

BEAST

Bayesian Ecological Analysis of Statistical Trends

User Manual v. 1.0.7



Eric J. Ward
University of Washington
warde@u.washington.edu

Introduction

The BEAST software has been written to aid ecologists in conducting Bayesian analyses of trends in population abundance. This manual contains an overview of the software , and several examples of its use.

	<u>Page</u>
1) Models	
Models supported	1
Types of error	5
Likelihoods	6
Advanced options: time lags , change points	6
Parameters	7
2) Data	
Entering Data	9
3) MCMC Analysis	
Burn-In	10
MCMC	10
4) Projections	
Entering Projections	11
Creating Probability Tables	11
5) Decision Analysis	
Conducting Decision Analyses	12
6) Viewing Output	15
7) Appendix of Distributions	
Distributions Supported	17
8) Examples	
Example 1	31

Model Form

Three forms of single-species models are currently supported:

- 1) [Single stage / age models](#)
- 2) [Stage-structured models](#)
- 3) [Age-structured models](#)

Single Stage / Age Models

There are 7 single stage / age models available:

- 1) **Random Drift Model:** $N_t = N_{t-L}$, where 'L' is a positive integer representing the time lag.
- 2) **Exponential Model:** $N_t = N_{t-L} \cdot \exp(r)$
 1. Parameter 'r' = growth rate
- 3) **Theta-Logistic Model:** $N_t = N_{t-L} + N_{t-L} \cdot r \cdot \left(1 - \left(\frac{N_{t-L}}{K}\right)^c\right)$
 1. Parameter 'r' = growth rate
 2. Parameter 'K' = carrying capacity
 3. Parameter 'c' = strength of density dependence
- 4) **Theta-Ricker Model:** $N_t = N_{t-L} \cdot \exp\left(r \cdot \left(1 - \left(\frac{N_{t-L}}{K}\right)^c\right)\right)$
 1. Parameter 'r' = growth rate
 2. Parameter 'K' = carrying capacity
 3. Parameter 'c' = strength of density dependence
- 5) **Depensation Model (Turchin 2000):** $N_t = \left(\frac{a \cdot N_{t-L}^c}{1 + b \cdot N_{t-L}^c}\right)$
 1. Parameter 'a' = controls the growth rate
 2. Parameter 'b' = scaling parameter
 3. Parameter 'c' = strength of density dependence
- 6) **Quadratic Allee Model (Lewis and Kareiva 1993):**
$$N_t = N_{t-L} + \frac{b \cdot N_{t-L} \cdot (N_{t-L} - a)(K - N_{t-L})}{K^2}$$
 1. Parameter 'a' = population required for non-negative growth
 2. Parameter 'b' = maximum possible growth rate
 3. Parameter 'K' = carrying capacity
- 7) **Gompertz Model:** $N_t = N_{t-L} \cdot a \cdot \exp^{-b \cdot \ln(N_{t-L})^c}$
 1. Parameter 'a' = growth rate

2. Parameter 'b' = scaling parameter (controls carrying capacity)
3. Parameter 'c' = strength of density dependence

Stage Structured Models

Up to 100 individual stages are available for any particular model. When the stage-structured modelling table appears, the user is able to specify density dependence in any of the yellow cells (cells in the first row represent fecundities, cells in the sub-diagonal represent survival probabilities from one stage to the next). In addition to the three density dependent [functions](#) available, each cell may be assigned a single parameter (or constant, including 0.0) by selecting the first option in the check box or leaving the check box empty to begin with. A third option is that [children](#) may be assigned to any cell. If no function is assigned to any of the cells in the model, the model will exhibit exponential population growth.

The image shows a screenshot of a software interface for stage-structured models. It features a grid with columns labeled 3 through 8. The first row contains yellow cells with dropdown arrows, representing fecundities. The sub-diagonal cells (e.g., row 2, column 3; row 3, column 4) are also yellow with dropdown arrows, representing survival probabilities. A dropdown menu is open over the cell at row 2, column 7, showing the following options: Logistic, Logarithmic, Hassell, Child (1), Child (2), Child (3), and Child (4). The grid has a scroll bar on the right and a horizontal scroll bar at the bottom.

Age Structured Models

Like with stage structured models, the range of ages supported is 2-100. Unlike stage structured models, however, each age is not allowed to have its own density dependence function. Density dependence may be incorporated into the model in one of three ways:

- 1) Fecundity
 - a. Fecundity is assumed to occur equally for all ages greater than the age at reproductive maturity (ASM)
- 2) Juvenile Survival
 - a. Juvenile survival is assumed to occur equally for all ages less than or equal to the age of transition to adult survival (ATM)
- 3) Adult Survival

- a. Adult survival is assumed to occur equally for all ages greater than the age of transition to adult survival (ATM)

Each aspect of the model may be assigned a unique [function](#), or no functions may be assigned (again, exponential growth will result).

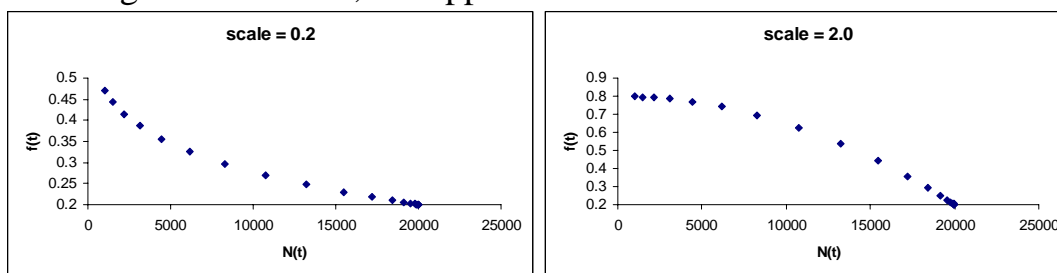
Density Dependent Functions

Three density dependent functions are available for both the survival and fecundity parameters (only the fecundity equations are shown, but survival equations are identical when 'f' is replaced by 's'):

1) **Logistic:**
$$f_t = f_0 + (f_{\max} - f_0) \cdot \left(1 - \left(\frac{N_t}{K} \right)^{f_{\text{scale}}} \right)$$

1. Parameter 'f₀' = minimum fecundity
2. Parameter 'f_{max}' = maximum fecundity
3. Parameter 'f_{scale}' = scaling parameter
4. Parameter 'N_t' = current population size
5. Parameter 'K' = carrying capacity

Notes: when the scale parameter is 1.0, the fecundity rate decreases linearly with population size. When the parameter is less than 1.0, the fecundity rate decreases exponentially with population size. When the scale is greater than 1.0, the opposite is true.



