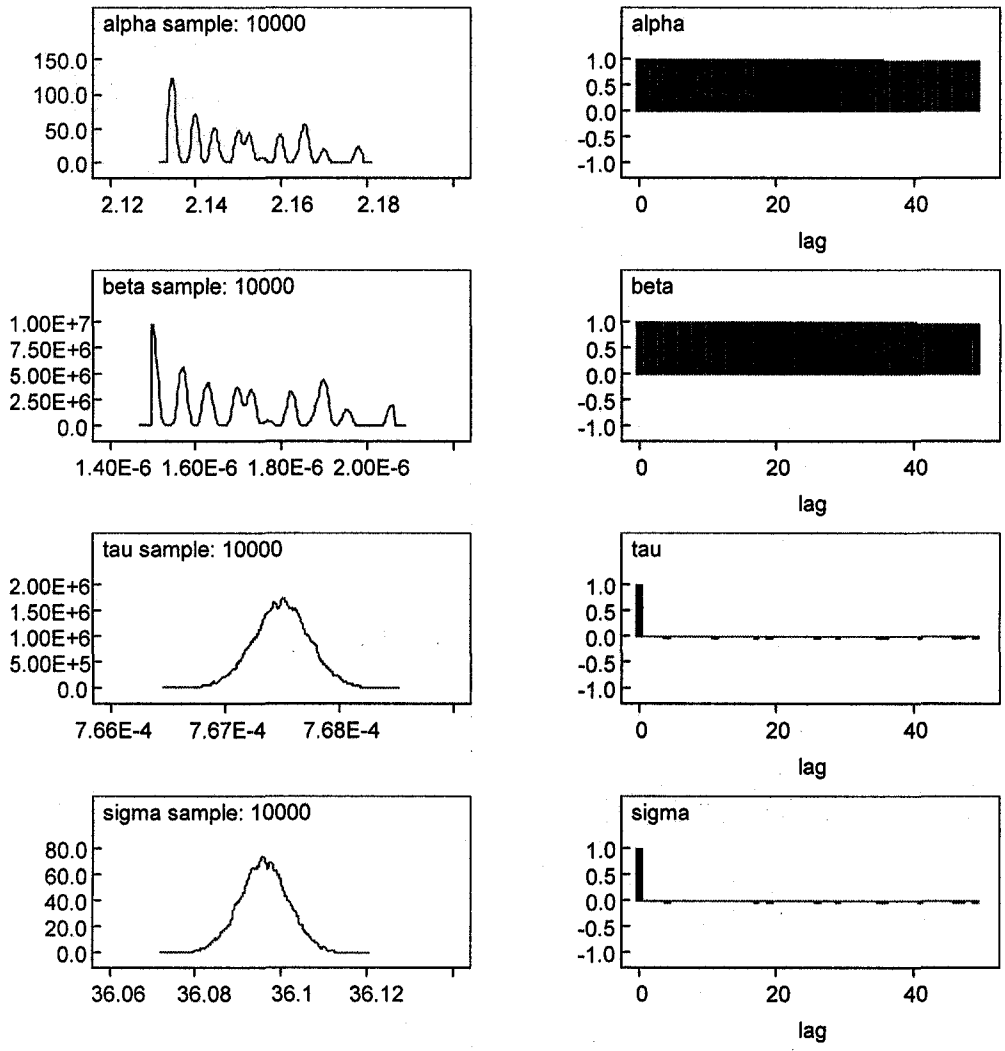


Figure 3.13 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the results in Table 3.14 (log R &S)

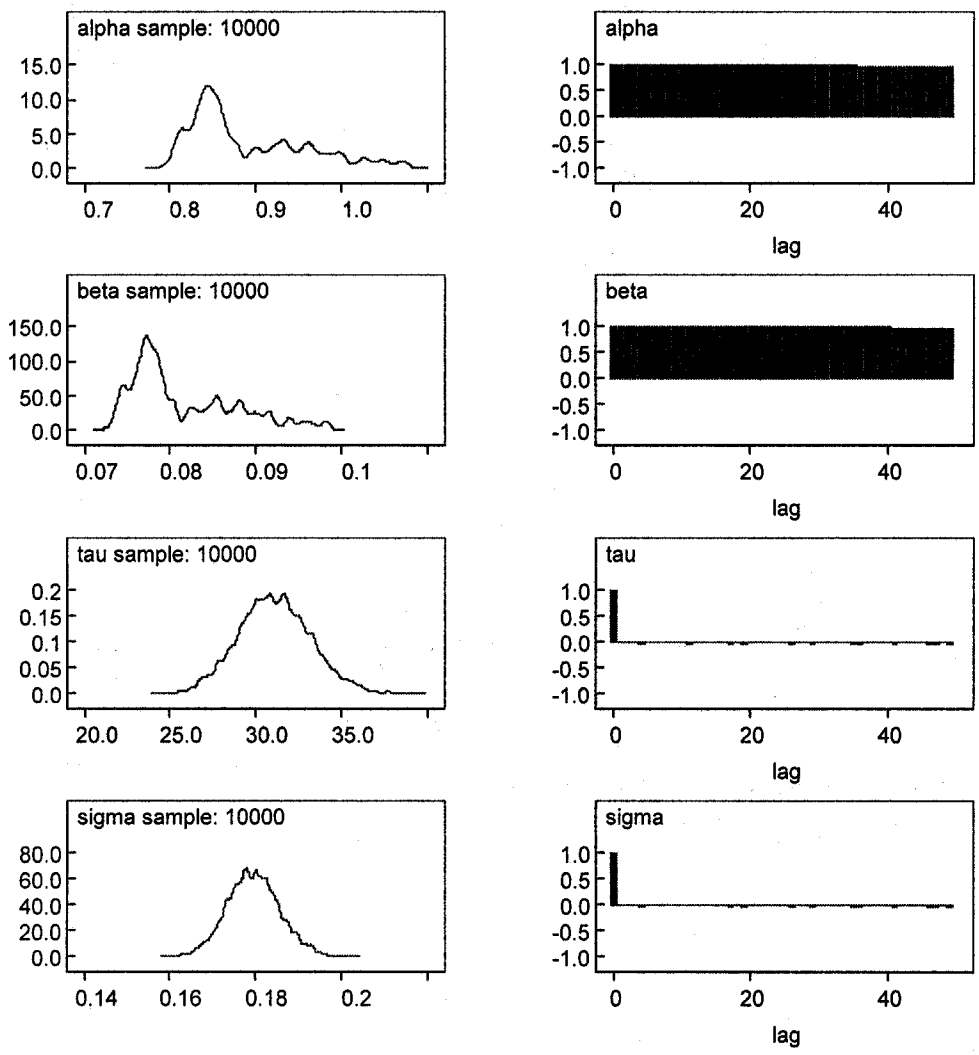
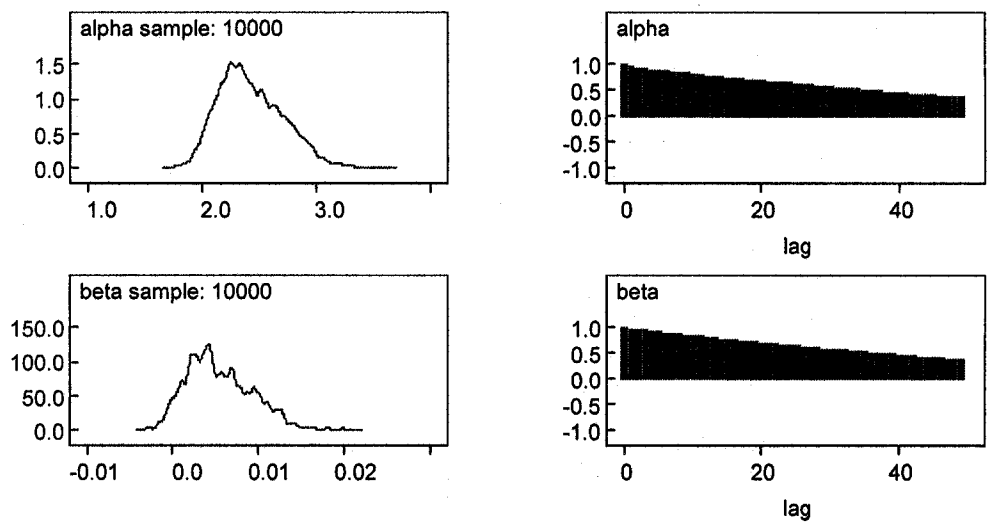


