BORMICON

User's manual

Draft

Halldór Narfi Stefánsson
Hersir Sigurgeirsson
Höskuldur Björnsson

Marine Research Institute

August 7, 1997

Contents

Chapter 1

Invoking and using Bormicon.

1.1 Invoking Bormicon.

The bormicon program can be invoked with several options. It uses the file called 'main' in the directory given with the environmental variable BORMICON_DATA_DIR.

Typical command to set the environmental variable is

$$setenv\ BORMICON\_DATA\_DIR\ .$$

Which assumes that main is in the current directory.

Currently two minimizing alorithms are implemented in the program.

- Hooke and Jeeves
- Simulated annealing.

The options allowed are:

No options The program reads information from the files and performs a simulation. Likelihood function is not calculated.

-s A stochastic run is performed. In a stochastic run the values of the some switches are usually given in a file following the -i option. If the -i option is not specified one simulation is made using the values in the datafiles. (Calculating the likelihood function)

-l Maximum likelihood run is performed.

-i filename Name of inputfile for stochastic runs or file with initial values for optimizing runs.

-o filename Name of file to write values of variables and likelihood components during optimization or stochastic runs. If an optimizing run is performed and the -o option not specified an error is issued and the program stops.

-co filenname Name of file to write values of variables and likelihood components during optimization or stochastic runs. The format of the file is different from the format of the file written by the -o option. The output is more column oriented (-co stands for column output).

-printinitial filename Print initial conditions to a file. The file will be very large. Only used as a debugging tool.

-printfinal filename Print final conditions to a file. The file will be very large (same format as initial file). Only used as a debugging tool.

-opt filename Name of optimizing information file. If no file is given and optimizing run is performed Hooke an Jeeves algorithm with default values of parameters is used.

-print1 How often to print using the -o option. Default value is one which means printing in every simulation. (if the -o option is specified)

-print2 How often to print using the -co option. Default value is one which means printing in every simulation. (if the -co option is specified)

-m filename A name of file with all the other options.

1.2 File-format for stochastic runs or initial values for optimization.

A stochastic run is simply repeated simulations with different values of the marked variables each time. The value of the likelihood function is calculated in the end of it and the output is written to a file, entirely similar to what is done in a maximum likelihood run.

Input data files for stochastic runs and initial value files for optimization have the same file format. The only difference is that when used as initial values for optimization only the first set of data given is used while stochastic runs can use multiple datasets.

First line. 3 possibilities

- Numbers with space or tab in between. These numbers must be exactly as many as the number of different switches read by the program and the order must be the same as the order in which the program reads the variables. A good way to see that order is to look at the output from bormicon for.example by typing **borm -s stoch.out**. In stochastic the number of lines in the file will determine the number of runs while only the first line is used in optimization. Comments are allowed in the file.

- The first line contains the text **switches**. This tells the program that the next line is to be understood as a list of switches. The remaining lines then contain data. The numbers given in each line must be exactly as many as the number of switches given in the second line. Here the program matches the switches in the datafile against the switches in the datafiles and sets the corresponding values. If a switch is found in the datafiles but not in the stochastic data files the data in the datafiles is used. If a switch is found in the stochastic data file but not in the data files a warning is given and the switch and corresponding number ignored.

The first line contains some or all of the of the words

```
switches
optimize
values
upperbound
lowerbound
```

The number of columns in the file must then be exactly as many as the number of words in the first line. The words given in the first line control What is in each column. Numbers in each line after the first one must be exactly as many as the words in the first line.

switches The switches to be matched against the values in datafiles

optimize Should the variable be changed during optimization. If the value is 1 the variable can be changed by the optimizing program, if the value is zero it is not changed.

values Initial values for optimization or values for stochastic runs.

lowerbound Lower bound for the variable. If the value of a variable is less than the lower bound a penalty can be added to the likelihood function. Lower bounds can also be given through optimizing information files.

upperbound Upper bound for the variable. If the value of a variable is less than the lower bound a penalty can be added to the likelihood function. Upper bounds can also be given through optimizing information files.

If switches are not given the number of lines with data must be exactly as many as the number of different switches in the datafiles and appear in the same order. Else the switches are matched against the switches in the datafiles. If a given switch dosent contain match in any datafile a warning is given and the corresponding numbers ignored. If a switch is given in a datafile but not in the stochastic datafile the value in the datafile is given. If optimize is not in the first line all variables are allowed to change during optimization else only the values where optimize is 1. This can be convenient if the user wants to fix some variables that have been allowed to change during optimization. If either of the first 2 types of files is used all variables must be allowed to change during optimization.

This fileformat can only be used to give initial values for optimization or **one** stochastic run.

Examples of files. Taken from an example where there are 10 different switches in the datafiles. The switches are read in the order 100-101-1100-98-97-300-301-302-303-304.

First type of stochastic files

```
0.5 0.76 0.98 0.89 -0.1 0.234 0.568 0.98 0.67 0.77
0.55 0.86 0.88 0.89 -0.01 0.234 0.568 0.98 0.68 0.77
0.45 0.7 0.91 0.8 -0.08 0.234 0.57  0.99 0.69 0.75
```

Second type of stochastic files. 3 switches varied during optimization.

```
switches
300 303 97
0.234 0.67 -0.1
0.234 0.68 -0.01
0.234 0.69 -0.08
```

Third type of stochastic files. Bounds given. Switches not given.

```
values opt lowerbound upperbound
0.5 0 0.25  0.75
0.76  0 0.5 1
0.98 1 0.89 1
0.89 1 0.8 1
-0.1 1 -0.2 -0.05
0.234 0 0.2 0.25
0.568 1 0.5 0.6
0.98 0 0.9 1
0.67 1 0.6 0.7
0.77 1 0.75 0.78
```

Here variables 1,2,6 and 8 will not be changed by the optimizing program (if optimizing) but their values set to 0.5, 0.76, 0.234 and 0.98. The value does not nessecarily have to be between upper and lower bound. A possibility to add penalty to variables outside the bounds will be added to the likelihood function. The bounds can also be added through the optimizing information file which is convenient if bounds are for example only set on two variables. Another example of the third type of files is.

```
switches values opt lowerbound upperbound
100 0.5 0 0.25  0.75
98 0.89 1 0.8 1
302 0.98 0 0.9 1
```

Here only one variable will change during optimization the variable with switch 98. The value of opt does not matter in a stochastic run.

Variable that is now set in the file prey.cc but will later be read is the maximum mortality in one month. Earlier all fish in an area could be caught but this was changed to 95% (Z = 3)

1.3 Marking variables.

Variables in the input files can be marked for later reference. Take as an example the natural mortality of an agegroup of a sub-stock. Then the file format is:

```
naturalmortality     0.2#1
```

Now the variable is marked with the number 1. The marks have to be ≥ 0 and all the variables with the same mark have to be of equal value. The number of the marks do **not** have to be sequential starting with #1. The marks are used in stochastic runs and optimizations In both cases the program performs multiple runs with (possibly) different values of the marked variables, not distinguishing between variables with the same mark.

Recently the syntax for marked variables has been extended so now the format is.

```
183*1#201*2#202
```

Here the value of the variable is 183 times marked variable 201 times marked variable 202. Initial values for marked variables 201 and 202 are 1 and 2 except listed otherwise in initial value file or input file for stochastic runs.

This syntax is now implemented in initial conditions and recruitment. By using this syntax simple main effect model can be used. If marked variable #*yy* describes the size of the *yy* yearclass and marked variable #*aa* the factor for area *aa* the the number recruited could be for the years 87 and 88 and areas 1-3.

```
202*1#87
202*1#87*1#2
202*1#87*1#3
191*1#88
191*1#88*1#2
191*1#88*1#3
```

1.4 The method of maximum likelihood.

When a maximum likelihood run is to be performed the program has to read at least one marked variable in the input files. Then the program runs repeated simulations with different values for the marked variables, using the same values for all the variables with the same mark, in order to find the maximum value of the likelihood function.

The names of the marked variables, their values in the repeated simulations and the value of the likelihood function are written to an output file.

1.5 Stochastic run.

A stochastic run is simply repeated simulations with different values of the marked variables each time. The value of the likelihood function is calculated in the end of it and the output is written to a file, entirely similar to what is done in a maximum likelihood run.

1.5.1 File format of the stochastic run input file.

The input file for the stochastic run contains a matrix whose number of columns equals the number of marked variables read from the input files. Element $x_{i,j}$ of that matrix is used as the value of the j-th marked variable in the i-th simulation.

Warning: Note that $x_{i,j}$ is used as the value of the **j-th marked variable**, i.e. the j-th mark read from the input files. Since it is not obvious in what order the input files are read, you should check the output file of the run to see if the the marks were not surely read in the order expected.