2) **Logarithmic:**
$$f_t = f_0 + (f_{\max} - f_0) \cdot \left(1 - \left(\frac{\ln(N_t)}{\ln(K)} \right)^{f_{\text{scale}}} \right)$$

1. Parameter 'f₀' = minimum fecundity
2. Parameter 'f_{max}' = maximum fecundity
3. Parameter 'f_{scale}' = scaling parameter
4. Parameter 'N_t' = current population size
5. Parameter 'K' = carrying capacity

Notes: behaves similarly to the logistic, but is a function of the logarithm of population size. When the scale is ~ 8.6, fecundity declines linearly with population size.

3) **Hassell:**
$$f_t = \frac{f_{\max}}{\left(1 + \frac{N_t}{K}\right)^B}$$

1. Parameter 'f_{max}' = maximum fecundity
2. Parameter 'B' = controls scale
3. Parameter 'N_t' = current population size
4. Parameter 'K' = carrying capacity
5. Notes: can behave similarly to logistic, depending on value of the scale parameter. The value of the scale, B, can also be used to solve for the minimum fecundity rate: $f_0 = \frac{f_{\max}}{2^B}$.

4) **Linear :**
$$f_t = f_0 + (f_{\max} - f_0) \cdot \left(\frac{K - N_t}{K}\right)$$

1. Parameter 'f_{max}' = maximum fecundity
2. Parameter 'f_{min}' = minimum fecundity
3. Notes: When the population is very far from carrying capacity, the term in brackets will be close to 1.0, and as the population approaches zero, the fecundity at time (t) will converge to the minimum fecundity.

5) **Ricker:**
$$f_t = f_0 \cdot \exp(-b \cdot N_t^c)$$

1. Parameter 'f_{max}' = maximum fecundity
2. Parameter 'b' = scale parameter
3. Parameter 'c' = scaling exponent

6) **Gompertz:**
$$f_t = f_0 \cdot \exp(-b \cdot \ln(N_t)^c)$$

1. Parameter 'f_{max}' = maximum fecundity
2. Parameter 'b' = scale parameter
3. Parameter 'c' = scaling exponent

Children in Stage / Age Structured Models

The “child” option in stage / age structured model operates on the same principle that is used for assigning individual parameters to be children. This option is useful when two cells are assumed to have the same value. For example, consider a 3-stage model, consisting of juveniles, sub-adults, and adults. If juvenile survival is assumed to be logistic, and sub-adult survival is assumed to be logistic, the parameters for each function may be estimated independently (by assigning ‘Logistic’ to each combo box). If sub-adult survival is assumed to follow the same form, and take on the same values as juvenile survival, we could select the ‘Child(1)’ option. The

converse is also true – sub-adult survival could be assigned ‘Logistic’, and juvenile survival could be assigned ‘Child(2)’ (Caution: this second approach could lead to confusion for more complicated models).

Error Form and Type

Error may be incorporated into any of the models either as observation error or process error, and the error may be either an additive or multiplicative term. The key differences involve (1) how the likelihoods are calculated, and (2) how the population projections are done.

1) Observation Error Model

- a. Multiplicative Error: $\hat{N}_t = P_t \cdot \varepsilon_t$, where \hat{N}_t is the observed population size at time t , P_t is the predicted population size at time t , and ε_t is the error term
- b. Additive Error: $\hat{N}_t = P_t + \varepsilon_t$, where \hat{N}_t is the observed population size at time t , P_t is the predicted population size at time t , and ε_t is the error term

2) Process Error Model

- a. Multiplicative Error: $S_t = \hat{S}_t \cdot \varepsilon_t$, where S_t is the observed population state at time t , \hat{S}_t is the predicted population state at time t , and ε_t is the error term
- b. Additive Error: $S_t = \hat{S}_t + \varepsilon_t$, where S_t is the observed population state at time t , \hat{S}_t is the predicted population state at time t , and ε_t is the error term

Error Distribution (Likelihoods)

Selecting the appropriate error distribution is one of the most important aspects of the Bayesian analysis. Currently, the BEAST program supports a total of 10 likelihood functions (some of which are symmetric, some of which are asymmetric). The likelihood options are determined by whether the error is additive or multiplicative. For more information on these distributions, see the [appendix](#).

Additive Error Distributions ($E[X] = 0.0$)

Normal ($0, \sigma$)

Gamma (a, b)

Student's-t ($0, \sigma, df$)

Double Exponential (0, b)
Gamma (u, b)

Multiplicative Error Distributions ($E[X] = 1.0$)

Lognormal (0, s)
Gamma (a, b)
Exponential (b)
Half-Normal (a, b)
Gamma (u, b)

Time Lag

The time lag in the model is allowed to be changed (1-5), but is not allowed to be a random variable. The reason for this is that assigning different lags to the same data set results in different portions of the data being used by the likelihood function. For the likelihoods to be comparable (in MCMC), the same portions of the data must be used.

Change Point

The change point of the model allows the user to create ‘broken stick’ models, where some (or all) of the population parameters in one portion of the time series are different from the same parameters in different portion of the time series. Such models might result from drastic shifts in ecosystem productivity or from human disturbances.