Figure 3.14 Kernel density and autocorrelation plots of the posterior distributions of α and β for the results in Table 3.19 (R is Poisson)



Chapter 4

NAFO subdivision 3Ps Data

4.1 Background

The NAFO subdivision 3Ps cod stock is located off southern Newfoundland from Cape St. Mary's to just west of Burgeo Bank, and over the St. Pierre Bank and most of Green Bank. Cod from this stock generally grow faster than those from other areas further northward but slower than in the Baltic. At least 50% of the females are mature by age 5 (~53 cm) in recent cohorts compared to age 6 (~58 cm) among cohorts in the 1970's and early 1980s (DFO 2004).

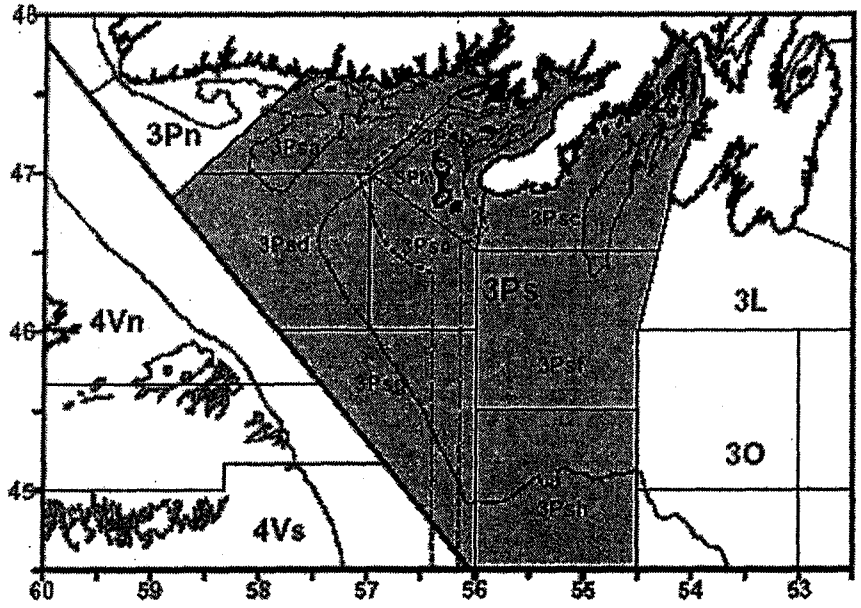


Figure 4.1 Map of NAFO subdivision 3Ps.

4.2 Exploratory Data Analysis

Analysis of the 3Ps data began with a look at some exploratory plots (Figure 4.2). The first plot shows the relationship of recruitment to spawning stock biomass. There appears to be a weak positive relationship between recruitment and spawning stock. The second plot shows the autocorrelation plot for recruitment. The 95% confidence limits about zero are included on the plot. Autocorrelation for recruitment does not seem to be of concern.

The third plot is of recruitment vs. time. This plot shows a decrease in the number of recruits over time. The fourth plot is of spawning stock biomass vs. time. Spawning stock biomass decreases over the first 18 years before increasing again.

4.3 Ricker Model

4.3.1 NLS Estimates for the Ricker Model

The Ricker model (2.8) was fitted to the 3Ps data, using non-linear least squares. Again, the estimates obtained from NLS were used to find non-informative prior distributions for α, β and τ in BUGS.

Four different sets of initial values for α and β were chosen as specified below. All four sets of starting values yielded similar results for both cases (Table 4.1).

1) $\alpha = 2.69,\ \beta = 0.000009937$
2) $\alpha = 0,\ \beta = 0$
3) $\alpha = 5,\ \beta = 0.0001$
4) $\alpha = 1.8988,\ \beta = 0.00003057$

Table 4.1 NLS estimates of the Ricker model

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. error)	$\hat{\beta}$ (std. error)	$\hat{\tau}$	$\hat{\sigma}$
R	S	0.3490 (0.2272)	9.014 e-6 (2.496 e-6)	0.001600	24.9966
log R	S	-7.882 (0.04810)	1.225 e-5 (5.513 e-7)	36.6300	0.1652

4.3.2 Results for the Normal and Lognormal Distributions

In order to replicate the NLS estimates, non-informative prior distributions need to be assigned to α, β and τ. Through trial and error, the following prior distributions were selected.

i) S & R: $\alpha \sim N(0,10000)$

$\beta \sim N(0,10^{-10})$

$\tau \sim Gamma(7000000,1)$

ii) S & log R: $\alpha \sim N(0,10000)$

$\beta \sim N(0,10^{-6})$

$\tau \sim Gamma(400,1)$

The following estimates were obtained (Table 4.2). A 30,000 iteration burn-in period with an additional 10,000 iterations was needed for convergence. Severe autocorrelation was present for both cases. Running each chain for a longer burn-in period did not correct this problem. Both cases yielded estimates that were similar to those obtained from using non-linear least squares (Table 4.1). Kernel density plots of the posterior distributions and autocorrelation plots are given in Figures 4.3 and 4.4.

Table 4.2 Estimates of Ricker S-R model using the prior distributions: $\alpha \sim N(0,10000)$, $\beta \sim N(0,10^{-10})$ and $\tau \sim Gamma(7000000,1)$ for the normal case and $\alpha \sim N(0,10000)$, $\beta \sim N(0,10^{-6})$ and $\tau \sim Gamma(400,1)$ for the lognormal case

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
R	S	0.349 (2.761 e-4)	9.014 e-6 (3.01 e-9)	0.00147 (5.501 e-7)	26.09 (0.004882)
log R	S	-7.863 (0.005878)	1.246 e-5 (6.18 e-8)	41.0 (2.008)	0.1563 (0.003834)

The DIC value for each of the above models was:

1) S & R: 14000400.000

2) S & log R: 692.720

The kernel density plots of α and β (Figure 4.4), for the S & log R model, were quite different than the kernel density plots obtained from previous models. These plots were not as smooth and appear to be multimodal. If the chain was run for a longer period, the plots become smoother.