1.6 Error messages.

When an error occurs in reading from the input files the program should give information about in what file the error occured, as well as some information on where the error occurred. A typical error message consists of the name of the file where the error occured and some details on the nature of the error, followed with a hint on where it occurred.

```
Warning in file cod.mat:
codmatnaturalm12.
Error occurred in ReadIndexVectorWithSwitches when element
 no. 8 out of 8 was read.
Error in file cod.mat:
Was reading naturalm
```

This means that the failure occurred in the file cod.mat when reading natural mortality for the 12 years old of the stock codmat.

Another simple example is:

```
Error in file cod.mat.migration:
Expected coefficients but found instead a1
```

Here the error occurred when reading the ’coefficients’ section of the migration data for cod.mat.

But sometimes the errors messages may be different from what you expected, but nevertheless lead you to the cause. The extra ’a’ in the last line of:

```
minlength       39.5
maxlength       135.5
dl              0a.5
```

causes the error message

```
Failed to create length group division with
minimum length 39.5
maximum length 135.5
and dl 0
Kept string for explanations is codmat
```

to be emitted. A quick inspection of the input file should lead the user to the cause.

1.7 Bugs.

The bugs are probably very many. Rule number 1 is to be critical when the output is of some surprise.

Known bugs are:

- Errors in the output files.

When writing the standard stock output, it is possible that on the time step where the age update occurs, the predation data for an age group is not printed in the same line as its abundance data. I.e. the predation for age group 4 may be printed in the line for age group 3.

- Unnecessary error messages when reading fleet data.
  When reading data on fleet, a error message may be emitted saying that unexpected end of file was encountered. This may happen when the last entry in the file is for the last time step in the simulation.

  Solution: Add an entry in the file for the time step after the last one in the simulation. Then it will not be used in the simulation but the error message will disappear.

1.8 Not impelemented.

The following is not fully implemented:

- Age dependent migration.
- Stocks with no length group division.
- Catch.

Chapter 2

Main input files.

2.1 What does the $^{\#}$ mean?

In the description of the input files, $g^{\#}$ means that the variable g may be marked – see section 1.3 on marking variables.

2.2 Comments in the input files.

Comments are allowed in the input files. If a semicolumn, ';', is encountered in a line, the rest of it is ignored.

2.3 Main file.

The names of the files that keep the information on the time of the simulation and the areas in the simulation are in the main file, which must bear the name `main`. This information will be used when reading the subsequent data.

All filenames are considered to be relative to the directory in which the main file resides.

timefile <name of time file> *(sec. 2.4)*
areafile <name of area file> *(sec. 2.5)*

catch str_1 <name of catch file 1> *(sec. 3.1)*
$\vdots$
str_l <name of catch file l>

stockfiles	$sname_1$ <name of stock file 1>	*(sec. 4)*
	⋮	
	$sname_m$ <name of stock file m>	
otherfoodfiles	$oname_1$ <name of other food file 1>	*(sec. 5.1)*
	⋮	
	$oname_n$ <name of other food file n>	
totalfleet	$totname_1$ <name of totalfleet file 1>	*(sec. 5.2)*
	⋮	
	$totname_o$ <name of totalfleet file o>	
linearfleet	$linname_1$ <name of linear fleet file 1>	*(sec. 5.2)*
	⋮	
	$linname_p$ <name of linear fleet file p>	
printing	<name of printfile>	*(sec. 6)*
likelihood	<name of likelihood file 1> w_1	*(sec. 7)*
	⋮	
	<name of likelihood file q> w_q	

Here $sname_i$ are names of stocks and the information for the corresponding stocks are found in <name of stock file i> (corresponding for the otherfood, totalfleet and linearfleet); $str_1 \ldots str_l$ are text strings. It is expected that each of them is the beginning of the name of at least one stock.

In the file format for the likelihood, w_i is a multiplication constant for the likelihood value. If there is a total of q likelihood files, w_i is the weight for likelihood component number i and ℓ_i is the calculated likelihood value for component number i, then the total likelihood value is

$$\ell_{total} = \sum_{i=1}^{q} w_i \ell_i.$$

2.4 Time.

The format of the time file is:

firstyear	y_f	
firststep	s_f	
lastyear	y_l	
laststep	s_l	
notimesteps	n	$s_1 \ldots s_n$

This means that the simulation is to start on the year y_f, step s_f and the last step in the simulation is s_l in the year y_l.

On each year there are n steps and the the length of the i-th step is s_i. The sum $s_1 + \ldots + s_n$ is expected to equal 12.

2.5 Area.

The format of the area file is:

$$\begin{array}{llll}
\text{areas} & a_1 & \ldots & a_n \\
\text{size} & s_1 & \ldots & s_n \\
\text{temperature} & & &
\end{array}$$

$$\begin{array}{llllll}
y_1 & s_1 & T_{1,1} & \ldots & T_{1,n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
y_r & s_r & T_{r,1} & \ldots & T_{r,n}
\end{array}$$

The numbers of the areas are first, $a_1 \ldots a_n$ and next their size. Then the temperature. In each line is the year y_i, the step s_i and then the temperatures $T_{i,1} \ldots T_{i,n}$ on the corresponding areas $a_1 \ldots a_n$.

It is assumed that the temperature data is ordered in time. The data in the file must cover the time period in question.

Chapter 3

Catch input files.

3.1 Catch.

The catch is **not fully implemented** in the program, but the file format is nevertheless described here below for completeness only.

minage	A_1	
maxage	A_2	
minlength	B_1	
maxlength	B_2	
dl	dl	
numberofstocks	no	
areas	$a_1 \ldots a_n$	
filenames	<catch data file 1>	*(sec. 3.2)*
	$\vdots$	
	<catch data file n>	

The numbers A_1, A_2,B_1, B_2 and dl describe the data in the files <catch data file 1>, . . . , <catch data file n>. As expected, A_1 is the minimum age, A_2 is the maximum age, and B_1, B_2 and dl describe the length groups of the data.

The catch in the files is aggregated into length groups, each of length dl, the minimum length in the first one is B_1 and the maximum length in the last one is B_2. This means of course that dl has to divide the difference $B_2 - B_1$.

It is quite possible that the same catch data read should be used for many stocks; say for female and male part of a stock. Therefore the parameter numberofstocks is needed to know between how many stocks the read catch should be **equally** divided.

3.2 Catch data.

The catch data itself, in the files <catch data file 1>, ... <catch data file n> from above, has to be on this format:

Year y_1 step s_1
<Matrix with $A_2 - A_1 + 1$ lines and $(B_2 - B_1)/dl$ columns.>
$\vdots$
Year y_r step s_r
<Matrix with $A_2 - A_1 + 1$ lines and $(B_2 - B_1)/dl$ columns.>

The matrices must be in correct order in time, but it is not necessary that they are given for every timestep of the period in question.

Chapter 4

Stock input files

The format of the stock file is:

livesonareas	$a_1 \dots a_n$	
minage	A_1	
maxage	A_2	
minlength	B_1	
maxlength	B_2	
dl	dl	
lowerabslength	$l_{A_1} \dots l_{A_2}$	
upperabslength	$L_{A_1} \dots L_{A_2}$	
growthandeatlengths	$gl_1 \dots gl_m$	
doesgrow	1	
	<file format for growth>	*(sec. 4.1)*
naturalmortality	${M_{A_1}}^{\#} \dots {M_{A_2}}^{\#}$	
iscaught	1	
iseaten	1	
	<file format for stockprey>	*(sec. 4.2)*
doeseat	1	
	<file format for stockpredator>	*(sec. 4.3)*
initialconditions	<file format for initial conditions>	*(sec. 4.5)*

doesmigrate	1	
	AgeDependentMigration 0	
	migrationfile <name of migration file>	*(sec. 4.6)*
doesmature	1	
	maturitytype m	
	<file format for maturity>	*(sec. 4.7)*
doesmove	1	
	<file format for movement>	*(sec. 4.8)*
doesrenew	1	
	renewaldatafile <name of renewal file>	*(sec. 4.9)*
doesspawn	1	
	spawnfile <name of spawning file>	*(sec. 4.10)*

This means that the stock occupies the areas a_1, ... ,a_n, the youngest agegroup is of age A_1, and the oldest A_2. Its population's length is between B_1 and B_2 and the length intervals are of length dl. Note that dl has to divide the difference B_2 - B_1.

In order to allow further reduction of the age-length keys, the parameters l_a and L_a define the minimum and maximum allowed length for the agegroup a. Note that dl has to divide the difference l_a - B_1 and L_a - l_a.

Growth and consumption are possibly calculated on a coarser scale than the population itself. The endpoints of the length groups used for that are given with $gl_1, \ldots, gl_m$ and each of these length groups has to be a union of one or more of the length groups given with B_1, B_2 and dl. The growthandeatlengths must be supplied even if the stock neither grows nor predates.

Different values of m, the maturity type, are possible and the file format of the maturity file may be different for different values of m.

In the example given above, all the parameters doesgrow, iscaught, iseaten doeseat, doesmigrate, doesmature, doesmove, doesrenew and doesspawn are set to 1. Of course this needs not be so, and if a stock should not have one of these properties, the 1 should be replaced with a 0 and the corresponding further information on it be omitted. (E.g. exclude the <file format for movement> if doesmove is set to 0).