Assigning Parameters

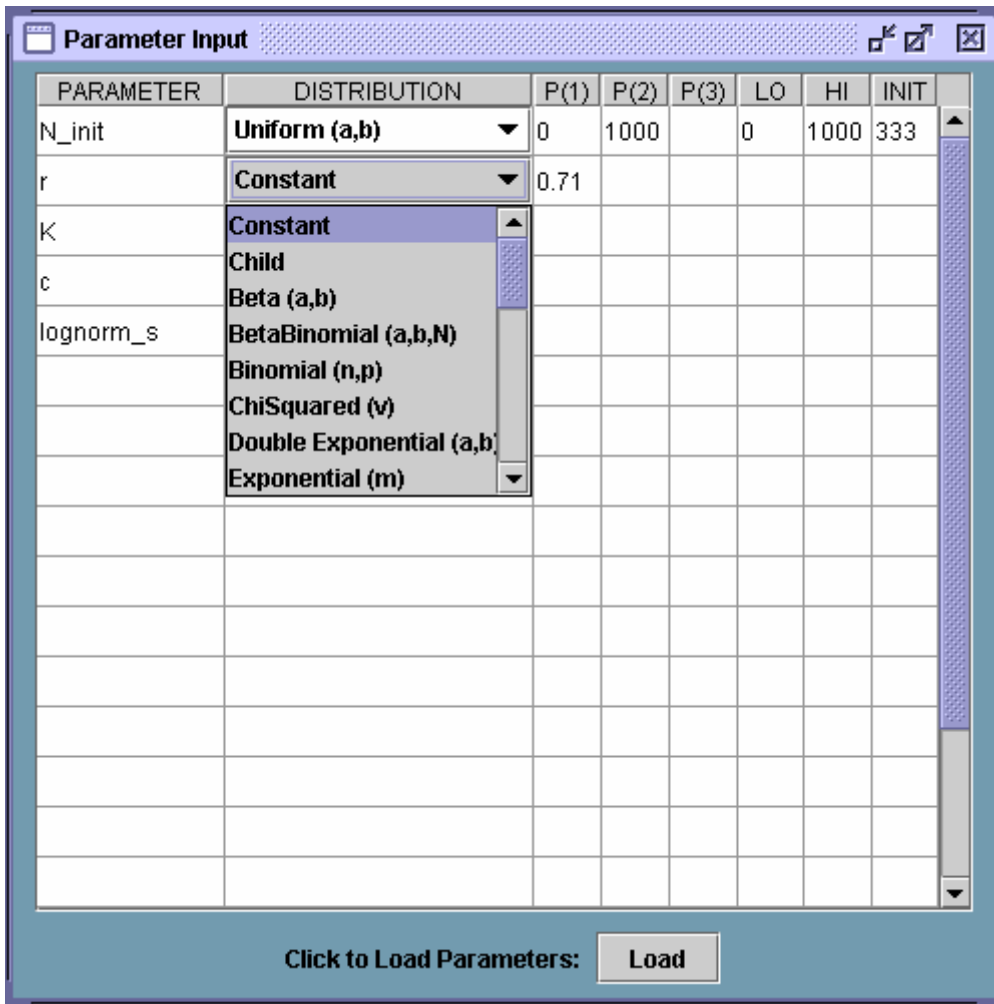
Once the parameter window is opened, the user has three options for each parameter:

- 1) Parameters may be assigned constant (known) values
 - a. All parameters that are constants must have ‘Constant’ selected from the corresponding combo box in the second column. The value of the parameter must be entered in the 3rd column, under the heading ‘P0’.
- 2) Parameters may be assigned prior distributions
 - a. Columns 3-5 of the parameter window represent the parameters necessary for each prior distribution. For distributions that require a single parameter (e.g. Geometric), only the cell that corresponds to ‘P0’ must be entered. For distributions requiring two parameters (e.g. Normal (u, σ)), the cells corresponding to ‘P0’ and ‘P1’ must be filled in – ‘P0’

representing u , and 'P1' representing σ . All three parameters must be filled in for distributions requiring three parameters (e.g. Student's-t). Columns 6-7 of the parameter window represent the lower and upper bounds for the parameter, and are required for all distributions. If bounds are infinite, the lower and upper bounds can be set to values that are very large in magnitude. The final 8th column of the parameter window represents the initial value that the MCMC chain should be initialized at, and is required for all parameters.

- 3) Parameters may be children of another parameter
 - a. Child parameters cannot be assigned in models where the change point is not present. In change-point models, children can only be assigned in the second half of the time series. For example, if a population is assumed to follow a change-point logistic model, any of the growth parameters (r , K) or the error parameters in the second half of the time series may be assigned to be children. No values need to be entered in columns 3-8.

Example: The following model is a 'Theta-logistic model' with lognormal observation error, and no change point. In this model, the parameter ' N_0 ' has been assigned a prior distribution that is \sim Uniform (0, 1000). The lower and upper bounds for the parameter are assigned values of (0, 1000), which is redundant for the uniform distribution, but required for all parameters that are assigned priors. The final column in the first row of the table indicates that the parameter ' N_0 ' has been initialized to a value of 333.0. The second parameter, ' r ' is assigned a constant value of 0.71 – note that no bounds or values need to be entered.



Data

The data panel consists of a table compatible with Excel copying / pasting. There are four columns in the table: 'Year', 'Abundance', 'Harvest', 'Scale'. The 'Year' and 'Abundance' columns are relatively straightforward – 'Year' representing the time period, and 'Abundance' the corresponding abundance estimate. The 'Harvest' column represents the annual harvest (harvest occurring after population growth).

The final column, 'Scale', represents the uncertainty involved in the estimate. By default, when all cells in the column are left empty (or all are set to equal values), the errors of the estimates are assumed to be equal (for the lognormal model, this assumes the CVs to be equal). The 'Scale' column of the data offers additional flexibility when the CVs (or variances)

are assumed to be unequal between observations. When a single parameter (e.g. sigma in the normal or Student's-t distribution) is responsible for controlling the variance, the scale parameter is treated as a proportionality constant. When multiple parameters are responsible for controlling the variance (e.g. gamma distribution), the scale is incorporated into the second of the two parameters (for the gamma, the 'b' parameter). When all data are entered, click the 'Load' button. To clear the table, click 'Clear'.