4.3.3 ML Estimates for the Ricker Model

In order to calculate estimates of the Ricker model using the Poisson distribution, the Poisson log-likelihood needs to be maximized as before. Once again, we minimize the negative of the Poisson log-likelihood in R using the non-linear minimization function (NLM).

The following estimates were obtained for α and β (Table 4.3). These estimates were not sensitive to the choice of starting values for α and β.

Table 4.3 Poisson ML estimates

Starting Values (α, β)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)
(0, 0)	0.3414	0.008935
(1, 0.1)	0.3415	0.008933
(0.1, 1)	0.3414	0.008932
(10, 0.00001)	0.3416	0.008934

4.3.4 Results for the Poisson Distribution

In BUGS, the following prior distributions were investigated to see how sensitive the estimates for α and β were to the choice of prior distribution. Each model was run for a 5,000 iteration burn-in period with an additional 10,000 iterations.

i) $\alpha \sim N(0,10^6)$, $\beta \sim N(0,10^{-10})$

ii) $\alpha \sim N(0,10000)$, $\beta \sim N(0,10^{-6})$

iii) $\alpha \sim N(0,1)$, $\beta \sim N(0,1)$

iv) $\alpha \sim \exp(0.1)$, $\beta \sim Gamma(0.1,0.1)$

v) $\alpha \sim Gamma(1,1)$, $\beta \sim \exp(1)$

The following estimates were obtained (Table 4.4). In all cases, the estimates obtained for α and β change very little. This indicated that these estimates were not sensitive to the choice of the prior distribution. Also, these estimates were not sensitive to the choice of initial starting values of α and β. Figure 4.5 gives the kernel density plots and autocorrelation plots of α and β for the first case in Table 4.4. Similar plots were obtained for the other cases.

Table 4.4 Poisson estimates for the Ricker S-R model

Prior Distributions	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	DIC
$\alpha \sim N(0,10^6)$ $\beta \sim N(0,10^{-10})$	0.2798 (0.08459)	0.008219 (9.397 e-4)	359.045
$\alpha \sim N(0,10000)$ $\beta \sim N(0,10^{-6})$	0.3415 (0.08705)	0.008937 (9.743 e-4)	358.658
$\alpha \sim N(0,1)$ $\beta \sim N(0,1)$	0.3426 (0.08701)	0.008953 (9.717 e-4)	358.618
$\alpha \sim \exp(0.1)$ $\beta \sim Gamma(0.1,0.1)$	0.3316 (0.08845)	0.008826 (9.845 e-4)	358.741
$\alpha \sim Gamma(1,1)$ $\beta \sim \exp(1)$	0.3322 (0.08517)	0.008839 (9.501 e-4)	358.652

4.4 Beverton-Holt Model

4.4.1 NLS Estimates for the Beverton-Holt Model

The Beverton-Holt model (2.6) was fitted to the 3Ps data set using non-linear least squares. Numerous combinations of starting values for α and β were investigated. These estimates were not sensitive to the choice of starting values. However, for the normal case, the model would not run unless β was 0 or close to 0 ($-0.00001 \leq \beta \leq 0.001$). The following estimates were obtained (Table 4.5).

Table 4.5 Beverton-Holt NLS estimates

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
R	S	2.405 (1.676)	3.226 e-5 (3.081 e-5)	0.001592	25.0654
log R	S	2.6160 (1.3907)	0.1531 (0.1285)	82.1763	0.1103

4.4.2 Estimates for the Normal and Lognormal Distributions

The NLS estimates were re-created in BUGS. Through trial and error, the following prior distributions were selected.

i) S & R: $\alpha \sim N(0,10000)$

$\beta \sim N(0,10^{-10})$

$\tau \sim Gamma(10^{7},1)$

ii) S & log R: $\alpha \sim N(0,10000)$

$\beta \sim N(0,10^{-5})$

$\tau \sim Gamma(200,1)$

A burn-in period of 30,000 iterations with an additional 10,000 iterations was required for the lognormal case to converge, while a burn-in period of 150,000 iterations was required for the normal case to converge. Severe autocorrelation was present in both cases (Figures 4.6 and 4.7), which could not be corrected by running the chains for longer burn-in periods. The following estimates were obtained (Table 4.6) for both the normal and lognormal cases.

Table 4.6 Beverton-Holt estimates using the prior distributions: $\alpha \sim N(0,10000)$, $\beta \sim N(0,10^{-10})$ and $\tau \sim Gamma(10^{7},1)$ for the Normal case and $\alpha \sim N(0,10000)$, $\beta \sim N(0,10^{-5})$ and $\tau \sim Gamma(200,1)$

Recruitment (R)	Stock (S)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	$\hat{\tau}$ (std. dev.)	$\hat{\sigma}$ (std. dev.)
R	S	2.7699 (0.003935)	3.948 e-5 (7.28 e-8)	0.002073 (6.582 e-7)	21.96 (0.003487)
log R	S	2.588 (0.1495)	0.2392 (0.01381)	69.21 (4.652)	0.1204 (0.004052)

The DIC value for each model was:

1) S & R: 19544000.00

2) S & log R: 219.332

4.4.3 ML Estimates for the Beverton-Holt Model

In order to calculate estimates of the Beverton-Holt model using the Poisson distribution, the Poisson log-likelihood needs to be maximized. The following estimates were obtained for α and β (Table 4.7).

Table 4.7 Poisson ML estimates for the Beverton-Holt model

Starting Values (α, β)	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)
(1, 1)	2.2171	0.0288
(1, 0)	2.2169	0.0288
(10, 1)	2.2170	0.0288
(1, 0.01)	2.2137	0.0287
(0.001, 1)	92.148	1.7204

The NLM function would not run when the starting value of α was 0. These estimates were not sensitive to the choice of the starting value of β. When the starting value of α was less than 0.1, the estimates were quite sensitive to the choice of starting value.

4.4.4 Estimates of the Poisson Distribution

The ML estimates were replicated in BUGS through trial and error. A burn-in period of 5,000 iterations with an additional 10,000 iterations was required for convergence. Several different prior distributions for α and β were investigated to see if there was an effect on the estimates. The following combinations of prior distributions were investigated.

i) $\alpha \sim N(0,10^{8})$, $\beta \sim N(0,10^{-10})$

ii) $\alpha \sim N(0,1)$, $\beta \sim N(0,10^{-8})$

iii) $\alpha \sim N(0,1)$, $\beta \sim N(0,1)$

iv) $\alpha \sim N(0,100)$, $\beta \sim N(0,1)$

v) $\alpha \sim Gamma(1,1)$, $\beta \sim \exp(1)$

The estimates of α and β (Table 4.8) changed as the prior distributions were changed, indicating that these estimates are sensitive to the choice of prior distribution. Kernel density plots of the posterior distributions and the autocorrelation plots of α and β are given in Figure 4.8 for the N(0, 1) case. Similar shaped plots were obtained for the other cases.