4.1 Growth.

The file format for growth is:

growthfunctionnumber	gf	
	<file format for growth function gf>	*(sec. 4.1.1-4.1.5)*
growthimplementfile	<name of growthimplement file>	*(sec. 4.1.7)*

The number *gf* is the number of the growth function that should be used (sections 4.1.1-4.1.5).

4.1.1 Growth function 1.

growthparameters ${p_0}^\#, \ldots, {p_3}^\#, {q_0}^\#, \ldots, {q_4}^\#$

4.1.2 Growth function 2.

Then the growth is read from an external file. When reading growth information for the areas $a_1, \ldots, a_n$, the file format is:

growthfiles <length growth file 1> <weight growth file 1>
$\vdots$
<length growth file n> <weight growth file n>

Then <length growth file i> contains information on the length increase on the area a_i and <weight growth file i> on the mean weight change on area a_i, where n is the number of areas.

The length and weight growth files have the same format:

$y_1 \quad s_1 \quad g_{1,1} \cdots g_{1,n}$
$\vdots$
$y_m \quad s_m \quad g_{m,1} \cdots g_{m,n}$

where n is the number of length groups, y_i is year, s_i is step and $g_{i,j}$ is either length increase or mean weight change of the j-th length group on the year y_i and step s_i.

4.1.3 Growth function 3.

Wgrowthparameters	${q_0}^\# \ldots {q_5}^\#$	
Lgrowthparameters	${p_0}^\# \ldots {p_8}^\#$	
refweightfile	<name of reference weight file>	*(sec. 4.1.6)*

4.1.4 Growth function 4.

Wgrowthparameters	${q_0}^\# \ldots {q_5}^\#$	
Lgrowthparameters	${p_0}^\# \ldots {p_7}^\#$	
refweightfile	<name of reference weight file>	*(sec. 4.1.6)*

4.1.5 Growth function 5.

Wgrowthparameters	${q_0}^{\#} \dots {q_5}^{\#}$	
Lgrowthparameters	${p_0}^{\#} \dots {p_8}^{\#}$	
YearEffect	${y_1}^{\#} \dots {y_{noyears}}^{\#}$	
StepEffect	${s_1}^{\#} \dots {s_{nosteps}}^{\#}$	
AreaEffect	${a_1}^{\#} \dots {a_{noareas}}^{\#}$	
refweightfile	<name of reference weight file>	*(sec. 4.1.6)*

Here *noyears* is the number of years in the simulation, *nosteps* is the number of timesteps in year (as given in the time file, see section 2.4), and *noareas* is the number of areas (as given in the area file, see section 2.5).

4.1.6 Reference weight.

The format of the reference weight file is:

$$\begin{array}{ll} l_1 & w_1 \\ \vdots & \vdots \\ l_n & w_n \end{array}$$

where $l_1, \dots, l_n$ are lengths and $w_1, \dots, w_n$ are weights to be used as reference points in a length-weight relationship, $l_1 < \dots < l_n$.

4.1.7 Growthimplement.

This file describes the connection between the growth routines and the stock.

resolution	*res*
maxmeangrowth	*mmg*
maxlengthgroupgrowth	*mlgg*
power	p
distribution	<matrix with *mlgg* columns and *mmg*/*res* lines.>

Here, *mmg* is the maximum growth expected and *res* is the resolution of in the calculations and p is the power in the length-weight relationship.

In line i in the distribution matrix, is the growth distribution when the mean growth is $i * res$. I.e. element no. (i, j) in that matrix is the proportion of a population that has the mean growth $i * res$, that is to move up j length groups . Then, obviously, *mlgg*, is the maximum growth, counted in number of length groups.

4.2 StockPrey.

The file format for stockprey is:

preylengths $l_1 \ldots l_n$

where $l_1, \ldots, l_n$ are the endpoints of the length groups to be used when acting as a prey.

4.3 StockPredator.

The file format for stockpredator is:

<file format for suitability> *(sec. 4.4)*
maxconsumption $m_0{}^{\#}, \ldots, m_3{}^{\#}$
halffeedingvalue $hf^{\#}$

Of course *hf* is the feeding level half value and $m_0, \ldots, m_3$ are the constants for the maximum consumption.

4.4 Suitability.

The file format for suitability is:

suitability

sp_1 <suitability for sp_1>
$\vdots$ $\vdots$
sp_n <suitability for sp_n>

where <suitability for sp_j> is either:

<file name> $mult^{\#}$

or:

fn_j $p_0{}^{\#}, \ldots, p_{l_j}{}^{\#}$

In both cases, sp_j is the name of a prey, followed by a format describing the predator's suitability for it.

If the file name is given, it should contain a matrix with the number of rows equal to the number of length groups in the predator and the number of columns equal to the number of length groups in the prey. The matrix is multiplied with *mult* to get the suitability of the predator for the prey. *mult* has to be > 0.

In short, element (i, j) of that matrix, multiplied with *mult*, describes the suitability of the length group i in the predator for the length group j of the prey.

If the latter format is used, fn_j is the number of the suitability function that is to be used and $p_1, \ldots, p_{l_j}$ are the constants that are needed to evaluate suitability function fn_j.

Note that the number of constants, $l_j + 1$, may depend on j, i.e. the function used.

4.5 Initial conditions.

When reading initial conditions for the areas $a_1, \ldots, a_n$ the file format is:

numbers

minage	A_1
maxage	A_2
minlength	B_1
maxlength	B_2
dl	dl
agemultiple	$M_{A_1}{}^{\#} \ldots M_{A_2}{}^{\#}$
distribution	$m_{1A_1}{}^{\#} \ldots m_{1A_2}{}^{\#}$
	$\vdots$
	$m_{nA_1}{}^{\#} \ldots m_{nA_2}{}^{\#}$

<**optional** file format for mean lengths> *(sec. 4.5.1)*

files	<number filename 1>	<weight filename 1>
	$\vdots$	$\vdots$
	<number filename n>	<weight filename n>

Then the files <number filename i> and <weight filename i> should contain numbers and mean weights for area a_i. The numbers and mean weights are contained in matrices, with $A_2 - A_1 + 1$ columns and $(B_2 - B_1)/dl$ lines. Element in column $A_k - A_1 + 1$ and row j in the matrix in the file <number filename i> is multiplied by $M_{A_k} m_{jA_k}$ to get the initial numbers. When mean lengths and standard deviations are specified the number file is not used but must be specified. (Observe that number file is used when meanlengthbasisfunc is specified). What is in it then does not matter except **it should not contain zeros**. Good way is to specify the weightfile twice.

4.5.1 File format for mean lengths.

The optional file format for mean lengths is either

meanbasisfunc *ncol*
;parameters
$p_{11}{}^{\#} \ldots p_{1,ncol}{}^{\#}$
$\vdots$
$p_{n1}{}^{\#} \ldots p_{n,ncol}{}^{\#}$
;basis
$b_{A_1 1} \ldots b_{A_1,ncol}$
$\vdots$
$b_{A_2 1} \ldots b_{A_2,ncol}$

or

meanlenghts $l_{1A_1}{}^{\#} \ldots l_{1A_2}{}^{\#}$
$\vdots$
$l_{nA_1}{}^{\#} \ldots l_{nA_2}{}^{\#}$

and

sdevbasisfunc *ncol*
;parameters
$p_{11}{}^{\#} \ldots p_{1,ncol}{}^{\#}$
$\vdots$
$p_{n1}{}^{\#} \ldots p_{n,ncol}{}^{\#}$
;basis
$b_{A_1 1} \ldots b_{A_1,ncol}$
$\vdots$
$b_{A_2 1} \ldots b_{A_2,ncol}$

or

sdev sdevmultiplier
$s_{1A_1}{}^{\#} \ldots s_{1A_2}{}^{\#}$
$\vdots$
$s_{nA_1}{}^{\#} \ldots s_{nA_2}{}^{\#}$

or

lengthmultiplebasisfunc	option
	ncol
;parameters	
	$p_{11}{}^{\#} \ldots p_{1,ncol}{}^{\#}$
	$\vdots$
	$p_{n1}{}^{\#} \ldots p_{n,ncol}{}^{\#}$
;basis	
	$b_{B_1 1} \ldots b_{B_1,ncol}$
	$\vdots$
	$b_{B_2 1} \ldots b_{B_2,ncol}$

if *meanlengths* and *sdev* is specified two matrices with mean length and standard deviation for each agegroup and area are read in. These are then used to generate normal distribution which is scaled so the total number in each agegroup and area is 10,000. All standard deviations in the matrix are multiplied by sdevmultiplier.