Example: Consider the following time series, spanning 16 years. Abundance estimates exist for all but two years (1983 is shown because harvest was known to occur, whereas 1982 is not entered because no abundance estimate exists, and no harvest took place). Along with the known harvest, we assume that the error involved with the estimate in 1983 is twice as large as the error of all other estimates (note: these values are relative, so instead of using 1.0 and 2.0, any other numbers – 0.5 and 1.0, 3.5 and 7.0 may be entered to produce the same result.

MCMC

The MCMC sampling proceeds from where the burn-in stopped. A new output file needs to be named for each run (old files can be overwritten). The format of the run name needs to appear in single quotes, e.g. 'D:/Beast/run01'. The amount of thinning for MCMC sampling and number of samples to save is used to calculate the total chain length. For example, if the amount of thinning is set at 2 (every other sample is saved), and the number of samples is set at 5,000,000, the total chain length will be 10,000,000. As a second example, if the thinning is set at 1000, and the number of samples is set at 5000, the total chain length will be 5,000,000. Of the total chain length, only samples designated as 'saved' are written to output files. Output files for the analysis include a file containing parameter inputs (thisRun.inp), a file containing population projections (thisRun.prj), a file containing MCMC posterior distributions (thisRun.par; the likelihood and joint prior are stored in the second to last and last columns), and a file containing the means, variances, and correlations between parameters (thisRun.cor).

Projections

The number of population projections is initially entered on the 'MCMC – Control' panel, and is optional – if set to 0, the estimated population size is output (if the number is greater than 0, a series of population estimates will be outputted for each MCMC sample). How the projections are done depends on whether the model is an observation error model or process error model. If the model is an observation error model, the population estimates will reflect the projected population size with no error (because the expected value of the error term in each year is 1.0 for multiplicative error and 0.0 for additive error). If the model is a process error model, future population projections will incorporate the stochastic error.

Probability Tables

Before conducting a formal decision analysis, it may be useful to conduct a simple matrix of probabilities for future population states. In simple terms, these represent the conditional probabilities for the future projections – “given that the population is currently between W and X, what is the probability that the population T time steps later will be between Y and Z?”

To conduct these types of probability tables, the number of bins for the current population size and future population size must be assigned. Because this analysis is relatively simple, the number of bins is limited to be less than 5. Once these values are selected (and the number of future years is entered), the intervals corresponding to those bins needs to be entered. The format needs to be comma-delimited, and can be pasted into the text fields. For example, if the bins are:

- Bin #1 (range = 0, 100)
- Bin #2 (range = 100, 500)
- Bin #3 (range = 500, 600)
- Bin #4 (range = 600, 900)
- Bin #5 (range = 900, 100000000)

then the text string that needs to be entered is “0, 100, 500, 600, 900, 100000000”. The range on Bin #5 is set to an arbitrarily large number, in place of infinity. When both strings are pasted into the text fields (they need not be equal), click ‘Analyze Projections’. A simple table will be displayed, with rows representing the future population states, and columns representing the current population state.

Future Current	Interval 1	Interval 2
Interval 1	0.0822	0.0
Interval 2	0.255	0.1218

These probabilities do not add up to 1.0 because not all of the estimated population sizes are contained within the 2 intervals. The table can be read as follows: given that the current population size lies in interval #1, there is an 8% chance that the population will remain in that interval, and a 25% chance that the population will end up in interval #2.

Decision Analysis

The ‘decision analysis’ panel was created to allow users to conduct formal Bayesian decision analyses through the use of cost functions and Bayes factors. When the number of intervals (and number of future years) is selected, a table will appear as follows:

DECISION	BOUND(L)	1	2	3	4
Interval 1					
Interval 2					
Interval 3					
Interval 4					

The first column, 'Bound(L)' of the table represents the lower bounds for the intervals. For example, for the following bins:

Bin #1 (range = 0, 500)

Bin #2 (range = 500, 600)

Bin #3 (range = 600, 900)

Bin #4 (range = 900, 100000000)

then the first column of data should be the vector (0, 500, 600, 900). The remaining column headings in the table represent the states of nature, while the rows represent the decisions made. To conduct a Bayesian decision analysis, the table needs to be filled in with costs, or penalties for making the wrong decision.

Example: the following table represents a decision analysis with 4 states:

	States of Nature			
Decision	Interval 1	Interval 2	Interval 3	Interval 4
Interval 1	0.0	0.05	0.25	1.0
Interval 2	0.05	0.0	0.05	0.25
Interval 3	0.25	0.05	0.0	0.05
Interval 4	1.0	0.25	0.05	0.0

Each cell value is the cost of making the corresponding decision, given the state of nature. The main diagonal of the table are zeros because there is no cost associated with making wrong decisions. The table is also symmetric for this problem – given the state of nature is Interval #1, the cost of deciding the state of nature is Interval #4 is the same as when the true state of nature is Interval #4 and Interval #1 is chosen.

After the table is filled in, and the 'Calculate Bayes Factor' button is checked, a new output file (thisRun.baf) is created, containing the bayes factor and decision analysis output.

Output

Because various programs already exist in other packages, BEAST has no graphical capabilities. MCMC output from BEAST is easily viewed in R. R can be downloaded freely from: <http://cran.r-project.org/>. After installation, make sure that the CODA package is installed (in the Windows version, under “Packages” → “Install Packages From CRAN” → “CODA”). Once CODA has been installed, load the library with the command: ‘library(coda)’.

The two most common types of output that could be analyzed in R are making density plots of parameters, and analyzing MCMC chains.