Table 4.8 Poisson estimates for the Beverton-Holt model

Prior Distributions	$\hat{\alpha}$ (std. dev.)	$\hat{\beta}$ (std. dev.)	DIC
$\alpha \sim N(0,10^{8})$ $\beta \sim N(0,10^{-10})$	1.231 (0.1067)	0.01058 (0.001863)	370.687
$\alpha \sim N(0,1)$ $\beta \sim N(0,10^{-8})$	1.69 (0.2101)	0.01906 (0.003832)	362.031
$\alpha \sim N(0,1)$ $\beta \sim N(0,1)$	2.007 (0.3491)	0.02498 (0.006467)	361.020
$\alpha \sim N(0,100)$ $\alpha \sim N(0,1)$	2.379 (0.5993)	0.03182 (0.01111)	361.093
$\alpha \sim Gamma(1,1)$ $\beta \sim \exp(1)$	2.342 (0.5059)	0.03119 (0.00941)	360.508

Figure 4.2 Exploratory plots of the NAFO subdivision 3Ps data

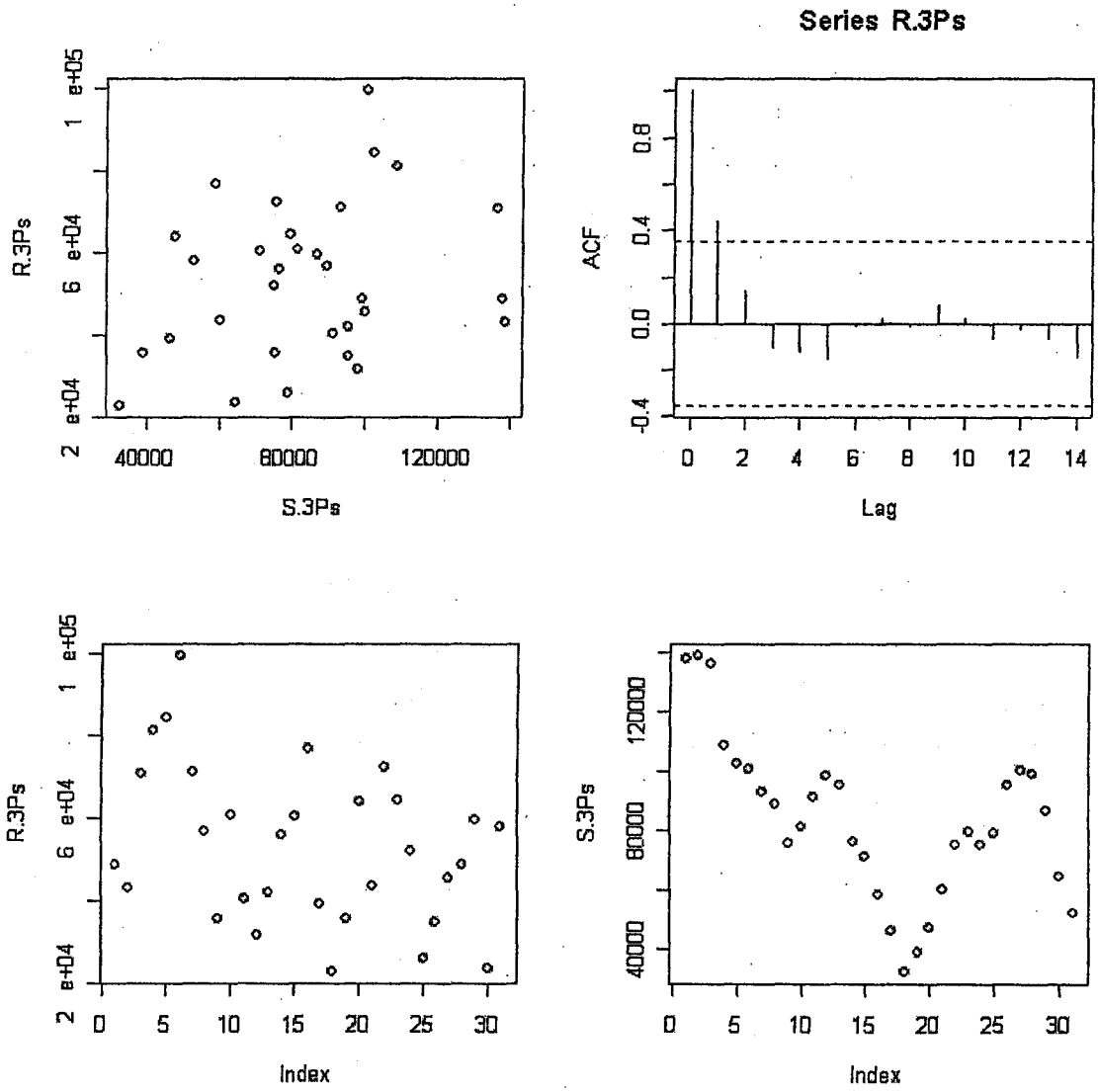


Figure 4.3 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the results in Table 4.2 (R &S)