In additions to specifying mean lengths and standard deviations directly they can be generated by multiplying(matrix-multiplication) together two matrices, parameter matrix and basisfunction matrix. The parameter matrix has one parameter for each area and basisfunction while the basisfunction matrix has one row for each agegroup and one column for each basisfunction. As an example if linear relationship between age and length and meanlength is assumed the basis functions are (ages 3 to 10)

1	3
1	4
1	5
1	7
1	8
1	9
1	10

Good basis functions can for example be generated by the ns command in Splus.

lengthmultiplebasisfunc is specified with stocks where age reading is not possible getting the initial length distribution can be a problem. In this case the basisfunctionmatrix and parametermatrix are multiplied (matrix-multiplication) together to get a matrix with number of columns equal to the number of areas and the number of rows equal to the number of lengthgroups. If that matrix is called M each coefficient in the numberfiles is divided by $\frac{1}{1+e^{-M_{la}}}$. if option is 1 or multiplied by M_{la}^2 if option is 2. The number in the number files can be looked at as result of surveys. To get the population length distribution the survey numbers must be divided by the survey's selection curve.

4.6 Migration.

When reading a migration file, it has to be known whether the migration is age dependent or not.

Assume the file containes information on the migration between the areas $a_1, \ldots, a_n$.

When the migration **is not** age dependent the file format is:

matrixnumbers

$$\begin{array}{lll} y_1 & s_1 & m_1 \\ \vdots & \vdots & \vdots \\ y_k & s_k & m_k \end{array}$$

variables

$$\begin{array}{ll} nv_1 & value_1{}^{\#} \\ \vdots & \vdots \\ nv_l & value_l{}^{\#} \end{array}$$

coefficients

<file format for varinfo> *(sec. 4.6.1)*

$\vdots$

<file format for varinfo>

migrationmatrices

n_1 <Migration matrix number n_1> *(sec. 4.6.2)*

$\vdots$ $\vdots$

n_l <Migration matrix number n_l>

Each number m_i is the number of the migration matrix that is to be used on step s_i on year y_i. The m_i's (and the corresponding y_i's and s_i's) are assumed to be in correct order in time. All the numbers m_i that are read for the period in question have to be found when reading the migration matrices. The corresponding matrices are read and kept.

The nv_j is an integer number, ≥ 0, followed by $value_j$ which is taken to be the value of the "migration variable number nv_j". This value is used when interpreting what is read from the "coefficients" section. The numbers $nv_1, \ldots, nv_l$ have to be distinct.

When the migration **is** age dependent, the format is:

ages

$$\begin{array}{lll} A_{1,1} & \ldots & A_{1,q_1} \\ \vdots & & \\ A_{p,1} & \ldots & A_{p,q_p} \end{array}$$

matrixnumbers

$$\begin{array}{ccccc} y_1 & s_1 & m_{1,1} & \ldots & m_{1,p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_k & s_k & m_{k,1} & \ldots & m_{1,p} \end{array}$$

variables

$$\begin{array}{ll} nv_1 & value_1{}^{\#} \\ \vdots & \vdots \\ nv_l & value_l{}^{\#} \end{array}$$

coefficients

<file format for varinfo> *(sec. 4.6.1)*

$\vdots$

<file format for varinfo>

migrationmatrices

n_1 <Migration matrix number n_1> *(sec. 4.6.2)*

$\vdots$ $\vdots$

n_l <Migration matrix number n_l>

Here, one number of migration matrix $m_{i,j}$ is read for each row in the ages matrix. I.e. all the yearclasses $A_{i,1} \ldots A_{i,q_i}$ use the same migration matrices, and they are found in the $i+2$–th column in the matrixnumbers matrix. Note that the lines in the ages matrix do not have to be of the same length.

4.6.1 Varinfo.

The file format for varinfo is for describing additions to elements of the migration matrices. This allows optimizing the migration matrices. The file format is:

$N \quad a_t \quad a_f \quad i_1\ c_1{}^{\#} \ldots i_n\ c_n{}^{\#}$

Let $j_1, \ldots, j_p$ be the numbers of the migration variables and $v_{j_1}, \ldots, v_{j_p}$ be their values, respectively. Then the line here above means that

$$\sum_{k=1}^{n} c_k * v_{i_k}$$

will be added to the element corresponding to migration from area a_f to area a_t of migration matrix number N. Of course all of the numbers $i_1, \ldots, i_n$ have to be contained in the list of all numbers of migration variables, $j_1, \ldots, j_p$.

4.6.2 Migrationmatrices.

The format of the migration matrices is:

$$\begin{matrix} p_{1,1} & \cdots & p_{1,n} \\ \vdots & \ddots & \vdots \\ p_{n,1} & \cdots & p_{n,n} \end{matrix}$$

for a migration matrix between n areas. Then $p_{i,j}$ is the proportion that moves from the j-th area to the i-th area, i.e. from area a_j to a_i.

4.6.3 Constraints.

There are two constraints on the migration matrices, the migration variables and the coefficients.

- The sum of each column of a migration matrix must equal 1.
- The sum of each column of a migration matrix, to which the linear combination of the migration variables have been added, must equal 1.

4.7 Maturity.

The file format for maturity type 1 is:

nameofmaturestocksandratio	$name_1$ $ratio_1$
	$\vdots$
	$name_n$ $ratio_n$
coefficients	α, β, γ
minmatureage	*minmatureage*

where $name_1, \ldots, name_n$ are the names of the mature stocks, i.e. the stocks to which the population that becomes mature should be transferred. α, β and γ are constants used when calculating the maturity. The $ratio_i$ is the number by which the size of the maturing part of stock will be multiplied just before adding it to the mature stock $name_i$.

Thus, the ratio allows e.g. for having female and male mature parts of stock by letting ratio equal 0.5 for them both. On the other hand, if there are, say, 2 mature stocks that only differ in their age compositions all of the younger part of the maturing stock should move to the younger mature stock, and likewise for the older part of the maturing stock. Hence, ratio should equal 1 for both of the mature stocks in this case.

minmatureage is the age at which the maturation process sets in. Immature cod for example be from 2-8 year and mature cod 4-11 year old. 2 and 3 year old cod that matures will be lost. This problem is solved by specifying *minmatureage* as 4.

The file format for maturity type 2 is:

nameofmaturestocksandratio	$name_1$ $ratio_1$
	$\vdots$
	$name_n$ $ratio_n$
maturitysteps	$s_1 \ldots s_o$
maturitylengths	$l_1 \ldots l_o$

These are the names of the mature stocks, the ratio and the constants that are used when calculating the maturity. On step s_i, the probability that an immature whose length is $\geq l_i$ becomes mature is 1. If there does not exist an i such that s_i is the current step and the length of the immature is $\geq l_i$, the probability that it becomes mature is 0.

The file format for maturity type 3 is the same as for maturyty type one except the coefficients are 4 instead of 3. Maturity type 3 is identical to maturity type 1 except the function is formulated differently in terms of l_{50} and a_{50}. This different formulation can help in minimization problems. The file format is:

nameofmaturestocksandratio	$name_1$ $ratio_1$
	$\vdots$
	$name_n$ $ratio_n$
coefficients	$\alpha, l_{50}, \beta, a_{50}$
minmatureage	*minmatureage*

For comments see comments on MaturityA.

4.8 Movement (transition).

The oldest agegroup can move from one stock to another. The file format for the movement (called transition in the program) is:

transitionstep	*ts*
transitionstock	*name*

The transition is expected to be an annual event and *ts* is the step on which it happens; *name* is the name of the stock to which the agegroup will move.

4.9 Renewal.

The renewal (should maybe be called recruitment) has two applications. First is the ordinary renewal, e.g. read the number and length distribution of recruits; and then there is the possibility of using it to set the length distribution and the mean weights, given the total number (which is not used in this version of the program).

Two options are available.

- Specifying the the complete lengthdistribution and the weight for each length.
- Giving only mean length and standard deviation as well as coefficients a and b in a length weight relation ship $w = \alpha l^{\beta}$. If the first line in the file contains the word "Normaldistribution" then this option is used

The first three lines with the first option and lines 2 to 4 with the second option (after the line containing "Normaldistribution") are: The first option was earlier in the program but the second option (Normaldistribution) is probably more useful.

minlength	B_1
maxlength	B_2
dl	dl

The rest of the file depends on which option is used. (Complete length distribution or normaldistribution.

4.9.1 Giving complete length and weight distribution

year	step	area	age	minl	No.	N	Distribution	Mean weights
y_1	s_1	a_1	A_1	l_1	n_1	$N_1{}^{\#}$	$t_{1,1} \dots t_{1,n_1}$	$W_{1,1} \dots W_{1,n_1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_p	s_p	a_p	A_p	l_p	n_p	$N_p{}^{\#}$	$t_{p,1} \dots t_{p,n_p}$	$W_{p,1} \dots W_{p,n_p}$

The header is to be omitted from the file. The columns year, step, area and age are obvious. The length distribution is in Distribution and the corresponding mean weights in Mean weights. The minimum length in the distribution is in minl, and the number of columns in Distribution and Mean weights is in No. The total number is in N; Distribution multiplied with N gives the absolute numbers in each length group.

At the top of the file is the mininum length possible in renewal, B_1, the maximum length, B_2, and the size of the length groups, dl.

