

Bayesian meta-analysis of inter-annual variance and serial correlation in calving rates of baleen whales

Justin G. Cooke¹

ABSTRACT

A Bayesian meta-analysis of time series of baleen whale calving rates and intervals, supplied to the 3rd Scientific Committee Workshop on baleen whale MSYR, was conducted in order to determine probability distributions of the process variance and serial correlation coefficient for calving rates and intervals. Ten series of calf counts or calving proportions, and eight series of calving intervals were used in the analysis. These two sets were analysed separately, because for most stocks both types of series were available. The analysis of calf count/calving proportion series and calving interval series show that in both cases the data are approximately equally consistent with the entire possible range $(-1, 1)$ of serial correlation coefficients. The calf count and calving proportion time series suggest that all values of inter-annual process variance in the range $[0, 1]$ are plausible, but with values near 1 being slightly less likely. The calving interval series show lower values of process variance, with $\sigma > 0.4$ being unlikely. To elucidate the relation between calving interval variance and calving rate variance probably requires explicit modelling of the calving cycle. The results do not enable direct inference of the variance of net recruitment rates, because variance in mortality rates has not been considered.

1. INTRODUCTION

The issue of the maximum sustainable yield rate (MSYR) has long been considered of central importance for the management of baleen whale populations and their exploitation. The IWC Scientific Committee has held a number of recent workshops on the topic (IWC 2009, 2010, 2011). The MSYR is related to r_0 , the rate of increase at low population sizes, sometimes called the intrinsic rate of increase. Butterworth and Best (1990) argued that, given certain assumptions, one would expect $MSYR \geq r_0/2$ (although McCall and Tatsukawa (1994) noted that this would not always be the case). In the context of deterministic models, r_0 is often referred to as the maximum rate of increase. However, as noted by Cooke (2007), it is important, in cases where reproduction and/or mortality are influenced by time-varying environmental factors, to distinguish between r_0 , the rate of increase at low population sizes under average environmental conditions, and r_{max} , the maximum rate of increase at low population size that is realised only under favourable environmental conditions.

¹ *Centre for Ecosystem Management Studies, Höllenbergstr. 7, 79312 Emmendingen, Germany.
Email: jgc@cems.de*

The three workshops on MSYR have devoted considerable attention to the issue of environmentally-caused variability. Cooke (2007) showed that estimates of MSYR derived from fitting deterministic population models to observed increasing trends in recovering populations can be positively biased in the presence of environmentally-caused variation in net recruitment rate, and that the degree of bias depends on both the level of inter-annual process variability, σ , and the inter-annual serial correlation, ρ . At the most recent workshop (IWC 2011), about 20 time series of data relating to calving rates or intervals in baleen whales were compiled for the purpose of estimating the level of variability in calving rates. In this paper, a Bayesian meta-analysis of these data sets is conducted to estimate the distribution of variability and serial correlation in calving rates in baleen whale populations, on the assumption that these data sets are typical for baleen whales.

2. DATA AND METHODS

2.1. Data

The time series relating to calf production in baleen whales that were submitted to the 3rd MSYR Workshop (IWC 2011) are listed in Table 1.

There are four types of series; raw calf counts; calving proportions (as a fraction of either known mature females or of the total stock); and population abundance.

More than one kind of series is available for several stocks. In order not to include multiple versions of the calving data from a single stock into an analysis, which would lead to pseudo-replication, the analysis was divided into two batches:

Batch A. Calf counts or proportions (10 stocks)

Batch B. Calving intervals (8 stocks)

In cases where both calf counts and calving proportion data were available, the calving proportion series was used, unless the calf count series was longer. For Gulf of St Lawrence fin whales, the calf count series (1983-2009) was used in preference to the calving proportion series (2000-2008). For Southeast Alaska humpback whales, the calving proportion series was used in preference to the calf count series (both series 1975-2008). In the case of eastern gray whales, a calving proportion series was constructed by dividing the calf count series by the estimated abundance series. In the case of raw calf count series, the possible confounding effect of population trend was allowed for by fitting a trend parameter as described below.