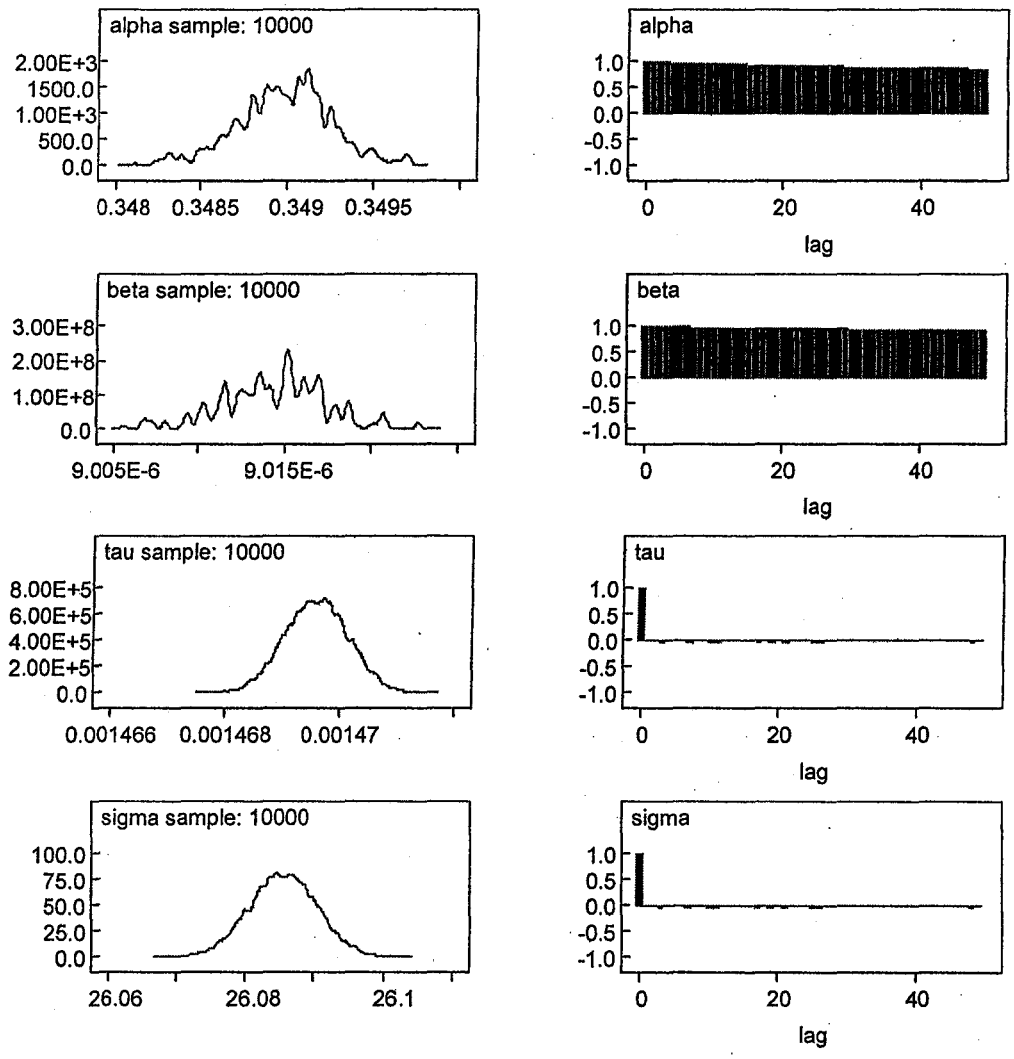


Figure 4.4 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the results in Table 4.2 (log R &S)

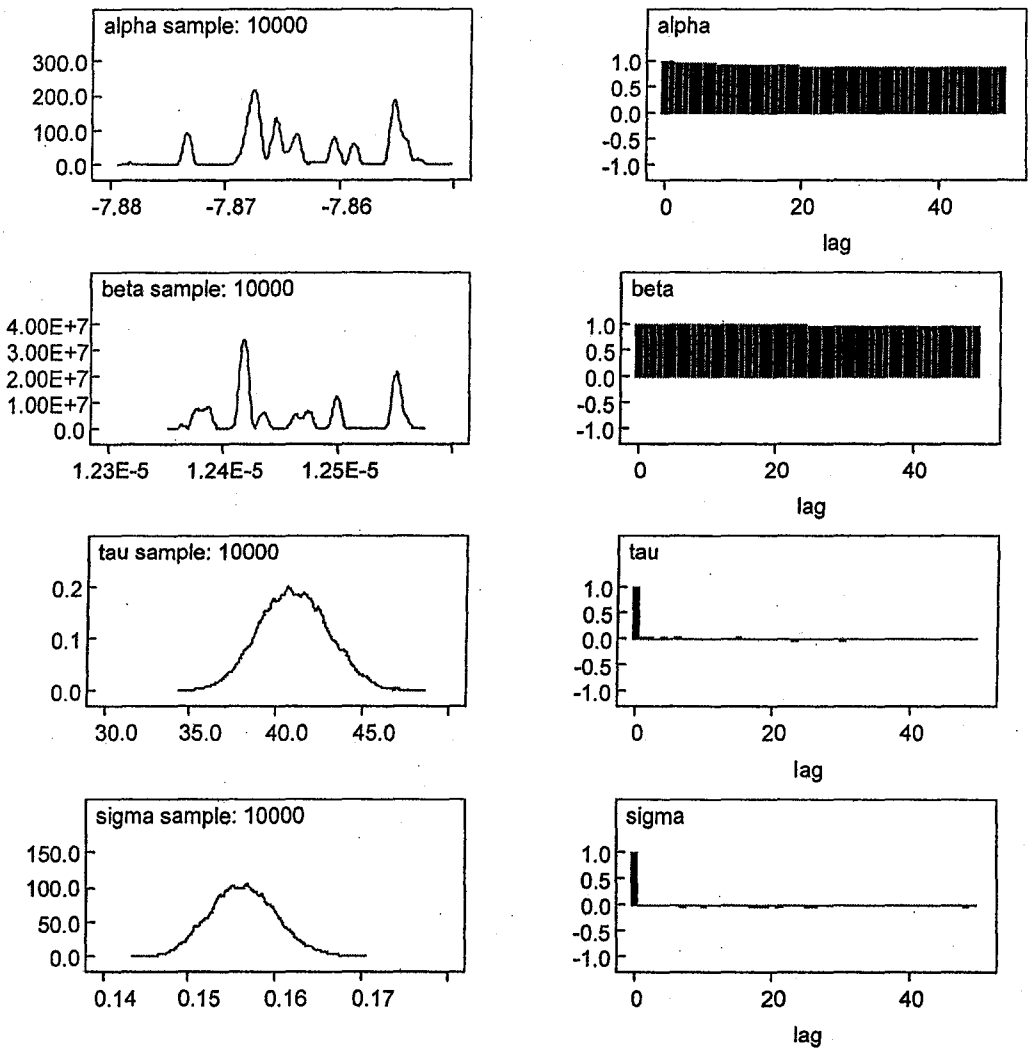


Figure 4.5 Kernel density and autocorrelation plots of the posterior distributions of α and β for the results in Table 4.4 (R is Poisson)

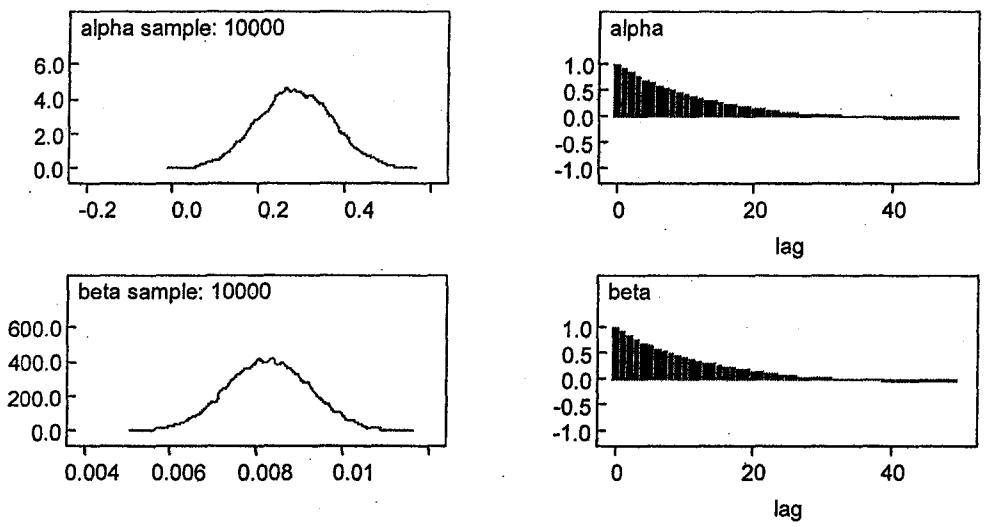


Figure 4.6 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the results in Table 4.6 (R &S)