It is assumed that the data is ordered w.r.t. the first two columns, i.e. in correct order in time.

To use the renewal for only setting the length distribution, put 0 in the total number column (N).

The total number column (N) is the only variables that can be marked. (formula)

4.9.2 Using normal length distribution and length-weight relationships

year	step	area	age	Number	Mean length	Standard deviation	L-W coefficient α	L-W coeffieicnt β

The number of fishes is scaled so the total number is $10000N$. The variables here that can be marked (as formula) are N, L, σ, α and β The values of the normal distribution are calculated at mean lengths of the lengthgroup distribution specified in lines 2 to 4. Most of the distribution can end outside the lengthgroup distibution. In that case the total number of fishes will still be correct but the distribution will not be similar to a normal distribution.

4.10 Spawning.

When reading spawning information for the age groups $A_1 \dots A_2$, the file format is:

spawnsinareas	$a_1 \dots a_n$
spawningstep	$s_{A_1} \dots s_{A_2}$
spawningratio	$r_{A_1} \dots r_{A_2}$
spawningmortality	m
spawningmortalitypattern	$m_{A_1} \dots m_{A_2}$
spawningweightloss	p
spawningweightlosspattern	$p_{A_1} \dots p_{A_2}$

Then the spawning mortality for age group a is $m * m_a$ and the weight loss in spawning is $p * p_a$, $A_1 \leq a \leq A_2$.

Chapter 5

Other input files.

5.1 Other food.

The format of the other food file is:

livesonareas $a_1 \ldots a_n$

lengths l_1 l_2

amount

$$\begin{array}{lllll} y_1 & s_1 & A_{1,1} & \ldots & A_{1,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_r & s_r & A_{r,1} & \ldots & A_{r,n} \end{array}$$

This defines other food on areas $a_1 \ldots a_n$. The amount of the other food is read from the amount matrix, where y_i is year and s_i is step. The i-th line in the amount matrix describes the density of other food on step s_i on the year y_i for the areas $a_1, \ldots, a_n$.

The lengths l_1 and l_2 describe the length of the other food. From these lengths the mean length is derived which is used when calculating the suitability for the other food as a prey. The lengths should not be interpreted as a physical decsription of the other food.

The amount matrix should be sorted by the first two columns, i.e. in correct order in time. The data there must cover the period of time in the simulation.

5.2 Fleetfiles, linear and total.

Fleet can be linearfleet and totalfleet. The file format for totalfleet is:

livesonareas $a_1 \dots a_n$
lengths l_1 l_2
<file format for totalpredator> *(sec. 5.2.1)*

amount

$$\begin{array}{lllll} y_1 & s_1 & A_{1,1} & \dots & A_{1,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_r & s_r & A_{r,1} & \dots & A_{r,n} \end{array}$$

and the file format for linearfleet is:

livesonareas $a_1 \dots a_n$
lengths l_1 l_2
<file format for linearpredator with multiplication> *(sec. 5.2.2)*

amount

$$\begin{array}{lllll} y_1 & s_1 & A_{1,1} & \dots & A_{1,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_r & s_r & A_{r,1} & \dots & A_{r,n} \end{array}$$

As usual, this means that the fleet will operate on the areas $a_1, \dots, a_n$ and that the fleet will have one length group with lower length l_1 and upper length l_2. Of course these lengths are not to be taken as a physical description of the fleet, only the mean length in the length group is used and only when calculating the suitability for the preys. It is expected that the amounts are given for every timestep in the simulation, and are given in correct order in time.

5.2.1 TotalPredator.

The file format for the total predator is simply the same as for suitability (section 4.4).

5.2.2 LinearPredator.

The file format for linear predator depends on whether it shall read "multiplication" or not. If not, the format is simply the same as for suitability (section 4.4). Else, the file format for suitability is to be followed with a nonnegative number, a multiplicative constant, i.e.

<file format for suitability> *(sec. 4.4)*
multiplicative $mult^{\#}$

5.3 Action At Times.

The file format for action at times is quite simple. It is just

<list of years and step>

and it specifies the steps at which some action is to take place.

The <list of years and step> contains two columns, the first one for years, the second for steps. The allowed values are either numbers or the word “all”. E.g.

```
1980  7
1983  5
all   3
1990  all
```

means that action should take place when year = 1980, step = 7 and when year = 1983, step = 5, when step = 3 and when year = 1990. It is also possible to supply “all” in both the year and step column, so the action will take place on every timestep.

Chapter 6

Input files for printing.

The file format for the printfile is:

stockstdprinter

<file format for standard stock printer> *(sec. 6.1.1)*

predpreystdlengthprinter

<file format for standard predator-prey printer by length> *(sec. 6.1.2)*

predpreystdageprinter

<file format for standard predator-prey printer by age> *(sec. 6.1.3)*

stockfullprinter

<file format for stock full printer> *(sec. 6.1.4)*

stockpreyfullprinter

<file format for stockprey full printer> *(sec. 6.2.1)*

stockprinter

<file format for stock printer> *(sec. 6.2.2)*

predatorprinter

<file format for predator printer> *(sec. 6.2.3)*

predatoroverprinter

<file format for predator overconsumption printer> *(sec. 6.2.4)*

preyoverprinter

<file format for prey overconsumption printer> *(sec. 6.2.5)*

Each of these blocks may be omitted or repeated as the user wishes. Their order is irrelevant. The purpose of the blocks is to describe what information should be written to files and how it should be written. For description of the output files, see chapter 8.

6.1 Standard printing.

6.1.1 Standard stock printer.

The file format is:

stockname	*name*	
scale	s	
areas	$a_1 \ldots a_n$	
printfile	*file*	
YearsAndSteps	<file format for action at times>	*(sec. 5.3)*

Then standard stock printout for the stock *name* will be written to *file* for the areas $a_1, \ldots a_n$ and stock sizes will be scaled with s. This will be done for the time steps in <file format for action at times>.

6.1.2 Standard predator-prey printer by length.

predator	*pred*	
prey	*prey*	
areas	$a_1 \ldots a_n$	
printfile	*file*	
YearsAndSteps	<file format for action at times>	*(sec. 5.3)*

Then information on the predation of each length group of *pred* on each length group of *prey* on the areas $a_1, \ldots, a_n$ will be written to *file* for the time steps in <file format for action at times>.

6.1.3 Standard predator-prey printer by age.

predator	*pred*	
prey	*prey*	
areas	$a_1 \ldots a_n$	
printfile	*file*	
YearsAndSteps	<file format for action at times>	*(sec. 5.3)*

Then information on the predation of each age group of *pred* on each age group of *prey* on the areas $a_1, \ldots, a_n$ will be written to *file* for the time steps in <file format for action at times>.

6.1.4 Stock full printer.

Full printout on stock sizes can be requested; the file format is:

stockname	*name*	
areas	$a_1 \ldots a_m$	
ages	$A_1 \ldots A_n$	
printfile	*file*	
YearsAndSteps	<file format for action at times>	*(sec. 5.3)*

6.2 Nonstandard printing.

Much of the information in the simulation could be calculated at a much too fine scale than may be of interest for actual output. Therefore the length groups into which the output should be divided (or actually aggregated, since it is usually calculated at a finer scale) when written to file must be supplied. It may be desirable to combine several areas into one, many agegroups into one and even several stocks into one when writing to file. This information is detailed in the blocks.

Due to similarity between the formats, it will only be explained in some detail for stock printer (section 6.2.2).

6.2.1 Stockprey full printer.

For each stock that acts as a prey, a full printout can be requested. The file format is:

stockname	*name*	
areas	$a_1 \ldots a_m$	
printfile	*file*	
YearsAndSteps	<file format for action at times>	*(sec. 5.3)*

6.2.2 Stock printer.

The stock printer defines how the abundance numbers of stocks should be written to file.

stocks	$name_1$ $\vdots$ $name_l$	
areas	$a_{1,1} \ldots a_{1,n_1}$ $\vdots$ $a_{m,1} \ldots a_{m,n_m}$	
ages	$A_{1,1} \ldots A_{1,q_1}$ $\vdots$ $A_{p,1} \ldots A_{p,q_p}$	
lengths	$l_1 \ldots l_r$	
printfiles	<number filename> <weight filename>	
YearsAndSteps	<file format for action at times>	*(sec. 5.3)*

Here, $name_1$, ... , $name_l$ are the names of the stocks over which the summation should take place and the lengths l_1, ... l_r are the endpoints of the length groups into which the output is divided. Note that the length groups given here must be a union of those of the stocks $name_1$, ... , $name_n$.

On every timestep when the data is written into file, it will be written for m "areas", where the i-th "area" is actually the aggregation of the areas $a_{i,1}$, ... , a_{i,m_i}.

Likewise, for every "area", the output will have p lines, where the i-th line is the aggregation of the agegroups $A_{i,1}$, ... , A_{i,q_i}.

The <file format for action at times> specifies the steps at which the action is to take place, the action here being that of writing to file.