1) Producing Density Plots in R

Producing density plots does not require the CODA package. For each set of plots, the number of parameters must be known (they may be counted on the BEAST ‘Parameter’ panel). For the data to be read into R, you must also include the sample index, prior, and likelihood values (the example below has 5 real parameters, but ‘Npar’ – the number of columns – is set to $5 + 3 = 8$). The other information that must be entered is the number of samples (from the ‘MCMC’ panel in BEAST).

```
Nsamp <- 50000           # number of samples – rows in the matrix
Npar <- 8                # the total number of parameters + 3
DirN <- "D:/Eric/beastOutput/" # the name of the directory
FileN <- "run01.par"     # the name of the file, typically the .par file
FileName <- paste(DirN, FileN, sep = "") # combine the directory and file names
# load the data into a matrix called ‘MCMC’ – the rows are samples
MCMC <- matrix(scan(FileName,n=Nsamp*Npar),ncol=Npar,byrow=T)
# call the density function within a plot function to produce a plot of the
# parameter in the second column
plot(density(MCMC[1:50000,2]), xlab = 'Parameter Name', main = 'Plot Title',
ylab = 'Density')
```

2) Analyzing MCMC chains with CODA

The commands for loading data are the same as above. The only difference is that an MCMC object needs to be created before the CODA package is run.

```
Nsamp <- 50000           # number of samples – rows in the matrix
Npar <- 8                # the total number of parameters + 3
DirN <- "D:/Eric/beastOutput/" # the name of the directory
FileN <- "run01.par"     # the name of the file, typically the .par file
FileName <- paste(DirN, FileN, sep = "") # combine the directory and file names
```

```
# load the data into a matrix called 'MCMC' – the rows are samples
MCMC <- matrix(scan(FileName,n=Nsamp*Npar),ncol=Npar,byrow=T)
mcmcObj <- mcmc(MCMC)          # create an mcmc object
codamenu()                    # start coda
# when prompted, type 'mcmcObj' as the requested MCMC object.
```

There are a range of convergence diagnostics and plots available that will not be discussed here.

3) DIC output

Summary DIC statistics are outputted to the *.cor file. There are four values computed:

- 1) $DIC(E[X])$, the deviance of the mean parameter vector
- 2) $E[DIC]$, the mean deviance of all parameter vectors
- 3) pD , the effective number of parameters estimated
- 4) DIC , the deviance information criterion statistic. For further details, see Spiegelhalter (2002).

Appendix of Distributions

Continuous

- 1) [Normal](#)
- 2) [Uniform](#)
- 3) [Log Uniform](#)
- 4) [Gamma \(a, b\)](#)
- 5) [Gamma \(a, b, d\)](#)
- 6) [Lognormal](#)
- 7) [Exponential](#)
- 8) [Inverse Gamma](#)
- 9) [Logistic](#)
- 10) [Beta](#)
- 11) [F](#)
- 12) [Chi-squared](#)
- 13) [Student's t](#)
- 14) [Double Exponential](#)
- 15) [Gamma \(u,b\)](#)
- 16) [Gamma \(u, b, d\)](#)
- 17) [Half-Normal](#)

Discrete

- 18) [Beta-Binomial](#)
- 19) [Uniform](#)
- 20) [Poisson](#)
- 21) [Negative Binomial](#)
- 22) [Binomial](#)
- 23) [Geometric](#)

Normal (continuous): $X \sim N(u, \sigma)$

$$f(\underline{x} | u, \sigma^2) = \left(\frac{1}{\sqrt{2\pi} \cdot \sigma} \right)^N \cdot \exp \frac{-\sum_{i=1}^N (x_i - u)^2}{2\sigma^2}$$

$-\infty \leq x \leq \infty; -\infty \leq u \leq \infty; \sigma > 0$

$$E[X] = u$$

$$\text{Var}(X) = \sigma^2$$

$$L(u, \sigma^2 | \underline{x}) = -N \cdot \ln(\sigma) - (N/2) \cdot \ln(2\pi) - \left[\frac{(N-1) \cdot S^2 + N(u - \bar{x})^2}{2\sigma^2} \right]$$

Notes: Also referred to as the Gaussian distribution.

Uniform (continuous): $X \sim \text{Uniform}(a, b)$

$$f(\underline{x} | a, b) = \left(\frac{1}{b-a} \right)^N$$

$x, a, b = \text{real number}; a < b$

$$E[X] = \left(\frac{b+a}{2} \right)$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

$$\log L(a, b | \underline{x}) = -N \cdot \ln(b-a)$$

Notes: If both parameters equal 1.0, this is a special case of the Beta distribution.

Log Uniform (continuous): $X \sim \text{Log Uniform}(a, b)$

$$f(\underline{x} | a, b) = \prod_{i=1}^n \left(\frac{1}{x_i \cdot (b-a)} \right)$$

$x, a, b = \text{real number}; a < b$

$$E[X] = \left(\frac{b+a}{2} \right)$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

$$\log L(a, b | \underline{x}) = -N \cdot \ln(b-a) + \sum_{i=1}^n \frac{1}{x_i}$$

Gamma (continuous): $X \sim \text{Gamma}(a, b)$

$$f(\underline{x} | a, b) = \left(\frac{1}{\Gamma(a) \cdot b^a} \right)^N \cdot \left(\prod_{i=1}^N x_i \right)^{(a-1)} \cdot \exp \frac{-\sum_{i=1}^N x_i}{b}$$

$$0 \leq x < \infty; a > 0; b > 0$$

$$E[X] = ab$$

$$\text{Var}(X) = ab^2$$

$$\log L(a, b | \underline{x}) = -N \cdot \ln(\Gamma(a)) - N \cdot a \cdot \ln(b) + (a-1) \cdot \sum_{i=1}^N \ln(x_i) - \frac{\sum_{i=1}^N x_i}{b}$$

Notes: Has many special cases, including Exponential, Chi-squared, Maxwell, Inverse Gamma.

3 Parameter Gamma (continuous): $X \sim \text{Gamma}(a, b, d)$

$$f(\underline{x} | a, b) = \left(\frac{1}{\Gamma(a) \cdot b^a} \right)^N \cdot \left(\prod_{i=1}^N (x_i + a \cdot b - d) \right)^{(a-1)} \cdot \exp \frac{-\sum_{i=1}^N (x_i + a \cdot b - d)}{b}$$

$$0 \leq x < \infty; a > 0; b > 0$$

$$E[X] = d$$

$$\text{Var}(X) = ab^2$$

$$\log L(a, b | \underline{x}) = -N \cdot \ln(\Gamma(a)) - N \cdot a \cdot \ln(b) + (a - 1) \cdot \sum_{i=1}^N \ln(x_i + a \cdot b - d)$$

$$\frac{\sum_{i=1}^N (x_i + a \cdot b - d)}{b}$$

Notes: Identical to gamma distribution, only difference being that values are de-meant.