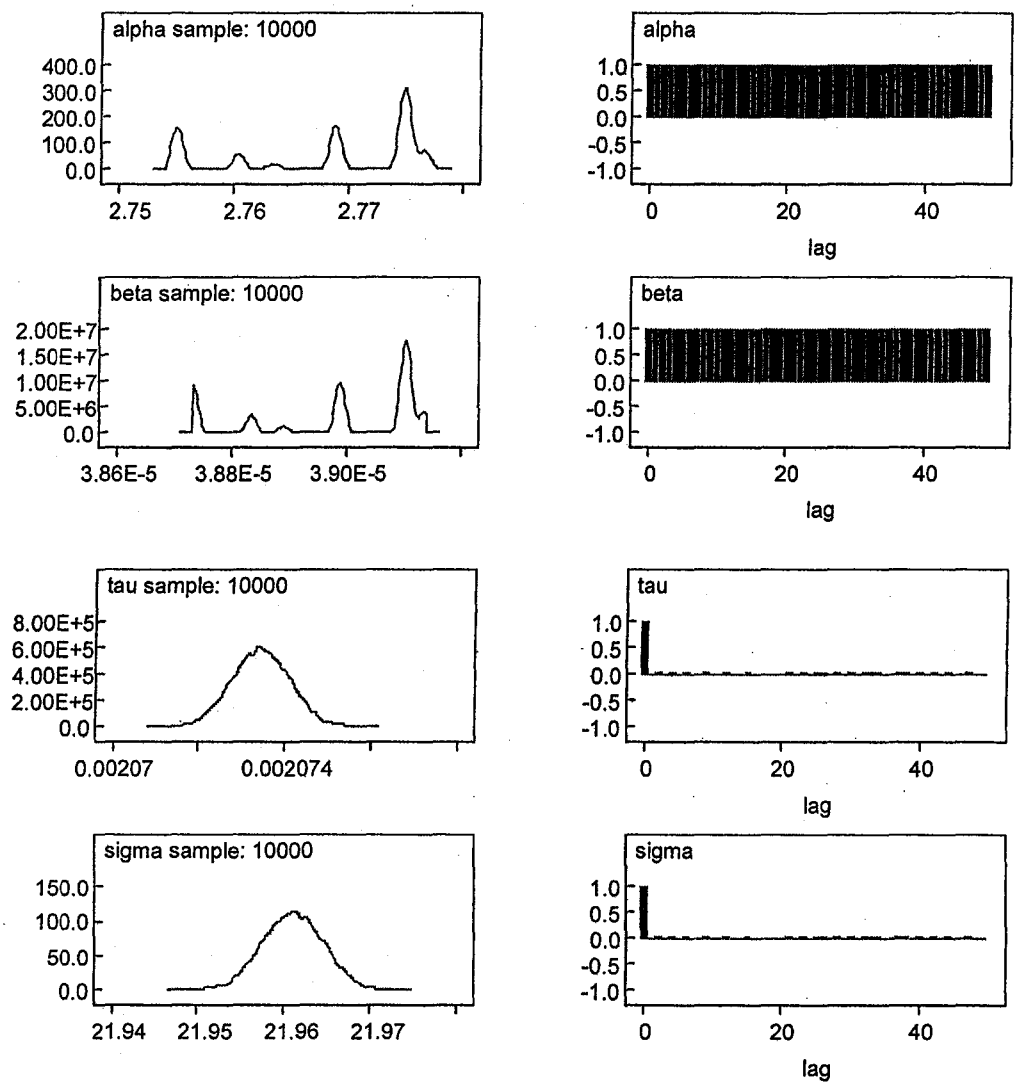


Figure 4.7 Kernel density and autocorrelation plots of the posterior distributions of α, β, τ and σ for the results in Table 4.6 (log R &S)

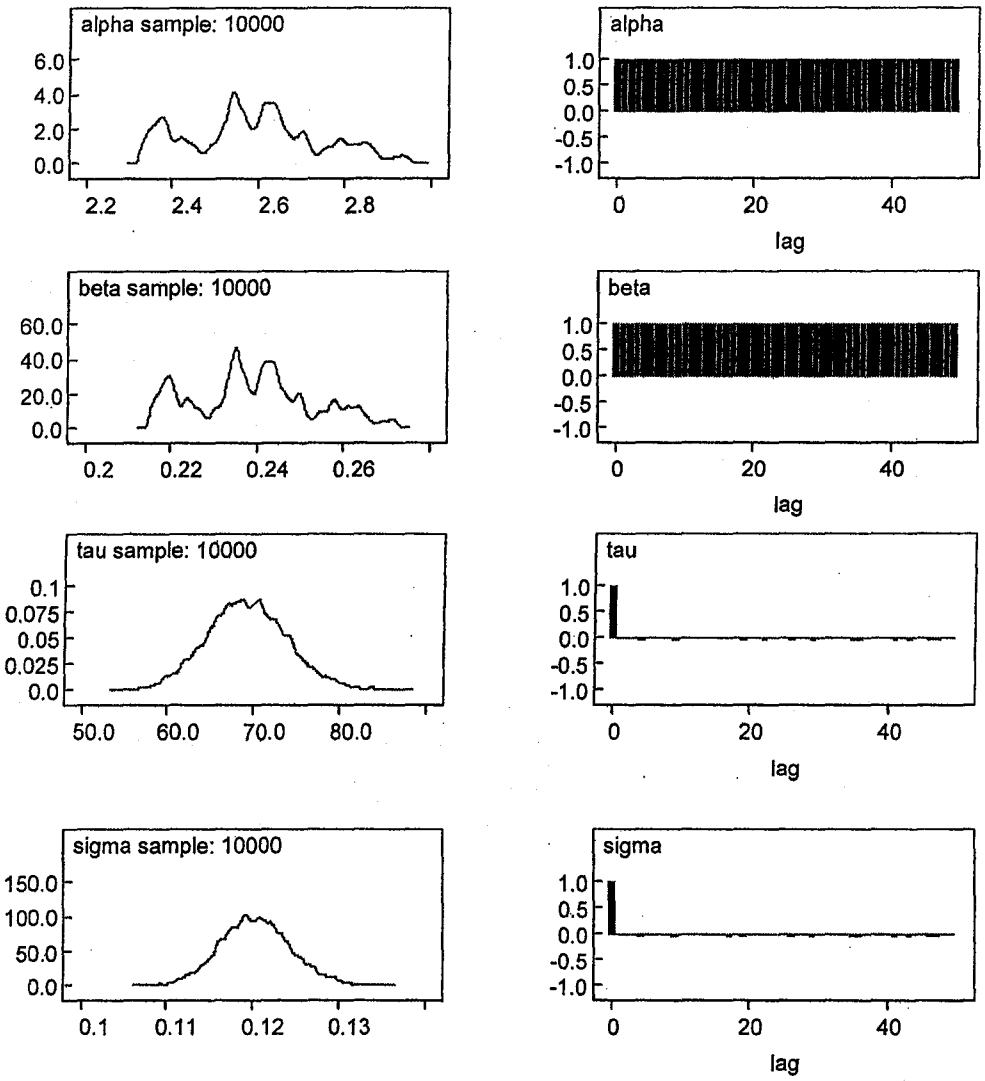
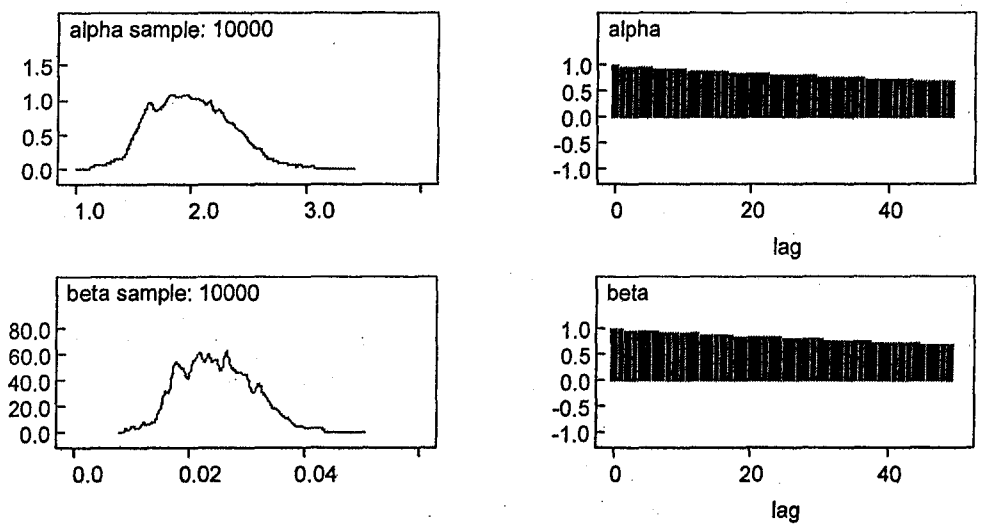