The abundance numbers will be written to the file <number filename> and the corresponding mean weights to <weight filename>.

6.2.3 Predator printer.

The predator printer file defines how the eating data should be written to file.

predators	$prednam e_1$	
	$\vdots$	
	$prednam e_l$	
preys	$preynam e_1$	
	$\vdots$	
	$preynam e_q$	
areas	$a_{1,1} \quad \ldots \quad a_{1,n_1}$	
	$\vdots$	
	$a_{m,1} \quad \ldots \quad a_{m,n_m}$	
predatorlengths	$L_1 \quad \ldots \quad L_r$	
preylengths	$l_1 \quad \ldots \quad l_s$	
printfile	<filename>	
YearsAndSteps	<file format for action at times>	*(sec. 5.3)*

The amount the predators $prednam e_1$, ... , $prednam e_l$ eat of the preys $preynam e_1$, ... , $preynam e_q$ will be written to the file <filename>. The predators will be divided into length groups with the endpoints L_1, ... , L_r and the preys into the length groups l_1, ... , l_s.

6.2.4 Predator overconsumption printer.

It is possible that predators try to eat more of preys than exists of them. In order to keep track of that, i.e. the overconsumption, there is a file format describing how it should be written to file:

predators	$prednam e_1$	
	$\vdots$	
	$prednam e_l$	
areas	$a_{1,1} \quad \ldots \quad a_{1,n_1}$	
	$\vdots$	
	$a_{m,1} \quad \ldots \quad a_{m,n_p}$	
predatorlengths	$L_1 \quad \ldots \quad L_r$	
printfile	<filename>	
YearsAndSteps	<file format for action at times>	*(sec. 5.3)*

This will print the total amount that the predators wanted to eat, but were refused to.

6.2.5 Prey overconsumption printer.

When a predator tries to eat more of a prey than exists of it, the prey records the overconsumption of itself. It is possible to keep track of the total overconsumption of preys:

preys	$preyname_1$			
	$\vdots$			
	$preyname_q$			
areas	$a_{1,1}$	$\ldots$	a_{1,n_1}	
	$\vdots$			
	$a_{m,1}$	$\ldots$	a_{m,n_p}	
preylengths	l_1	$\ldots$	l_s	
printfile	<filename>			
YearsAndSteps	<file format for action at times>			*(sec. 5.3)*

Chapter 7

Input files for likelihood.

The format of the likelihood file is **one of**:

SurveyIndices	<file format for survey indices>	*(sec. 7.2)*
UnderStocking	<file format for understocking>	*(sec. 7.6)*
CatchDistribution	<file format for catch distribution>	*(sec. 7.3)*
CatchStatistics	<file format for catch statistics>	*(sec. 7.4)*
StomachContent	<file format for stomach content>	*(sec. 7.5)*

Only one of these can appear in any file, but the number of files is unlimited (see the discussion in section 2.3).

7.1 On aggregation.

Often, the calculations in a simulation are done to a much too fine scale, so it is sometimes necessary to aggregate areas, ages, length groups, stocks, or fleets before comparing to data from measurements. Most of the likelihood component formats offer aggregation, and the format of the aggregation is always the same.

When aggregating by areas or ages, the file format is usually

$$\begin{array}{ll} \text{areas} & a_{11} \dots a_{1n_1} \\ & \vdots \\ & a_{m1} \dots a_{mn_m} \end{array}$$

This means that the areas $a_{i1}, \dots, a_{in_i}$ are aggregated to the area group, say $\hat{a}_i$, for $i = 1, \dots m$. In all calculations of the likelihood component and the format for the likelihood data, the areas $a_{i1}, \dots, a_{in_i}$ will then be treated as the single area $\hat{a}_i$. In the description of the file formats for the likelihood components, the areas $\hat{a}_i$ are usually referred to as “area i”.

When aggregating by length groups, the file format is usually

lengths $l_1 \dots l_n$

The length group division given by the file format must be coarser than of the stock (or the predator, the prey etc.) in question, i.e. each length group must be the union of one or more adjacent length groups in the stock. This length group division is then used in all calculations of the likelihood component, and the file format for the likelihood data. In the description of the file format for the likelihood component, the length group $[l_i, l_{i+1}]$ is usually referred to as "length group i".

When aggregating by stocks or fleets, the file format is ususally

stocks $stock_{11} \dots stock_{1n}$
$\vdots$
$stock_{m1} \dots stock_{mn_m}$

with the case $m = 1$ being the most common. As before, this means that the stocks $stock_{i1}, \dots, stock_{im_i}$ are aggregated and treated as a single stock in all calculations and the file format for the data.

7.2 Survey Indices.

The file format for Survey Indices is either:

stocknames $name_1 \dots name_k$
areas $a_1 \dots a_l$
ages $A_{11} \dots A_{1m_1}$
$\vdots$
$A_{n1} \dots A_{1m_n}$
<file format for SI on step> *(sec. 7.2.1)*
$\vdots$
<file format for SI on step>

or

stocknames $name_1 \dots name_k$
areas $a_1 \dots a_m$
lengths $l_1 \dots l_o$
<file format for SI on step> *(sec. 7.2.1)*
$\vdots$
<file format for SI on step>

7.2.1 SI on step.

The file format for SI on step is:

fittype		*fittype*
		slope α
		intersept β
y_1	s_1	$i_{11} \ldots i_{1n}$
$\vdots$	$\vdots$	$\vdots$
y_m	s_m	$i_{m1} \ldots i_{mn}$

Here *fittype* is `LogLinearFit`, `FixedSlopeLogLinearFit` or `FixedLogLinearFit`. In the first case, neither α nor β are to be specified, in the second case α must be given, and in the last case, both α and β must be given.

7.3 Catch Distribution.

The file format for catch distribution is:

functionnumber	*no*
overconsumption	1
minimumprobability	*mimprob*
fleetnames	*fleetname*$_1$ … *fleetname*$_t$
stocknames	*stockname*$_1$ … *stockname*$_k$
areas	$a_{11} \ldots a_{1n_1}$
	$\vdots$
	$a_{m1} \ldots a_{mn_m}$
ages	$A_{11} \ldots A_{1m_1}$
	$\vdots$
	$A_{p1} \ldots A_{pm_p}$
lengths	$l_1, \ldots, l_r$

DistributionData

y_1	s_1	<file format for distribution data 1>	*(sec. 7.3.1)*
$\vdots$	$\vdots$	$\vdots$	
y_f	s_f	<file format for distribution data f>	

Here *no* is the number of the likelihood function to be used, and overconsumption is set to 1 if overconsumption of the stocks is to be taken into account – 0 otherwise. minimumprobability is used if the outcome that occurs is very inprobable. If the proportion is a group is less than $\frac{1}{minpropN}$ the proportion is set to *minpropN*. N is here the total numer of individiuals whose number is to be compared to proportions calculated by the model. Typical value of *minprop* has been 100 and overconsuption has been set to 0

The number of individuals of the stocks $stockname_1$, ... , $stockname_k$ caught by the fleets $fleetname_1$, ... , $fleetname_t$ in each age–length group will be compared to the distribution data on year y_i and step s_i. Each length group of $l_1, \ldots, l_r$ must be a union of the length groups of the stocks.

7.3.1 Distribution Data.

The file format for the distribution data is:
<file format for distribution on areas 1>
⋮
<file format for distribution on areas m>
where the file format for distribution on areas is:

$$\begin{array}{lll} N_{11} & \ldots & N_{1,r-1} \\ \vdots & & \vdots \\ N_{p1} & \ldots & N_{p,r-1} \end{array}$$

Here N_{aj} is a number indicating how many individuals the fleets $fleetname_1$, ... , $fleetname_t$ caught of the stocks $stockname_1$, ... , $stockname_k$ in the age groups A_{a1}, ... ,A_{am_a} and the length group j.

To compare this data to the results of a simulation, the multinomial distribuiton is used. If only one agegroup is read in, i.e. if $p = 1$, the likelihood function used is

$$\ell = \sum_{steps} \sum_{areas} \left(\log m! - \sum_{l=1}^{r-1} \log N_{1l}! + \sum_{l=1}^{r-1} N_{1l} \log \pi_l \right)$$

where π_l are the attribute frequencies from the simulation, and $m = \sum_l N_{1l}$ is the sample size. However, if $\pi_l = 0$ and $N_{1l} \neq 0$, for any l, then π_l is replaced by $1/(m * largenumber)$. Currently *largenumber* is set to 20. Note that $\sum_{l=1}^{r-1} \pi_l = 1$. In other words, length distribution within the agegroup is compared to the ratios from the simulation.

If on the other hand more than one agegroup is read in ($p > 1$) then the likelihood function value is

$$\ell = \sum_{steps} \sum_{areas} \sum_{l=1}^{r-1} \left(\log m_l! - \sum_{a=1}^{p} \log N_{al}! + \sum_{a=1}^{p} N_{al} \log \pi_{al} \right)$$

where π_{al} are the attribute frequencies from the simulation, and $m_l = \sum_a N_{al}$ is the sample size within each length group. However, as in the former case, if $\pi_{al} = 0$ and $N_{al} \neq 0$, for any a, l, then π_{al} is replaced by $1/(m * largenumber)$. Note that $\sum_{i=1}^{p} \pi_{ij} = 1$

for each $j = 1, \ldots, r-1$. In other words, the age distribution within each length group is compared to the ratios from the simulation.

7.4 Catch Statistics.

The file format for catch statistics is:

functionnumber	*no*
overconsumption	1
fleetnames	$\mathit{fleetname}_1 \ \ldots \ \mathit{fleetname}_t$
stocknames	$\mathit{stockname}_1 \ \ldots \ \mathit{stockname}_k$
areas	$a_{11} \ \ldots \ a_{1n_1}$
	$\vdots$
	$a_{m1} \ \ldots \ a_{mn_m}$
ages	$A_{11} \ \ldots \ A_{1m_1}$
	$\vdots$
	$A_{p1} \ \ldots \ A_{pm_p}$

StatisticsData
y_1 s_1 <file format for statistics data 1> *(sec. 7.4.1)*
$\vdots$ $\vdots$ $\vdots$
y_f s_f <file format for statistics data f>

Here *no* is the number of the likelihood function to be used.

The mean weight or length, and the standard deviation of the weight or length of the individuals of the stocks $\mathit{stockname}_1, \ldots, \mathit{stockname}_k$ caught by the fleets $\mathit{fleetname}_1$, $\ldots$, $\mathit{fleetname}_t$ in each age-area group will be compared to the statistics data on year y_i and step s_i, $i = 1, \ldots, f$.

7.4.1 Statistics Data.

The file format for the statistics data depends on the function number, *no.* For function numbers 2 and 3 it is:

$$\begin{array}{llll}
\text{numbers} & N_{11} & \dots & N_{1p} \\
 & \vdots & & \vdots \\
 & N_{m1} & \dots & N_{mp} \\
\text{mean} & x_{11} & \dots & x_{1p} \\
 & \vdots & & \vdots \\
 & x_{m1} & \dots & x_{mp} \\
\text{variance} & s^2_{11} & \dots & s^2_{1p} \\
 & \vdots & & \vdots \\
 & s^2_{m1} & \dots & s^2_{mp}
\end{array}$$

but for function numbers 1 and 4 the variance should be omitted.

Here x_{ra} is either the mean length or mean weight of individuals the fleets *fleetname*$_1$, … , *fleetname*$_t$ caught of the stocks *stockname*$_1, \dots,$ *stockname*$_k$ on the (combination of) areas $a_{r1}, \dots, a_{rn_r}$, in the age groups $A_{a1}, \dots, A_{am_a}$, s^2_{ra} is the corresponding variance (in length or weight) and N_{ra} the sample size. Note that the columns are indexed by age-groups, and the rows by area-groups.

The likelihood function in case of function numbers 2 or 3 is weighted sum of squares:

$$\ell = \sum_{steps} \sum_{\substack{area \\ r}} \sum_{\substack{age \\ a}} \frac{(x_{ra} - \mu_{ra})^2}{s^2_{ra}} N_{ra},$$

where x_{ra} is either the sample mean length (function number 2) or weight (function number 3), μ_{ra} is the corresponding mean from the simulation, s^2_{ra} is the sample variance (in length or weight) and N_{ra} the sample size. If any one of the numbers x_{ra}, μ_{ra} or s^2_{ra} is zero (or negative) the corresponding term in the sum is omitted.

For function number 1, the likelihood function is the same as for 2, except that instead of the read-in variances, the corresponding variances from the model are used, i.e.

$$\ell = \sum_{steps} \sum_{\substack{area \\ r}} \sum_{\substack{age \\ a}} \frac{(x_{ra} - \mu_{ra})^2}{\sigma^2_{ra}} N_{ra}.$$

For function number 4, the likelihood function is the same as for 3, except that s^2_{ra} is set to 1 (i.e. no weighing):

$$\ell = \sum_{steps} \sum_{\substack{area \\ r}} \sum_{\substack{age \\ a}} (x_{ra} - \mu_{ra})^2 N_{ra}.$$

7.5 Stomach Content.

The file format for stomach content is:

functionnumber	no
minimumprobability	$minp$
predators	$predator_1 \dots predator_t$
predatorlengths	$l_1 \dots l_n$
preys	$prey_{11} \dots prey_{1m_1}$
	lengths $l_{11} \dots l_{1n_1+1}$
	digcoeff $d_{11} \dots d_{1n_1}$
	$\vdots$
	$prey_{k1} \dots prey_{km_k}$
	lengths $l_{k1} \dots l_{kn_k}$
	digcoeff $d_{k1} \dots d_{kn_k+1}$
areas	$a_{11} \ \dots \ a_{1s_1}$
	$\vdots$
	$a_{m1} \ \dots \ a_{ms_m}$

StomachContentData

y_1 s_1 <file format for stomach content data 1> *(sec. 7.5.1)*

$\vdots$ $\vdots$ $\vdots$

y_f s_f <file format for stomach content data f>

Here no is the number of the likelihood function to be used. The length group division given with predatorlengths must be coarser than the length group division for each of the predators (growthandeatlengths), and the same applies to the length group divisions given by lengths for each combination of preys. The comsumption of the preys by the predators will be calculated for each combination of areas.

The user must be careful when specifying the length group division of other food that is kind of arbitrary. If the lengthgroupdivision of other food is specified as 30 40 in the otherfood file and 20 30 in the stomachcontent file the calculated amount of other food in the stomachs will be 0. (the intersection of 20-30 and 30-40 is empty).

For each prey there are 5 digestioncoefficients which allow in a simple way to take into account that different preys and different lengthgroup of preys do not have to be digested equally fast by a predator. Most commonly used values of them (1 0 0 0 0). Of the digestioncoefficients only the first three are currently used.

The digestionrate of a prey of length l is calculated as follows

$$dr_{prey,l} = p[0]$$

7.5.1 Stomach Content Data.

The file format for stomach content data is:
<file format for stomach content data on areas 1>
⋮
<file format for stomach content data on areas m>
where the file format for stomach content data on areas depends on the type of likelihood function to be used.

7.5.2 Sum of Squares of Errors in Ratios.

This is indicated by specifying functionnumber 1. In this case the file format for stomach content on areas is

$$\begin{array}{llllllll} \text{amount} & w_{11}, & \dots, & w_{1n_1}, w_{1,n_1+1}, & \dots, & w_{1,n_1+n_2}, & \dots, & w_{1,\sum n_i} \\ & \vdots & & \vdots & & \vdots & & \vdots \\ & w_{n-1,1}, & \dots, & w_{n-1,n_1}, w_{n-1,n_1+1}, & \dots, & w_{n-1,n_1+n_2}, & \dots, & w_{n-1,\sum n_i} \end{array}$$

Here $w_{i,j}$, where $j = n_1 + \cdots + n_{s-1} + l, \quad 1 \leq l \leq n_s, \quad 1 \leq s \leq k$, is the amount (weight) length group i of the predators $predator_1, \dots, predator_t$ ate of length group l of the preys $prey_{s1}, \dots, prey_{sm_s}$.

In calculating the likelihood value, the ratio of the amount length group i of the predators $predator_1, \dots, predator_t$ ate of length group l of the preys $prey_{s1}, \dots, prey_{sm_s}$, to the total amount the same length group of the predators ate of the preys $prey_{11}, \dots,$ $prey_{1m_1}, \dots, prey_{k1}, \dots, prey_{km_k}$ is calculated, subtracted from the corresponding ratio from the simulation and the square of the result, added to the likelihood value, i.e. for each simulation

$$\ell = \sum_{steps} \sum_{areas} \sum_{L} \sum_{l} (r_{Ll} - \hat{r}_{Ll})^2,$$

where

$$\mathcal{S}_L = \sum_{l} w_{Ll} \qquad \text{and} \qquad r_{Ll} = \frac{w_{Ll}}{\mathcal{S}_L},$$

and $\hat{r}_{Ll}$ is the corresponding ratio from the simulation. Note however, that if $\mathcal{S}_L = 0$ for any L on any combination of areas on any timestep, the corresponding term is omitted from the calculations.

7.5.3 Weighted Sum of Squares of Errors in Ratios.

This is indicated by specifying functionnumber 2. In this case additional information is needed on the number of samples in each length group of predator, and the standard deviation of the amount each length group of the predators consumed of each length group of the preys. Hence, the file format is

$$\begin{array}{llllllll}
\text{amount} & w_{11}, & \dots, & w_{1n_1}, w_{1,n_1+1}, & \dots, & w_{1,n_1+n_2}, & \dots, & w_{1,\sum n_i} \\
 & \vdots & & \vdots & & \vdots & & \vdots \\
 & w_{n-1,1}, & \dots, & w_{n-1,n_1}, w_{n-1,n_1+1}, & \dots, & w_{n-1,n_1+n_2}, & \dots, & w_{n-1,\sum n_i} \\
\text{stddev} & \sigma_{11}, & \dots, & \sigma_{1n_1}, \sigma_{1,n_1+1}, & \dots, & \sigma_{1,n_1+n_2}, & \dots, & \sigma_{1,\sum n_i} \\
 & \vdots & & \vdots & & \vdots & & \vdots \\
 & \sigma_{n-1,1}, & \dots, & \sigma_{n-1,n_1}, \sigma_{n-1,n_1+1}, & \dots, & \sigma_{n-1,n_1+n_2}, & \dots, & \sigma_{n-1,\sum n_i} \\
\text{number} & t_1, \dots, t_{n-1} & & & & & &
\end{array}$$

The likelihood value is calculated in the same way as with functionnumber 1, except each term is weighed with $\frac{t_i}{\sigma_{ij}^2} \sum_k w_{ik}$, i.e.

$$\ell = \sum_{steps} \sum_{areas} \sum_{L} \mathcal{S}_L t_L \sum_{l} \frac{(r_{Ll} - \hat{r}_{Ll})^2}{\sigma_{Ll}^2},$$

where

$$\mathcal{S}_L = \sum_{l} w_{Ll} \qquad \text{and} \qquad r_{Ll} = \frac{w_{Ll}}{\mathcal{S}_L},$$

and $\hat{r}_{Ll}$ is the corresponding ratio from the simulation. Note however, that if $\mathcal{S}_L = 0$ for any L or if $\sigma_{Ll} = 0$ for any L and l, on any combination of areas on any timestep, the corresponding term is omitted from the calculations.

7.5.4 Weighted Sum of Squares of Errors.

This is indicated by specifying functionnumber 3. The file format is the same as for functionnumber 2, but in calculation of the likelihood value, absolute numbers, i.e. the amounts, are used instead of ratios. Hence, in this case

$$\ell = \sum_{steps} \sum_{areas} \sum_{L} t_L \sum_{l} \frac{(w_{Ll} - \hat{w}_{Ll})^2}{\sigma_{Ll}^2}.$$

and $\hat{w}_{Ll}$ is the amount, obtained from the simulation, corresponding to w_{Ll}. As before, if $\sigma_{Ll} = 0$ for any L and l on any combination of areas on any timestep, the corresponding term is omitted from the calculations. However, even though $\mathcal{S}_L = 0$, the corresponding term is not omitted.

7.6 Understocking

The file format for understocking is:

fleets	$\mathit{fleet}_1 \ldots \mathit{fleet}_t$	
areas	$a_{11} \; \ldots \; a_{1n_1}$	
	$\vdots$	
	$a_{m1} \; \ldots \; a_{mn_m}$	
YearsAndSteps	<file format for actionattimes>	*(sec. 5.3)*

The likelihood function is the sum of squares of the overconsumption on each combination of areas, i.e.

$$\ell = \sum_{steps} \sum_{r=1}^{m} U_r^2,$$

where U_r is the amount the fleets $\mathit{fleet}_1, \ldots, \mathit{fleet}_t$ wanted to catch, in biomass units, but were refused to.

Chapter 8

Output files.

Following is a description of the ouput files from the printers, whose input file format was given in chapter 6.

8.1 Standard output files.

The standard output files are all structured in the same way, every file contains the same number of columns.

1. **Standard stock output.**
   Columns: year, step (i.e. month), area, age, number (scaled numbers), mean length, mean weight, std.dev. of length, number caught (not implemented), number consumed, biomass consumed.
   The last column, containing the biomass consumed, is **not in the standard**.

2. **Standard predator-prey output.**
   The output is the almost the same for the output of standard printout of predator-prey interactions by age or length.
   Columns: year, step (i.e. month), area, predator length/age group, prey length/age group, consumption in numbers, consumption in biomass, mortality induced.

3. **Full stock output.**
   Columns: year, step (i.e. month), area, age, mean length of length group, number in length group, mean weight in length group

8.2 Nonstandard output files.

The non-standard printing is used when some aggregation of the models division into fleets, stocks, areas, length groups and age is wanted.

1. **Full output for stock prey.**
   Columns: year, step (i.e. month), area, age, length, number consumed of age-length group, biomass consumed of age-length group, number consumed of length group, biomass consumed of length group

2. **Stock output.**
   The output from the stock printer consists of two files, a *number file* and a *weight file*. Their format is the same in that for each time step there is a line on the form 'Year *year* Step *step*', followed by matrices, one for each combination of areas. Each matrix is preceded by a line of the form 'areas *areagrno*'. The columns are indexed by length groups, and the rows indexed by age groups. The *number file* contains the mean length in each age-length group, and the *weight file* the mean weight in each age-length group.

3. **Predator output.**
   Columns: year, step(i.e. month), areas, predator length group, prey length group, consumption in biomass.

4. **Predator overconsumption output.**
   Columns: year, step (i.e. month), areas, and one column for each length group. The last columns contain the total amount (in biomass units) the corresponding length group wanted to eat but was refused to.

5. **Prey overconsumption output.**
   Columns: year, step (i.e. month), areas, and one column for each length group. The last columns contain the total amount (in biomass units) that the predators of the model tried to eat of the corresponding length group, but were refused to.

Chapter 9

Functions.

9.1 Maximum Consumption.

The maximum consumption is:

$$m_0 e^{m_1 T - m_2 T^3} l^{m_3},$$

where T is the temperature and l is the length.

9.2 Suitability functions.

The suitability functions implemented are:

$$S_1(l, L) = \frac{\delta}{1 + e^{-\alpha - \beta l - \gamma L}},$$

where α, β and γ are constants $(\alpha = p_0, \beta = p_1, \gamma = p_2, \delta = p_3)$.

And

$$S_2(l, L) = \alpha,$$

where $\alpha(= p_0)$ is a constant.

And

$$S_3(l, L) = \begin{cases} p_0 + p_2 e^{-\frac{(\ln \frac{L}{l} - p_1)^2}{p_4}} & \text{if } \ln \frac{L}{l} \leq p_1, \\ p_0 + p_2 e^{-\frac{(\ln \frac{L}{l} - p_1)^2}{p_3}} & \text{if } \ln \frac{L}{l} > p_1. \end{cases}$$

where p is a vector of parameters.

9.3 Maturity functions.

In maturity of type 1, the probability that immature fish becomes mature is:

$$P(l,a) = \frac{1}{1-M}\frac{dM}{dt},$$

where M is the maturity ogive,

$$M(l(t),a(t)) = \frac{1}{1+e^{-\alpha-\beta l(t)-\gamma a(t)}},$$

and α, β and γ are constants.

And in maturity of type 2, it is:

$$P(l,a) = \begin{cases} 1 & \text{if there is an } i \text{ such that } s_i \text{ is the current step and } l > l_i, \\ 0 & \text{else.} \end{cases}$$

where $s_1, \dots, s_n$ is a list of steps and $l_1, \dots, l_n$ is a list of lengths.

9.4 Growth functions.

Growth functions 1 are:

$$\Delta l = \Delta t p_0 l^{p_1} \psi(p_2 T + p_3),$$

and

$$\Delta w = \Delta t q_0 W^{q_1} (\psi - q_2)(q_3 T + q_4),$$

where p and q denote the vectors of parameters, T is the temperature, and ψ is the feeding level.

Growth functions 2 mean that the growth should be read from an external file.

Growth functions 3 are:

$$\Delta w = \Delta t q_0 e^{q_1 T} \left(\left(\frac{W}{q_2}\right)^{q_4} - \left(\frac{W}{q_3}\right)^{q_5} \right),$$

where q is the vector of parameters.

Now we let

$$r := \frac{W - (p_0 + p_8(p_1 + p_2 p_8))W_{ref}}{W}, \tag{9.1}$$

where W_{ref} is the reference weight and

$$f(x) := \begin{cases} 0 & \text{if } p_3 + p_4 x \leq 0, \\ p_5 & \text{if } p_3 + p_4 x \geq p_5, \\ p_3 + p_4 x & \text{else.} \end{cases} \tag{9.2}$$

Then we let

$$\Delta l = \frac{\Delta w}{p_6 p_7 l^{p_7 - 1}} f(r). \tag{9.3}$$

Growth functions 4 are:

$$\Delta w = \Delta t \left[\frac{C}{q_0 W^{q_1}} - q_2 W^{q_3} e^{q_4 T + q_5} \right],$$

where C is the consumption. The weight update is done the same way as in growth function 3, i.e. by equations (9.1)-(9.3), with $p_8 = \psi$, the feeding level.

Growth functions 5 are:

$$\Delta w = \Delta t M q_0 e^{q_1 T} \left(\left(\frac{W}{q_2} \right)^{q_4} - \left(\frac{W}{q_3} \right)^{q_5} \right),$$

where $M = y_{year} s_{step} a_{area}$.

The length update is done in exactly the same way as in growth function 3, i.e. by equations (9.1)-(9.3).