The resulting time series used in this analysis are listed in Table 2.

2.2. Observation error and likelihood

The observation error is presumed to be of Poisson, binomial, normal or lognormal kind depending on the series, as listed in Table 2. Poisson errors are assumed for raw calf counts, and binomial errors for directly observed calving proportions. In the case of the BCB bowhead calving proportion series, normal errors with the quoted standard error are assumed. The calving interval data are treated as normally distributed with the empirically observed variance of intervals for each data point. The calculated calving proportions for the eastern gray whale are assumed to be lognormally distributed, with the CV calculated from the CVs of the calf number and total abundance estimates.

2.3. Modelling of process variance

The expected value of the data point for series i in year t is assumed to be given by:

$$z_{it} = \exp(a_i + b_i t + \varepsilon_{it})$$

where a_i , b_i are stock-specific intercept and slope parameters. The slope parameter is omitted (set to zero) for all time series except the absolute calf count series. a_i is treated as a nuisance parameter that is eliminated by conditioning on the sufficient statistic for a_i .

Two alternative assumptions for the nature of the process variance were considered:

- (i) No serial correlation: the ε_{it} are independent normal random variates with mean zero and series-specific standard deviation σ_i ,
- (ii) Serial correlation: the ε_{it} are normal random variates with mean zero, series-specific standard deviation σ_i , are independent between series, but are serially correlated over time within a series. The series-specific serial correlation coefficient is denoted by:

$$\rho(\varepsilon_{i,t}, \varepsilon_{i,t+1}) = \rho_i.$$

2.4. General approach to Bayesian meta-analysis

The assumption behind Bayesian meta-analysis (Congdon 2003) is that the value of a given parameter for a given stock is drawn from a population of values of that parameter across stocks.

By scrambling (anonymizing) the stock labels, we can justify the assumption that the prior distribution of the parameter of interest is the same for each stock.

If data are collected for a number of stocks, these contribute to knowledge of the distribution of the parameter of interest across stocks, and influence our prior expectations of the likely value of the parameter for further stocks. This implies that the prior distributions are not independent across stocks. We suppose that the correlation of the priors for a given parameters between any pair of distinct stocks is ζ where $0 \leq \zeta \leq 1$. (Note that negative mutual correlation between three or more random variables is not possible).

The extreme case $\zeta = 0$ corresponds to the case where the priors are independent. In this case, regardless of for how many stocks we have observed the value of the parameter, this knowledge would have no influence on our expectation of the value of the parameter for the next stock. Apart from being scientifically unreasonable, this case corresponds to multiple single-stock analyses rather than a true meta-analysis.

The other extreme case, $\zeta = 1$, corresponds to the assumption that the value of the parameter is the same for all stocks. This would mean that data collected from stock A would be as informative about the value of the parameter for stock B as would data collected from stock B itself. Apart from being scientifically unreasonable, this case involves estimation of common parameters applicable to all stocks, and is also not a true meta-analysis.

Values of ζ that lie strictly between 0 and 1 correspond to a true meta-analysis.

In the case where the prior distribution for a parameter of is non-normal, it is more convenient to let ζ refer to the correlation between the prior distributions after the parameter has been transformed to makes its prior normal. If the cumulative prior distribution for a parameter p is $F(p)$, then we apply the transformation $p \rightarrow p^*$ where $p^* = \Phi^{-1}(F(p))$ and Φ^{-1} is the inverse of the cumulative normal distribution function. ζ then refers to the correlation between the priors of p_i^* and p_j^* for any two distinct stocks i, j . This form of correlation is sometimes called the normalized correlation.

2.5. Priors for each parameter

The unknown parameters for each stock and their assumed priors are:

σ_i : standard deviation of process error: prior is uniform on $[0,1]$.

ρ_i : the serial correlation of process error : prior is uniform on $(-1,1)$

b_i : trend parameter (for calf counts): prior is normal $N(0,0.1)$

The unknown meta-parameters and their priors are:

ζ_σ : normalized correlation between priors for σ_i , σ_j $i \neq j$: prior is uniform on $(0,1)$

ζ_ρ : normalized correlation between priors for ρ_i , ρ_j $i \neq j$: prior is uniform on $(0,1)$

ζ_b : normalized correlation between priors for b_i , b_j $i \neq j$: prior is uniform on $(0,1)$

The upper bound on the prior for σ_i is arbitrary at this stage but as found below is not critical.