Figure 4.8 Kernel density and autocorrelation plots of the posterior distributions of α and β for the results in Table 4.8 (R is Poisson)



Chapter 5

Discussion

Posterior parameter estimates in fisheries models are quite sensitive to the choice of prior distribution. In particular, posterior estimates can differ dramatically from the non-Bayesian (LS or ML) estimates.

The posterior distributions of α and β, obtained using the Ricker model on the Baltic Areas 22-24 data, were quite different than the specified prior distributions for the normal and lognormal cases. The DIC value for the lognormal case was 76.693. In both cases, the posterior estimates were not sensitive to the choice of initial starting values of the chain. The prior distribution specification for τ had little effect on the posterior results. The results varied greatly depending on the specification of the prior distribution for β. When the variance of the prior distribution of α was greater than 100, the estimates were highly sensitive to the choice of the prior distribution. For the Poisson case, the estimates obtained were not sensitive to the choice of the prior distributions for α and β. The DIC value for the Poisson case was 562.587.

The posterior distributions of α and β, obtained using the Beverton-Holt model on the Baltic Areas 22-24 data, also differed substantially from the specified prior distributions for the normal and lognormal cases. This model required a longer burn-in period in order to converge. The DIC value for the lognormal case was 325.530. The prior distribution specification for α and τ had little effect on the posterior results. The posterior estimates of β varied greatly depending on the specification of the prior distribution for β. For the Poisson case, the estimates obtained were not sensitive to the choice of the prior distributions for α and β. The DIC value for the Poisson case was 561.554.

The Ricker model provided the best fit for the lognormal distribution on the Baltic Area 22-24 data (DIC = 76.693 vs. 325.530 for the Beverton-Holt). For the Poisson distribution, the DIC values for both models were very close (562.587 for the Ricker model vs. 561.554 for the Beverton-Holt model).

Results obtained for the NAFO subdivision 3Ps data, using the Ricker model, were similar to those obtained for the Baltic Areas 22-24 data, using the Ricker model, for the normal and lognormal cases. The results were sensitive to the choice of the prior distribution for β. The DIC for the lognormal case was 692.720. For the Poisson case, the estimates obtained were not sensitive to the choice of the prior distributions for α and β. The DIC value for the Poisson case was 358.658.

Results obtained for the NAFO subdivision 3Ps data, using the Beverton-Holt model, were similar to those obtained for the Baltic Areas 22-24 data, using the Beverton-Holt model, for the normal and lognormal cases. Again, the results were sensitive to the choice of the prior distribution for β and the chain was slower to converge than the Ricker model. The DIC for the lognormal case was 219.332. For the Poisson case, the estimates varied as the prior distributions of α and β changed. The DIC for the Poisson case was 362.031.

The Beverton-Holt model provided the best fit for the lognormal distribution on the NAFO subdivision 3Ps data (DIC = 219.332 vs. 692.720 for the Ricker). For the Poisson distribution, both the Ricker and Beverton-Holt models yielded similar DIC values (358.658 for the Ricker model and 362.031 for the Beverton-Holt model).

Very different results were obtained for both of the data sets, indicating that model parameters are stock specific. Generally, results for the normal and lognormal distributions were very sensitive to the choice of the prior distribution of β. In practice, more emphasis should be placed on properly specifying the prior distribution of β.

For all cases, severe autocorrelation was present in the results which could not be corrected by running the chain for a longer burn-in period. The use of multiple chains

with different starting values also had little effect on this problem. This also implies that many of the posterior estimates of α and β may be suspect. This is worth further investigation. This analysis could also be expanded to include additional sampling distributions, such as the Gamma distribution or the Negative Binomial distribution. Additional cod stocks could also be investigated since the parameter estimates vary depending on the particular stock.

The posterior estimates were sensitive to our choice of prior distribution, which can have an important practical effect on the study of stock-recruitment issues in a number of different fisheries.

Bibliography

1. Adkinson, M.D, and Peterman, R.M. 1996. Results of Bayesian methods depend on details of implementation: an example of estimating salmon escapement goals. Fish. Res. 25:155-170.
2. Berger, J.O. 1985. Statistical decision theory and Bayesian analysis. Springer-Verlag, New York.
3. Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish population. Fish. Invest. Ser. II. Mar. Fish. G.B. Minist. Agric. Fish. Food. Vol 19.
4. Box, G.E.P., and Tiao, G.C. 1992. Bayesian Inference in Statistical Analysis. Wiley Classics Library Ed. John Wiley and Sons, New York.
5. Carlin, B.P., and Louis, T.A. 1998. Bayes and empirical Bayes methods for data analysis. Chapman and Hall/CRC, New York.
6. Casella, G., and George, E.I. 1992. Explaining the Gibbs Sampler. The American Statistician 46, No. 3, 167-174.

7. Chen, D.G., and Holtby, L.B. 2002. A regional meta-model for stock-recruitnemt analysis using an empirical Bayesian approach. Can. J. Fish. Aquat. Sci. 59: 1503-1514.

8. Chen, Y, and Fournier, D. 1999. Impacts of atypical data on Bayesian inference and robust Bayesian approach in fisheries. Can. J. Fish. Aquat. Sci. 56: 1525-1533.