Lognormal (continuous): $X \sim \text{LogN}(u, \sigma)$.

$$f(\underline{x} | u, \sigma) = \prod_{i=1}^N \left(\frac{1.0}{\sqrt{2\pi\sigma} \cdot x_i} \cdot \exp\left(\frac{-((\log(x_i) - u)^2)}{2\sigma^2}\right) \right)$$

$0 \leq x \leq \infty$; $u = \text{any real number}$, $\sigma > 0$

$$E[X] = \exp\left(u + \frac{\sigma^2}{2}\right)$$

$$\text{Var}(X) = \exp(2 \cdot u + \sigma^2) \cdot (\exp(\sigma^2) - 1)$$

$$\log L(u, \sigma | \underline{x}) \propto f(\underline{x} | u, \sigma)$$

Exponential (continuous): $X \sim \text{Exp}(b)$

$$f(\underline{x} | b) = \left(\frac{1}{b}\right)^N \cdot \exp\left(\frac{-\sum_{i=1}^N x_i}{b}\right)$$

$0 \leq x < \infty$; $b > 0$

$$E[X] = b$$

$$\text{Var}(X) = b^2$$

$$\log L(\mathbf{b} | \underline{\mathbf{x}}) = -N \cdot \ln(\mathbf{b}) - \frac{\sum_{i=1}^N x_i}{\mathbf{b}}$$

Notes: Also a special case of the gamma distribution with scale = 1. This distribution is unique in that it has a memoryless property. Special cases include Weibull, Rayleigh, Gumbel.

Inverse Gamma (continuous): $\mathbf{X} \sim \mathbf{IG}(\mathbf{a}, \mathbf{b})$

$$f(\underline{\mathbf{x}} | \mathbf{a}, \mathbf{b}) = \left(\frac{1}{\Gamma(\mathbf{a}) \cdot \mathbf{b}^{-\mathbf{a}}} \right)^N \cdot \left(\prod_{i=1}^N x_i \right)^{-(\mathbf{a}+1)} \cdot \exp^{-\frac{\mathbf{b}}{\sum_{i=1}^N x_i}}$$

$$0 \leq x < \infty; \mathbf{a} > 0; \mathbf{b} > 0$$

$$E[\mathbf{X}] = \frac{\mathbf{a}}{\mathbf{b} - 1}$$

$$\text{Var}(\mathbf{X}) = \frac{E[\mathbf{X}]^2}{\mathbf{b} - 2}$$

$$\log L(\mathbf{a}, \mathbf{b} | \underline{\mathbf{x}}) = -N \cdot \ln(\Gamma(\mathbf{a})) + N \cdot \mathbf{a} \cdot \ln(\mathbf{b}) - (\mathbf{a} + 1) \cdot \sum_{i=1}^N \ln(x_i) - \frac{\mathbf{b}}{\sum_{i=1}^N x_i}$$

Logistic (continuous): $\mathbf{X} \sim \mathbf{Logistic}(\mathbf{u}, \mathbf{b})$

$$f(\underline{\mathbf{x}} | \mathbf{u}, \mathbf{b}) = \prod_{i=1}^N \left(\frac{1}{\mathbf{b}} \cdot \frac{\exp^{-\frac{x_i - \mathbf{u}}{\mathbf{b}}}}{\left[1 + \exp^{-\frac{x_i - \mathbf{u}}{\mathbf{b}}} \right]^2} \right)$$

$$-\infty \leq x \leq \infty; \mathbf{u} = \text{any real number}, \mathbf{b} > 0$$

$$E[X] = u$$

$$\text{Var}(X) = \frac{(\pi \cdot b)^2}{3}$$

$$\log L \propto f(\underline{x} | u, b)$$

Beta (continuous): $X \sim \text{Beta}(a, b)$

$$f(\underline{x} | a, b) = \left(\frac{1}{\text{Beta}(a, b)} \right)^N \cdot \left(\prod_{i=1}^N x_i \right)^{(a-1)} \cdot \left(\prod_{i=1}^N (1-x_i) \right)^{(b-1)}$$

$$0 \leq x \leq 1; a = 1, 2, 3, \dots; b = 1, 2, 3, \dots$$

$$E[X] = \frac{a}{a+b}$$

$$\text{Var}(X) = \frac{(a+b)}{(a+b)^2(a+b+1)}$$

$$\log L(a, b | \underline{x}) = -N \cdot \ln(\text{Beta}(a, b)) + (a-1) \cdot \sum_{i=1}^N \ln(x_i) + (b-1) \cdot \sum_{i=1}^N \ln(1-x_i)$$

Notes: The beta constant can be expressed as $\text{Beta}(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$. The

Beta can be used to model rates because it is confined to the interval (0,1).

F (continuous): $X \sim F(v, w)$

$$f(\underline{x} | v, w) = \prod_{i=1}^N \left(\frac{\Gamma\left(\frac{v+w}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\Gamma\left(\frac{w}{2}\right)} \cdot \left(\frac{v}{w}\right)^{\frac{v}{2}} \cdot \frac{x_i^{\frac{v-2}{2}}}{\left(1 + \frac{v}{w} \cdot x_i\right)^{\frac{v+w}{2}}}\right)$$

$$0 \leq x < \infty; v, w = 1, 2, 3, \dots$$

$$E[X] = \frac{w}{w-2}, w > 2$$

$$\text{Var}(X) = 2 \cdot \left(\frac{w}{w-2}\right)^2 \frac{(v+w-2)}{v \cdot (w-4)}, w > 4$$

$$\log L(v, w | \underline{x}) \propto f(\underline{x} | v, w)$$

Notes: The F-distribution is simply two independent chi-squared random variables divided by their degrees of freedom.