2.6. Computation of posterior distributions

The distributions of interest are the posterior distributions of σ and ρ for a generic stock; that is, a stock for which there are no stock-specific data. We denote the generic stock as stock 0 and its parameters as ρ_0 and σ_0 .

Because the prior distributions of the stock-specific parameters are correlated across stocks, the posterior distribution must be evaluated for all parameters from all stocks jointly. A sample of the joint posterior distribution of all parameters is obtained using an MCMC algorithm, with sampling every 100th iteration, and discarding the first half of the chain (regardless of chain length) as “burn-in”. The chain is sampled long enough until the posterior distributions are deemed to have converged satisfactorily. Convergence is examined by computing the 10th, 50th and 90th percentiles of the posterior distribution computed from a rolling subset of the chain: this subset being the most recent 20% of iterations at each point. These percentiles should stabilise when the chain is long enough.

3. RESULTS AND DISCUSSION

3.1. Convergence of the MCMC

Figs 1a-b show the 10th, 50th and 90th percentiles of the distributions of σ_0 , ρ_0 , ξ_σ and ξ_ρ as a function of iteration number for a chain of length 5 million. As noted above, the percentiles are calculated on a rolling basis using the trailing 20% of the chain up to each point. The results suggest that the posterior distributions for σ_0 and ρ_0 have stabilised after approx. 1 million iterations. The posterior distributions of ξ_σ and ξ_ρ are not quite so stable, but appear satisfactory for low-precision work, since it is the stability of the distributions of the parameters of interest, σ_0 , ρ_0 , that is of primary importance. For the results that follow, a chain length of 5 million sampled every 100th iteration was used, with the first half the chain being discarded as burn-in.

3.2. Calf counts/proportions series (batch A)

Figs 2a-b show the posterior distributions of σ_0 and ξ_σ for the case of no serial correlation ($\rho = 0$).

Figs 3a-d show, for the case of estimated serial correlation, the posterior distributions of : (a) σ_0 ; (b) ρ_0 ; (c) ξ_σ ; (d) ξ_ρ .

The posterior distribution of σ_0 is rather insensitive to whether serial correlation is allowed or not. The posterior distribution of σ_0 is fairly flat out to $\sigma_0 = 0.5$, and declines only moderately for $\sigma_0 > 0.5$. The posterior implies that there are no strong grounds to exclude any of the range [0,1] except that values near the upper end of the range are slightly less likely. The posterior distribution for ρ_0 indicates no strong grounds to exclude any values within the range (-1,1).

3.3. Calving interval series (batch B)

Figs 4a-b and 5a-d show the same results as in Figs 2-3 but for the calving interval series (batch B). The posterior distribution for σ_0 is again rather insensitive to whether serial correlation is allowed or not. The mode of the posterior is near $\sigma_0 = 0.05$, and values of $\sigma > 0.4$ are effectively excluded. The posterior distribution of ρ_0 is almost uniform across almost the entire range (0,1): this implies that these data are uninformative with respect to ρ_0 .

3.4. Discussion

The lower estimates of σ_0 using the calving-interval data may be consistent with higher values of σ_0 for calf-count or proportion data, because the intervals by definition measure a rolling average of rates smeared over a number of consecutive years. To relate the observed variation in calving interval to variation in effective reproductive rate may require explicit modelling of the calving cycle and its variation, as done for example by Cooke (2003).

A further problem with interpreting the calving interval data in this context is that both the observation and process variance are estimated empirically from the calving interval data. The power to separate the two may be less than for the calf count and proportion data, where the variance is either given theoretically (Poisson, binomial) or estimated externally.