9. Cushing, D.H. 1973. Dependence of recruitment on parent stock. J. Fish. Res. Board Can. 30: 1965-1976.

10. Deriso, R.B. 1980. Harvesting strategies and parameter estimation for an age-structured model. Can. J. Fish. Aquat. Sci. 37: 268-282.

11. DFO, 2004. Subdivision 3Ps Cod. DFO Can. Sci. Advis. Sec. Stock Status Report. 2004/039.

12. Gelfand, A.E, and Smith, A.F.M. 1990. Sampling-based approaches to calculating marginal densities. J. Am. Stat. Assoc. 85: 398-409.

13. Getz, W.M., and Swartzman, G.L. 1981. A probability transition matrix model for yield estimation in fisheries with highly variable recruitment. Can. J. Fish. Aquat. Sci. 38: 847-855.

14. Geweke, J.F. 1989. Bayesian inference in econometric models using Monte Carlo integration. Econometria 57: 1317-1340.

15. Gilks, W.R., Thomas, A., and Spiegelhalter, D.J. 1994. A language and program for complex Bayesian modeling. The Statistican 43, No. 1, 169-177.

16. Harley, S.J., and Myers, R.A. 2001. Hierarchical Bayesian models of length-specific catchability of research trawl surveys. Can. J. Fish. Aquat. Sci. 58: 1596-1584.

17. Hastings, W.K. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57: 97-109.

18. Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York.

19. Jiao, Y., Chen, Y., Schneider, D., and Wroblewski, J. 2004a. A simulation study of impacts of error structure on modeling stock-recruitment data using generalized linear models. Can. J. Fish. Aquat. Sci. 61: 122-133.

20. Jiao, Y., Schneider, D., Chen, Y., and Wroblewski, J. 2004b. An analysis of error structure in modeling the stock–recruitment data of gadoid stocks using generalized linear models. Can. J. Fish. Aquat. Sci. 61: 134-146.

21. Kass, R.E., and Wasserman, L. 1996. The selection of prior distributions by formal rules. J. Am. Stat. Assoc. 91: 1343-1370.

22. Larkin, P.A. 1973. Some observations on models of stock and recruitment relationships for fishes. Conseil Int. Explor. Mer, Rapp et Proc.-Verb. 164: 316-324.

23. Liermann, M., and Hilborn, R. 1997. Depensation in fish stocks: a hierarchic Bayesian meta-analysis. Can. J. Fish. Aquat. Sci. 54: 1976-1984.

24. Ludwig, D., and Walters, C.J. 1989. A robust method for parameter estimation from catch and effort data. Can. J. Fish. Aquat. Sci. 46: 137-144.

25. McAllister, M.K., and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54: 284-300.

26. Millar, R.B. 2002. Reference priors for Bayesian fisheries models. Can. J. Fish. Aquat. Sci. 59: 1492-1502

27. Myer, R., and Millar, R.B. 1999. Bayesian stock assessment using a state-space implementation of the delay difference model. Can. J. Fish. Aquat. Sci. 56: 37-52.

28. Myers, R.A., and Barrowman, N.J. 1996. Is fish recruitment related to spawner abundance? Fish. Bull. 94: 707-724.

29. Myers, R.A., Barrowman, N.J., Hilborn, R., and Kehler, D.G. 2002. Inferring Bayesian Priors with limited direct data: applications to risk analysis. North Amer. J. Fish. Man. 22: 351-364.

30. Myers, R.A., Bridson, J., and Barrowman, N.J. 1995. Summary of worldwide spawner and recruitment data. Can. Tech. Rep. Fish. Aquat. Sci. No. 2024.

31. Myers, R.A., Hutchings, J.A, and Barrowman, N.J. 1997. Why fish stocks collapse? The example of cod in Canada. Ecological Applications 7: 91-106.

32. Punt, A.E., and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Review in Fish Biology and Fisheries 7: 35-63.

33. Quinn, T., and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, New York.

34. Ricker, W.E. 1954. Stock and recruitment. J. Fish. Res. Board Can. 11: 559-623.

35. Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Fish. Res. Board Can. Bull. No. 191.

36. Rivot, E., Prevost, E., and Parent, E. 2001. How robust are Bayesian posterior inferences based on a Ricker model with regards to measurement errors and prior assumption about parameters? Can. J. Fish. Aquat. Sci. 58: 2284-2297.

37. Robb, C.A., and Peterman, R.M. 1998. Applications of Bayesian decision analysis to management of a sockeye salmon (Oncorhynchus nerka) fishery. Can. J. Fish. Aquat. Sci. 55: 86-98.

38. Rothschild, B.J. 1986. Dynamics of Marine Fish Populations. Harvard Univ. Press, Cambridge, Massachusetts.

39. Schnute, J. 1985. A general theory for the analysis of catch and effort data. Can. J. Fish. Aquat. Sci. 42: 414-429.

40. Shelton, P.A. 1992. The shape of recruitment distributions. Can. J. Fish. Aquat. Sci. 49: 1754-1761.

41. Sheperd, J.G. 1982. A versatile new stock recruitment relationship for fisheries and the construction of sustainable yield curves. J. Cons. Int. Explor. Mer. 40: 67-75.

42. Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and van der Linde, A. 2002. Bayesian measures of model complexity and fit. J.R. Statist. Soc. B. 64, Part 4, 583-639.

43. Tanner, M.A., and Wong, W.H. 1987. The calculation of posterior distributions by data augmentation (with discussion). J. Amer. Statist. Assoc. 82: 528-550.

44. Walsh, B. 2004. Lecture notes for EEB 581, Version 26, April 2004.

(http://nitro.bio.sci.arizona.edu/courses/EBB581-2004/handouts/Gibbs.pdf).

Appendix

Sample BUGS Syntax

Below are samples of the BUGS command files that were used to perform the analysis for this paper. Sample code is given for each of the three distributions that were investigated (Normal, Lognormal and Poisson).

1. Normal distribution sample syntax:

```
model {
        for (i in 1:N) {
                mu[i] <- S[i]*exp(alpha-beta*S[i])
                R[i]~dnorm(mu[i],tau)
                        }
        alpha~dnorm(0, 0.01)
        beta~dnorm(0, 1000000)
        tau~dgamma(1, 5)
        sigma <- 1/sqrt(tau)
        }
```

2. Lognormal distribution sample syntax:

```
model {
        for (i in 1:N) {
                mu[i] <- S[i]*exp(alpha-beta*S[i])
                R[i]~dlnorm(mu[i],tau)
                        }
        alpha~dnorm(0, 0.1)
        beta~dnorm(0, 10000)
        tau~dgamma(1, 5)
        sigma <- 1/sqrt(tau)
        }
```

3. Poisson distribution sample syntax:

```
model {
        for (i in 1:N) {
                mu[i] <- S[i]*exp(alpha-beta*S[i])
                R[i]~dpois(mu[i])
                        }
        alpha~dnorm(0, 0.0001)
        beta~dnorm(0, 100000)
        }
```