Chi-squared (continuous): $\mathbf{X} \sim \text{Chi-squared}(p)$

$$f(\underline{x} | p) = \left(\frac{4.0}{\Gamma(2) \cdot p^2} \right)^N \cdot \left(\prod_{i=1}^N x_i \right) \cdot e^{-2 \cdot \sum_{i=1}^N x_i / p}$$

$$0 \leq x < \infty; p = 1, 2, \dots$$

$$E[\mathbf{X}] = p$$

$$\text{Var}(\mathbf{X}) = \frac{p^2}{2}$$

$$\log L(p | \underline{x}) = N \cdot \ln(4) - N \cdot \ln(\Gamma(2)) - 2 \cdot N \cdot \ln(p) + \sum_{i=1}^N \ln(x_i) - \frac{\sum_{i=1}^N x_i}{p}$$

$$= N \cdot \ln(4) - 2 \cdot N \cdot \ln(p) + \sum_{i=1}^N \ln(x_i) - \frac{\sum_{i=1}^N x_i}{p}, \text{ because } \Gamma(2) = 1$$

Notes: Chi-squared (p) is just a special case of the Gamma distribution $\sim \text{Gamma}(2, p/2)$. If $Z \sim \text{Chi-squared}(w)$ and $Y \sim \text{Chi-squared}(x)$, $Z+Y \sim \text{Chi-squared}(w+x)$.

Student's t (continuous): $\mathbf{X} \sim t(u, df, \sigma)$

$$f(\underline{x} | u, \sigma, df) = \prod_{i=1}^N \left(\frac{\Gamma((df+1)/2)}{\Gamma(df/2) \cdot \sqrt{df} \cdot \pi \cdot \sigma} \cdot \left(1 + \frac{1}{df} \left(\frac{x_i - u}{\sigma} \right)^2 \right)^{-(df+1)/2} \right)$$

$$-\infty \leq x \leq \infty; u = \text{any real number}, \sigma > 0, df = \text{integer} > 0$$

$$E[\mathbf{X}] = u$$

$$\text{Var}(X) = \frac{\text{df}}{\text{df} - 2} \cdot \sigma^2, \text{df} > 2$$

Notes: Small sample approximation to the Normal. Related to the F distribution... $F(1, v) \sim (t(v))^2$. Also the ratio of a Normal random variable to the square root of a Gamma (or Chi-Square).

$$L(u, \sigma, \text{df} | \underline{x}) \propto f(\underline{x} | u, \sigma, \text{df})$$

Double Exponential (continuous): $X \sim \text{Dexp}(a, b)$

$$f(\underline{x} | a, b) = \left(\frac{1}{2b}\right)^N \cdot \exp \frac{\sum_{i=1}^N |x_i - a|}{b}$$

$$-\infty \leq x \leq \infty; a = \text{any real number}, b > 0$$

$$E[X] = a$$

$$\text{Var}(X) = 2b^2$$

$$\log L(a, b | \underline{x}) \propto f(\underline{x} | a, b)$$

Notes: Also referred to as the Laplace distribution

Gamma(u) (continuous): $X \sim \text{Gamma}(u, b)$

$$f(\underline{x} | u, b) = \left(\frac{1}{\Gamma\left(\frac{u}{b}\right) \cdot b^{\left(\frac{u}{b}\right)}} \right)^N \cdot \left(\prod_{i=1}^N x_i \right)^{\left(\frac{u}{b} - 1\right)} \cdot \exp \frac{-\sum_{i=1}^N x_i}{b}$$

$$0 \leq x < \infty; u > 0; b > 0$$

$$E[X] = u$$

$$\text{Var}(X) = ub$$

$$\log L(u, b | \underline{x}) = -N \cdot \ln\left(\Gamma\left(\frac{u}{b}\right)\right) - N \cdot \frac{u \cdot \ln(b)}{b} + \left(\frac{u}{b} - 1\right) \cdot \sum_{i=1}^N \ln(x_i) - \frac{\sum_{i=1}^N x_i}{b}$$

3 Parameter Gamma(u) (continuous): $X \sim \text{Gamma}(u, b, d)$

$$f(\underline{x} | u, b) = \left(\frac{1}{\Gamma\left(\frac{u}{b}\right) \cdot b^{\left(\frac{u}{b}\right)}} \right)^N \cdot \left(\prod_{i=1}^N (x_i + u - d) \right)^{\left(\frac{u}{b} - 1\right)} \cdot \exp \frac{-\sum_{i=1}^N (x_i + u - d)}{b}$$

$$0 \leq x < \infty; u > 0; b > 0$$

$$E[X] = d$$

$$\text{Var}(X) = ub$$

$$\log L(u, b | \underline{x}) = -N \cdot \ln\left(\Gamma\left(\frac{u}{b}\right)\right) - N \cdot \frac{u \cdot \ln(b)}{b} + \left(\frac{u}{b} - 1\right) \cdot \sum_{i=1}^N \ln(x_i + u - d) - \frac{\sum_{i=1}^N (x_i + u - d)}{b}$$

Half-Normal (continuous): $X \sim \text{Half-Normal}(a, b)$

$$f(x | a, b) = \frac{1}{b} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{1}{2} \left(\frac{x_i - a}{b}\right)^2\right)$$

$$E[X] = a + b \sqrt{\frac{2}{\pi}}$$

$$\text{Var}(X) = b^2 \left(1 - \frac{2}{\pi}\right)$$

$$\log L(a, b | \underline{x}) = -N \cdot \ln(b) + \frac{N}{2} \cdot \ln\left(\frac{2}{\pi}\right) - \frac{1}{2} \left(\frac{x_i - a}{b}\right)^2$$

Notes: a is a location parameter, b is a scale parameter (controlling variance). If $X \sim \text{normal}(a, b)$, and the distribution is folded about the mean, the new distribution \sim half-normal.