It would appear more appropriate to use the results from the calf count and calf proportion data in the first instance. Based on the posterior distributions of σ_0 and ρ_0 , there seems little reason to change the practice of IWC (2010, Table 2) of using, for simulation purposes, values of 0.0, 0.5, and 1.0 as “low”, “medium” and “high” values of σ , and the values 0, 0.5 and 0.9 for low medium and high values of ρ , noting that values of $\rho < 0$, if they occur, are unproblematic in the sense that they tend to dampen rather than enhance the effect of a given level inter-annual variation.

It should be emphasised that variance in reproductive rates does not translate directly into variance in the net recruitment rate, because there may also be variability in mortality rates. An important question in that context is whether variations in mortality rates are likely to be positively or negatively correlated with variations in reproductive rates (Cooke, 2011).

ACKNOWLEDGEMENTS

Special thanks are due to those who supplied data sets to the 3rd Scientific Committee workshop on MSYR (IWC 2011), which are listed in Table 1 and used in this paper (Table 2): P. Best, R. Brownell Jr., J. Calambokidis, C. Gabriele, W. Koski, S. Kraus, J. Laake, W. Perryman, C. Ramp, J. Robbins, R. Sears, J. Straley, V. Rowntree.

REFERENCES

Butterworth D.S. and Best P.B. 1990. Implications of the recovery rate of the South African right whale population for baleen whale population dynamics. *Rep. int. Whal. Commn* 40:433-447.

Cooke J., Rowntree V., Payne R. 2003. Analysis of inter-annual variation in reproductive success of South Atlantic right whale (*Eubalaena australis*) from photo-identification of calving females observed off Peninsula Valdes, Argentina, during 1971-2000. Paper SC/55/O23.

Congdon P. 2003. *Applied Bayesian Analysis*. Wiley Series in Probability and Statistics. 450pp.

Cooke J.G. 2007. The influence of environmental variability on baleen whale sustainable yield curves. Paper SC/N07/MSYR1 presented to the MSYR workshop, November 2007.

Cooke J.G. 2011. A note on the expected relationship between variability in reproductive rate and variation in net recruitment rate based on life history trade-off models. *J. Cetacean Res. Manage* 12 (Suppl.) :99-101.

IWC. 2009. Report of the MSYR Workshop, November 2007. *J. Cetacean Res. Manage*. 11 (Suppl.): 467-80.

IWC. 2010. Report of the 2nd intersessional workshop on the review of MSYR for baleen whales, February 2009. *J. Cetacean Res. Manage*. 11 (Suppl. 2): 493-508.

IWC. 2011. Report of the 3rd intersessional workshop on the review of MSYR for baleen whales, April 2010. *J. Cetacean Res. Manage*. 12 (Suppl.): 401-11.

MacCall A.D. and Tatsukawa K. 1994. Theoretical effects of habitat selection on distribution and productivity of whales. *Rep. int. Whal. Commn* 44:407-12.

LIST OF TABLES

1. Table 1. Time series of calf numbers, proportions, intervals and associated data submitted to the 3rd Scientific Committee Workshop on baleen whale MSYR (adapted from IWC 2011, Table 1).
2. Table 2. Time series used in the analyses.

LIST OF FIGURES

Fig. 1. Convergence of MCMC.

Fig. 2. Posteriors for (a) sigma (b) zeta(sigma) (counts/propn) (rho = 0)

Fig. 3. Posteriors for (a) sigma (b) rho (c) zeta(sigma) (d) zeta(rho) (counts/propn)

Fig. 4. Posteriors for (a) sigma (b) zeta(sigma) (calving intervals) (rho = 0)

Fig. 5. Posteriors for (a) sigma (b) rho (c) zeta(sigma) (d) zeta(rho) (calving intervals)

Table 1. Time series of calf numbers, proportions, intervals and associated data submitted to the 3rd Scientific Committee Workshop on baleen whale MSYR (adapted from IWC 2011, Table 1)

Species	Stock	Series type	1st year	Last year	No. Years	Notes	Data supplier
Blue	Gulf of California	Calving proportion	1986	2009	18	for known rep. Females	Sears/Ramp
Blue	Gulf of California	Intervals	1990	2005	7		
Bowhead	Bering/Beaufort/Chuckchi seas	Calf proportion	1985	2004	8	as fraction of total stock	Koski
Fin	Gulf of St Lawrence	Calf Count	1983	2009	27		
Fin	Gulf of St Lawrence	Calving proportion	2000	2008	5	for known rep. Females	Sears/Ramp
Fin	Gulf of St Lawrence	Intervals	2007	2008	2		
Gray	Eastern North Pacific	Abundance	1968	2007	23	1+	Laake, Perryman & Brownell
Gray	Eastern North Pacific	Calf Count	1994	2009	16		
Humpback	Gulf of Maine	Calving proportion	1979	2005	27	for known rep. Females	Robbins
Humpback	Gulf of Maine	Intervals	1984	2005	22		
Humpback	Gulf of St Lawrence	Calving proportion	1983	2009	25	for known rep. Females	Sears/Ramp
Humpback	Gulf of St Lawrence	Intervals	1985	2009	18		
Humpback	Southeast Alaska	Calf Count	1975	2008	34		
Humpback	Southeast Alaska	Calving proportion	1975	2008	34	for known rep. Females	Gabrele/Straley
Humpback	Southeast Alaska	Intervals	1986	2008	23		
Humpback	US West Coast	Abundance	1991	2008	18	total	Calambokidis
Right	Northwest Atlantic	Calf Count	1980	2009	30		Kraus
Right	Northwest Atlantic	Intervals	1985	2009	25		
Right	Southeast Atlantic	Calf Count	1979	2006	28		Best
Right	Southeast Atlantic	Intervals	1984	2006	23		
Right	Southwest Atlantic	Calf Count	1971	2008	38		Rowntree
Right	Southwest Atlantic	Intervals	1977	2008	32		

Note: no. of years does not always equal the span of years, because of missing years

Table 2. Time series used in the analyses

Batch	Species	Stock	Series type	Observation error	1st year	Last year	No. Years	Notes
A	Gray	Eastern North Pacific	Calving rate	Lognormal	1994	2007	23	Ratio calves to 1+ stock for known rep. Females as fraction of total stock for known rep. Females for known rep. Females for known rep. Females
A	Fin	Gulf of St Lawrence	Calf Count	Poisson	1983	2009	27	
A	Right	Northwest Atlantic	Calf Count	Poisson	1980	2009	30	
A	Right	Southeast Atlantic	Calf Count	Poisson	1979	2006	28	
A	Right	Southwest Atlantic	Calf Count	Poisson	1971	2008	38	
A	Blue	Gulf of California	Calving proportion	Binomial	1986	2009	18	
A	Bowhead	Bering/Beaufort/Chuckchi seas	Calf proportion	Normal	1985	2004	8	
A	Humpback	Gulf of Maine	Calving proportion	Binomial	1979	2005	27	
A	Humpback	Gulf of St Lawrence	Calving proportion	Binomial	1983	2009	25	
A	Humpback	Southeast Alaska	Calving proportion	Binomial	1975	2008	34	
B	Blue	Gulf of California	Intervals	Normal	1990	2005	7	
B	Fin	Gulf of St Lawrence	Intervals	Normal	2007	2008	2	
B	Humpback	Gulf of Maine	Intervals	Normal	1984	2005	22	
B	Humpback	Gulf of St Lawrence	Intervals	Normal	1985	2009	18	
B	Humpback	Southeast Alaska	Intervals	Normal	1986	2008	23	
B	Right	Northwest Atlantic	Intervals	Normal	1985	2009	25	
B	Right	Southeast Atlantic	Intervals	Normal	1984	2006	23	
B	Right	Southwest Atlantic	Intervals	Normal	1977	2008	32	

Fig. 1. Convergence of MCMC

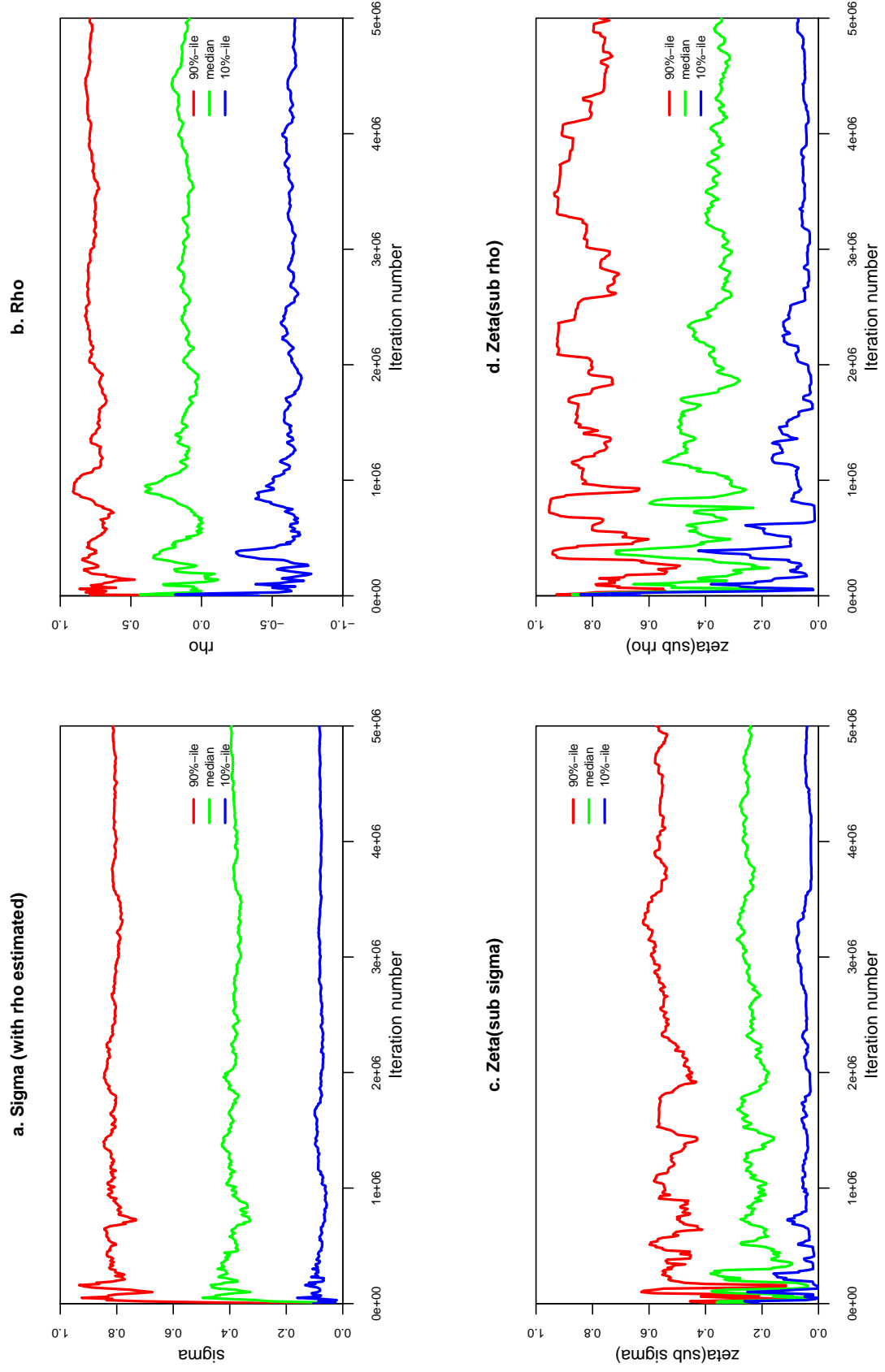


Fig. 2. Posteriors for sigma, zeta(sigma) (counts/propn) (rho = 0)

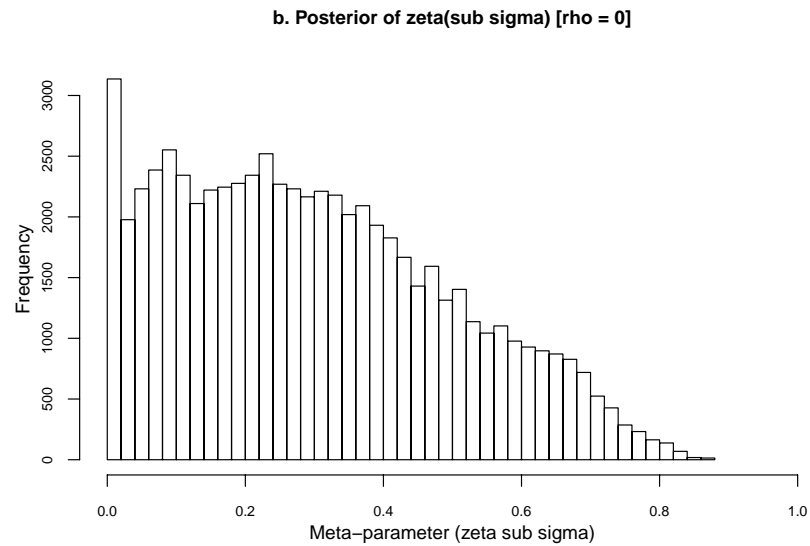
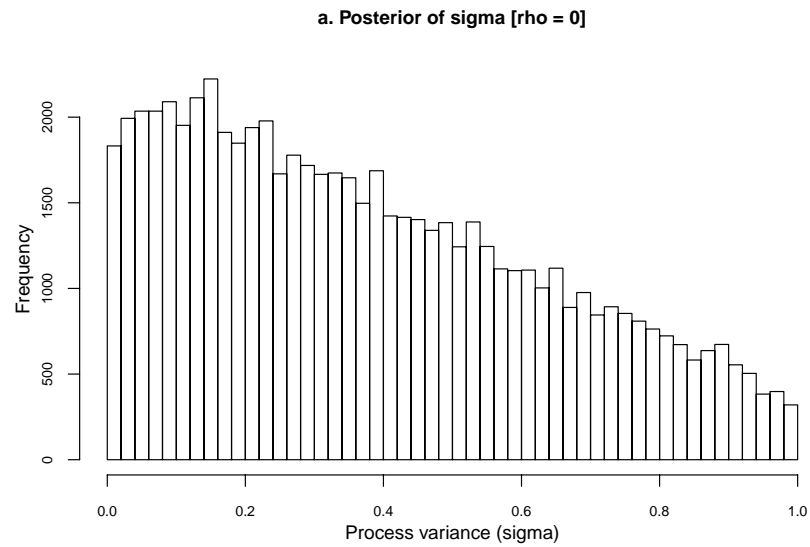


Fig. 3. Posteriors for sigma, zeta(sigma), rho, zeta(rho) (counts/propn)

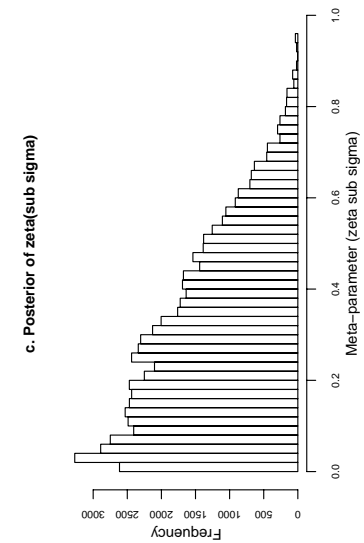
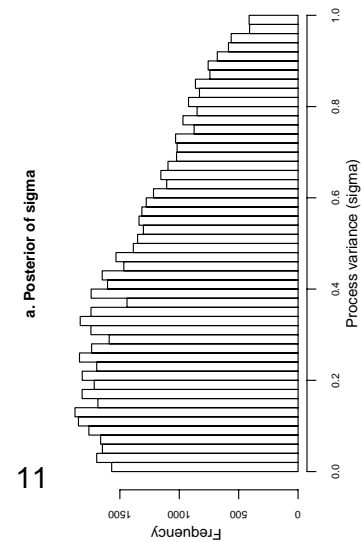
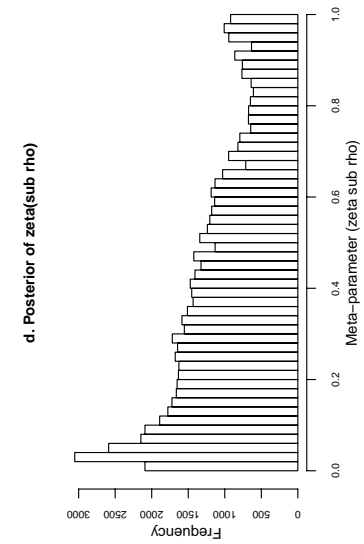
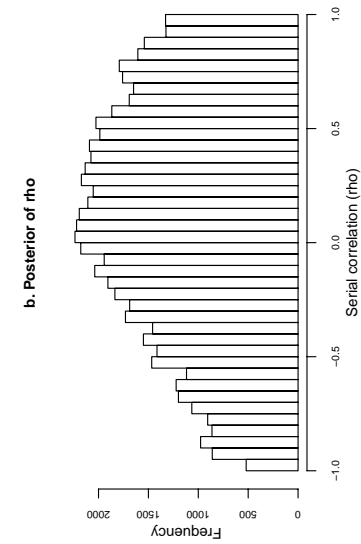


Fig. 4. Posteriors for sigma, zeta(sigma) (calving intervals) ($\rho = 0$)

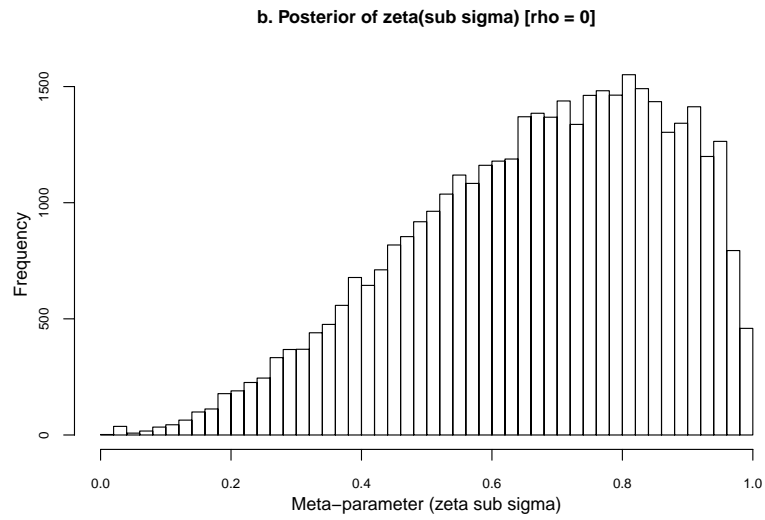
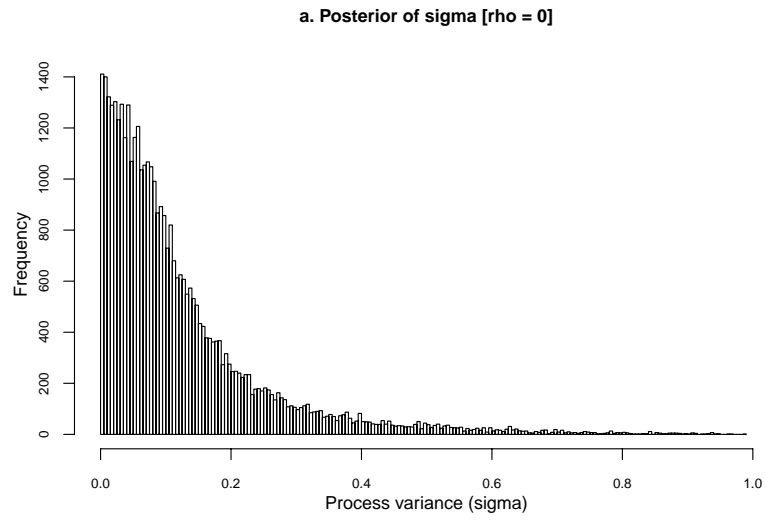


Fig. 5. Posteriors for sigma, zeta(sigma), ρ , zeta(ρ) (calving intervals)