BetaBinomial (discrete): **$X \sim \text{BetaBinomial}(a, b, N)$**

$$f(\underline{x} | a, b, N) = \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)\Gamma(a+b+N)} \right)^N \prod_{i=1}^N \left[\binom{N}{x_i} \cdot \Gamma(a+x_i)\Gamma(b+N-x_i) \right]$$

$a > 0$; $b > 0$; $N = \text{positive integer}$

$$E[X] = N \cdot \left(\frac{a}{a+b} \right)$$

$$\text{Var}(X) = \frac{N \cdot a \cdot b}{(a+b)^2} \cdot \frac{(a+b+n)}{(a+b+1)}$$

$$\log L(a, b, N | \underline{x}) \propto f(\underline{x} | a, b, N)$$

Uniform (discrete): **$X \sim \text{Uniform}(a, b)$**

$$f(\underline{x} | a, b) = \left(\frac{1}{b-a} \right)^N$$

$x, a, b = \text{integer}$; $a < b$

$$E[X] = \left(\frac{b+a}{2} \right)$$

$$\text{Var}(X) = \frac{(b-a+1)^2}{12}$$

$$\log L(a, b | \underline{x}) = -N \cdot \ln(b-a)$$

Poisson (discrete): **$X \sim \text{Poisson}(m)$**

$$f(\underline{x} | m) = \frac{\exp^{-N \cdot m} \cdot m^{\sum_{i=1}^N x_i}}{\prod_{i=1}^N (x_i!)}$$

$x = 0, 1, 2, \dots$; $m > 0$

$$E[X] = m$$

$$\text{Var}(X) = m$$

$$\log L(m | \underline{x}) = -N \cdot m + \ln(m) \cdot \left(\sum_{i=1}^N x_i \right) - \sum_{i=1}^N \ln(x_i!)$$

Notes: Commonly used to model count data. If $X \sim \text{Poi}(m)$ and $Y \sim \text{Poi}(n)$, and X and Y are independent, $X+Y \sim \text{Poi}(m+n)$.

Negative Binomial (discrete): $X \sim \text{NegBin}(r, p)$

$$f(\underline{x} | r, p) = \prod_{i=1}^N \left[\binom{r + x_i - 1}{x_i} \cdot p^r \cdot (1-p)^{x_i} \right]$$

$$x = 0, 1, 2, \dots; r = 1, 2, 3, \dots; 0 \leq p \leq 1$$

$$E[X] = \frac{r(1-p)}{p}$$

$$\text{Var}(X) = \frac{r(1-p)}{p^2}$$

$$\log L(r, p | \underline{x}) = \sum_{i=1}^N \ln((r + x_i - 1)!) - N \cdot \ln((r - 1)!)$$

$$- \sum_{i=1}^N \ln(x_i!) + r \cdot N \cdot \log(p) + \log(1-p) \cdot \sum_{i=1}^N x_i$$

Other parameterizations:

$$f(\underline{x}) = \prod_{i=1}^N \left[\binom{r + x_i - 1}{r - 1} \cdot p^r \cdot (1-p)^{x_i} \right]$$

$$E[X] = rp$$

$$\text{Var}(X) = \frac{r(1-p)}{p^2}$$

Notes: If $r = 1$, the distribution is geometric. The mode of the distribution also depends on r . Using the first parameterization, if $r(1-p) < 1$, the mode is at zero. If $r(1-p) = 1$, there are two modes at 0 and 1. If $r(1-p) > 1$, the mode is the smallest integer not less than $r(1-p)/p$. A useful result is that the sum of k independent random negative binomial variables with parameters (r_i, p) is also negative binomial with parameters $(r_1+r_2+\dots+r_k, p)$.

Binomial (discrete): $\mathbf{X} \sim \text{Binomial}(\mathbf{N}, p)$

$$f(\underline{x} | n, p) = \left(\prod_{i=1}^N \binom{n}{x_i} \cdot p^{\sum_{i=1}^N x_i} \cdot (1-p)^{n - \sum_{i=1}^N x_i} \right)$$

$$x = 0, 1, 2, \dots; n = 1, 2, 3, \dots; 0 < p < 1$$

$$E[\mathbf{X}] = np$$

$$\text{Var}(\mathbf{X}) = np(1-p)$$

Notes: The Binomial can be seen as the sum of independent Bernoulli trials. The multivariate case of the binomial is multinomial. The mode of the distribution is at the greatest integer not greater than $(n+1)p$ – if this value is an integer, then it is a mode along with $(n+1)p - 1$. A useful result is that the sum of k independent Binomial random numbers with parameters (n_i, p) $i = 1, \dots, k$ is also binomial with parameters $(n_1+n_2+\dots+n_k, p)$.

$$\log L(n, p | \underline{x}) = N \cdot \ln(n!) - \sum_{i=1}^N \ln[(n - x_i)!] - \sum_{i=1}^N \ln(x_i!) + \ln(p) \cdot \sum_{i=1}^N x_i + \ln(1-p) \cdot \left(N \cdot n - \sum_{i=1}^N x_i \right)$$

Geometric (discrete): $\mathbf{X} \sim \text{Geo}(p)$

$$f(\underline{x} | p) = p^N \cdot (1-p)^{\sum_{i=1}^N x_i - N}$$

$$x = 1, 2, 3, \dots; 0 \leq p \leq 1$$

$$E[\mathbf{X}] = \frac{1}{p}$$

$$\text{Var}(\mathbf{X}) = \frac{1-p}{p^2}$$

$$\log L(p | \underline{x}) = N \cdot \ln(p) + \left(\sum_{i=1}^N x_i - N \right) \cdot \ln(1-p)$$

Notes: The geometric distribution is used to model the number of failures until the first success (this definitions can be rearranged). Multiple iid geometric random variables are distributed \sim Negative Binomial.

Example 1

L.W. Krefting and C.E. Ahlgren conducted a study on deer mice in Lake County, Minnesota (Ecology, 55:1391-1398). The data from their study (abundance in log-space) are

1955	45
1956	10
1957	14
1958	11
1959	9
1960	2
1961	15
1962	6
1963	6
1964	1
1965	3.5
1966	1
1967	2

Using BEAST, we can answer several questions, including:

- 1) What is the estimate rate of decline
- 2) What is the distribution of projected abundance estimates in 1972?

Approach #1:

An exponential population model with no change point, lognormal likelihood, and observation error was chosen. The data above were pasted into BEAST, with no additional harvest or scale data entered. Uniform priors were placed on $\ln(N_0)$ and r , and a log-uniform prior was placed on σ . A burn-in of 1,000,000 was used, and 5000 samples were collected with the default value of thinning ($= 100$). The distribution of the population decline and 1972 population size is:

