

THE IBM User Manual

C. Marsh

2018

THE IBM User Manual (modified 2018-07-17) for use with
ibm-v2018-07-17 (rev. cd12b61)

Citation: C. Marsh (2018). THE IBM User Manual, v2018-07-17 (rev. cd12b61). . 72 p.

Contents

1	Introduction	1
1.1	Where to get THE IBM	1
1.2	System requirements	1
1.3	Necessary files	1
1.4	Getting help	2
1.5	Technical details	2
2	Model overview	3
2.1	Introduction	3
2.2	The population section	3
2.3	The estimation section	3
2.4	The observation section	3
2.5	The report section	3
3	Running THE IBM	5
3.1	Using THE IBM	5
3.2	The input configuration file	5
3.3	Redirecting standard output	5
3.4	Command line arguments	6
3.5	Constructing the THE IBM input configuration files	7
3.5.1	Commands	7
3.5.2	Subcommands	7
3.5.3	The command-block format	8
3.5.4	Commenting out lines	9
3.5.5	Determining THE IBM parameter names	9
3.6	THE IBM exit status values	10
4	The population section	11
4.1	Introduction	11
4.2	Population structure	11
4.3	The state object and the partition	12
4.4	Time sequences	12
4.4.1	Annual cycle	12
4.4.2	Mortality blocks	12
4.4.3	Initialisation	13
4.5	Population processes	14
4.5.1	Ageing	14
4.5.2	Recruitment	14
4.5.2.1	Constant recruitment	14
4.5.2.2	Beverton-Holt recruitment	14
4.5.3	Mortality	14

4.5.3.1	Constant mortality rate	15
4.6	Derived Quantities	15
4.7	Growth	15
4.8	Length-weight relationship	15
4.9	Selectivities	15
4.9.1	Constant	16
4.9.2	Knife-edge	16
4.9.3	All-values	16
4.9.4	All-values-bounded	17
4.9.5	Increasing	17
4.9.6	Logistic	17
4.9.7	Inverse logistic	17
4.9.8	Logistic producing	17
4.9.9	Double-normal	18
4.9.10	Double-exponential	18
5	The observation section	19
5.1	Observations	19
6	The report section	21
7	Population command and subcommand syntax	23
7.1	Model structure	23
7.2	Initialisation	25
7.2.1	Iterative	25
7.3	Time-steps	25
7.4	Processes	26
7.4.1	Growth Von Bertalanffy With Basic Growth	26
7.4.2	Maturity	26
7.4.3	Mortality Constant Rate Mortality	26
7.4.4	Nop	26
7.4.5	Recruitment Beverton Holt	26
7.4.6	Recruitment Constant	27
7.5	Time varying parameters	27
7.5.1	Annual Shift	28
7.5.2	Constant	28
7.5.3	Exogenous	28
7.5.4	Linear	29
7.5.5	Random Draw	29
7.5.6	Random Walk	29
7.6	Derived quantities	30
7.6.1	Abundance	30
7.6.2	Biomass	31

7.6.3	Mature Biomass	31
7.7	Selectivities	31
7.7.1	All Values	32
7.7.2	All Values Bounded	32
7.7.3	Constant	32
7.7.4	Double Exponential	32
7.7.5	Double Normal	33
7.7.6	Increasing	33
7.7.7	Inverse Logistic	34
7.7.8	Knife Edge	34
7.7.9	Logistic	34
7.7.10	Logistic Producing	35
7.8	Layers	35
7.8.1	Int Layer	35
7.8.2	Numeric Layer	35
8	Observation command and subcommand syntax	36
8.1	Observation types	36
8.2	Likelihoods	37
8.2.1	Binomial	37
8.2.2	Binomial Approx	37
8.2.3	Dirichlet	37
8.2.4	Log Normal	37
8.2.5	Log Normal With Q	37
8.2.6	Multinomial	37
8.2.7	Normal	37
8.2.8	Pseudo	37
9	Report command and subcommand syntax	37
9.1	Report commands and subcommands	37
9.1.1	Derived Quantity	38
9.1.2	Initialisation Partition	38
9.1.3	Process	38
9.1.4	Standard Header	38
9.1.5	Summarise Agents	38
9.1.6	World Age Frequency	38
10	Including commands from other files	38
11	Syntax conventions, examples and niceties	41
11.1	Input File Specification	41
11.1.1	Keywords And Reserved Characters	41
11.2	Processes	45

12 Post processing output using R	47
13 Troubleshooting	49
13.1 Introduction	49
13.2 Reporting errors	49
13.3 Guidelines for reporting a problem with THE IBM	49
14 THE IBM software license	51
15 Acknowledgements	57
16 References	59
Appendices	61
A An Appendix	61
17 Index	62

1. Introduction

THE IBM is a project I embarked on during my PhD and I hope that it will be useful to some one out in the wide world. THE IBM simulates a generalised individual based model that allows a great deal of choice in specifying the agent dynamics and model outputs. THE IBM is designed for flexibility. The THE IBM is a spatially explicit as I thought that would be a valuable model characteristic and was the area of interest for me.

The time period and annual cycle of THE IBM is completely defined by the user. It can simulate many different user defined IBM, for example removals-at-length or -age from an anthropogenic or exploitation event (e.g. fishery or other human impact), scientific survey and other biomass indices, and mark-recapture data.

The real power of THE IBM comes with management strategy evaluation and population assessment model investigation.

1.1. Where to get THE IBM

THE IBM source code is hosted on github, and can be found at <https://github.com/Craig44/IBM>.

Currently you have to compile the code, to get an executable but, the repository contains all the required thirdparty libraries and has been developed for ease of compilation.

1.2. System requirements

THE IBM is available for most IBM compatible machines running 64-bit Linux and Microsoft Windows operating systems.

Several of THE IBM's tasks are highly computer intensive and a fast processor is recommended. Depending on the model implemented, some of the THE IBM tasks can take a considerable amount of processing time.

The program itself requires a few gigs of hard-disk space but output files can consume large amounts of disk space. Depending on the number and type of user output requests, the output could range from a few hundred kilobytes to several hundred megabytes. When estimating model fits, several hundred megabytes of RAM may be required, depending on the spatial size of the model, number of categories, and complexity of processes and observations. For extremely large models, several gigabytes of RAM may occasionally be required.

1.3. Necessary files

For both 64-bit Linux and Microsoft Windows, only the binary executable `ibm` or `ibm.exe` is required to run THE IBM. No other software is required. We do not provide a version for 32-bit operating systems.

THE IBM offers little in the way of post-processing of model output, and a package available that allows tabulation and graphing of model outputs is recommended. We suggest software such as **R** (R Core Team, 2014) to assist in the post processing of THE IBM output. We provide the `CASAL2` **R** package for importing the THE IBM output into **R** (see Section 12).

1.4. Getting help

THE IBM is distributed as unsupported software. The Development Team would appreciate being notified of any problems or errors in THE IBM, please use the github page to post issues, see Section 13.2 for the recommended template for reporting issues.

1.5. Technical details

THE IBM was compiled on Linux using gcc (<http://gcc.gnu.org>), the C/C++ compiler developed by the GNU Project (<http://gcc.gnu.org>). The 64-bit Linux version was compiled using gcc version 5.2.1 20151010 Ubuntu Linux (<http://www.ubuntu.com/>). The Microsoft Windows (<http://www.microsoft.com>) version was compiled using MingW (<http://www.mingw.org>) gcc (tdm64-1) 5.1.0 (<http://gcc.gnu.org>). The Microsoft Windows(<http://www.microsoft.com>) installer was built using the Inno Setup 5 (<http://www.jrsoftware.org/isdl.php>).

The random number generator used by THE IBM uses an implementation of the Mersenne twister random number generator (Matsumoto and Nishimura, 1998). This, the command line functionality, matrix operations, and a number of other functions use the BOOST C++ library (Version 1.58.0).

2. Model overview

2.1. Introduction

THE IBM is a generalised individual based model.

THE IBM is run from the console window in Microsoft Windows or from a terminal window in Linux. THE IBM gets its information from input data files, the main one of which is the *input configuration file*. Commands and subcommands in the input configuration file are used to define the model structure, provide observations, define parameters, and define the outputs (reports) for THE IBM. Command line switches tell THE IBM the run mode and where to direct its output. See Section 3 for details.

We define the model in terms of the *state*. The state consists of two parts, the *partition*, and any *derived quantities*. The state will typically change in each *time-step* of every year, depending on the *processes* defined for those time-steps in the model.

2.2. The population section

2.3. The estimation section

2.4. The observation section

2.5. The report section

3. Running THE IBM

THE IBM is run from the console window (i.e., the command line) on Microsoft Windows or from a terminal window on Linux. THE IBM uses information from input data files – the *input configuration file* being the key file.

The input configuration file is compulsory and defines the model structure, processes, observations, parameters, and the reports (outputs) requested. The following sections describe how to construct the THE IBM configuration file. By convention, the name of the input configuration file ends with the suffix `.ibm`. However, any file name is acceptable. Note that the input configuration file can ‘include’ other files as a part of its syntax. Collectively, these are called the input configuration file.

Other input files can, in some circumstances, be supplied depending on what is required. For example adding additional layers so you do not clutter a single file and improve readability.

3.1. Using THE IBM

To use THE IBM, open a console (i.e. the command prompt) window (Microsoft Windows) or a terminal window (Linux). Navigate to a directory of your choice, where your input configuration files are located. Then enter `ibm` with any arguments (see Section 3.4 for the the list of possible arguments) to start your THE IBM job running. THE IBM will print output to the screen and return you to the command prompt when the job has completed. Note that the THE IBM executable (binary) and shared libraries (extension `.dll`) must be either in the same directory as the input configuration files or in your systems `PATH`. The THE IBM installer (**doesn’t exist**) should update your path on Windows in any case, but see your operating system documentation for help identifying or modifying your `PATH`.

3.2. The input configuration file

The input configuration file is made up of four broad sections; the description of the population structure and parameters (the population section), the observations and their associated likelihoods (the observation section), and the outputs and reports that THE IBM will return (the report section). The input configuration file is made up of a number of commands (many with subcommands) which specify various options for each of these components.

The command and subcommand definitions in the input configuration file can be extensive (especially when you have a model that has many observations), and can result in a input configuration file that is long and difficult to navigate. To aid readability and flexibility, we can use the input configuration file command `!include file` (e.g. Figure ??). The command causes an external file, *file*, to be read and processed, exactly as if its contents had been inserted in the main input configuration file at that point. The file name must be a complete file name with extension, but can use either a relative or absolute path as part of its name. Note that included files can also contain `!include` commands. See Section 10 for more detail.

3.3. Redirecting standard output

THE IBM uses the `standard output stream` to display run-time information. The standard error stream is used by THE IBM to output the program exit status and run-time errors. We suggest redirecting both the standard output and standard error into files. With the bash shell (on Linux systems), you can do this using the command structure,

```
(ibm [arguments] > out) >& err &
```

It may be useful to redirect the standard input, especially if you're using THE IBM inside a batch job software, i.e.

```
(ibm [arguments] > out < /dev/null) >& err &
```

On Microsoft Windows systems, you can redirect to standard output using,

```
ibm [arguments] > out
```

And, on some Microsoft Windows systems (e.g., Windows10), you can redirect to both standard output and standard error, using the syntax,

```
ibm [arguments] > out 2> err
```

Note that THE IBM outputs a few lines of header information to the output (e.g. Figure ??). The header consists of the program name and version, the arguments passed to THE IBM from the command line, the date and time that the program was called (derived from the system time), the user name, and the machine name (including the operating system and the process identification number). These can be used to track outputs as well as identifying the version of THE IBM used to run the model.

3.4. Command line arguments

The call to THE IBM is of the following form:

```
ibm[-c config_file] [task] [options]
```

where,

-c *config_file* Define the input configuration file for THE IBM (if omitted, THE IBM looks for a file named `config.ibm`)

and where *task* must be one of the following ([] indicates a secondary label to call the task, e.g. **-h** will execute the same task as **--help**),

-h [--help] Display help (this page)

-l [--licence] Display the reference for the software license (GPL v2)

-v [--version] Display the THE IBM version number

-r [--run] Run the model once using the parameter values in the input configuration file, or optionally, with the values from the file denoted with the command line argument **-i *file***

and where the following optional arguments [*options*] may be specified,

-g [--seed] *seed* Seed the random number generator with *seed*, a positive (long) integer value (note, if **-g** is not specified, then THE IBM will generate a random number seed based on the computer clock time)

--loglevel *arg* = {trace, finest, fine, medium} (see Section 6)

3.5. Constructing the THE IBM input configuration files

The model definition, parameters, observations, and reports are specified in input configuration files:

Population input (Section 4) specifies the model structure, population dynamics, and other associated parameters;

Observation input (Section 5) contains all the observations data available to the model and describes how the observed values should be formatted, how THE IBM calculates the expected values, and the likelihoods available for each type of observation; and

Report input (Section 6) specifies any output required.

The command and subcommand syntax to be used in each of these configuration files are listed in Sections 7 (Population), 8 (Observation) and 9 (Report).

3.5.1. Commands

THE IBM has a range of commands that define the model structure, processes, observations, and how tasks are carried out. There are three types of commands,

1. Commands that have an argument and do not have subcommands (for example, `!include file`)
2. Commands that have a label and subcommands (for example `@process` must have a label, and has subcommands)
3. Commands that do not have either a label or argument, but have subcommands (for example `@model`)

Commands that have a label must have a unique label, i.e., the label cannot be used on more than one command of that type. The labels can contain alpha numeric characters, period ('.'), underscore ('_') and dash ('-'). Labels must not contain white-space, or other characters that are not letters, numbers, dash, period or an underscore. For example,

```
@process NaturalMortality
or
!include MyModelSpecification.cs12
```

3.5.2. Subcommands

THE IBM subcommands are used for defining options and parameter values related to a particular command. Subcommands always take an argument which is one of a specific *type*. The argument *types* acceptable for each subcommand are defined in Section 10, and are summarised below.

Like commands (`@command`), subcommands and their arguments are not order specific — except that that all subcommands of a given command must appear before the next `@command` block. THE IBM may report an error if they are not supplied in this way. However, in some circumstances a different order may result in a valid, but unintended set of actions, leading to possible errors in your expected results.

The argument type for a subcommand can be either:

switch	true/false
integer	an integer number,
integer vector	a vector of integer numbers,
integer range	a range of integer numbers separated by a colon, e.g. 1994:1996 is expanded to an integer vector of values (1994 1995 1996),
constant	a real number (i.e. double),
constant vector	a vector of real numbers (i.e. vector of doubles),
estimable	a real number that can be estimated (i.e. estimable double),
estimable vector	a vector of real numbers that can be estimated (i.e. vector of estimable doubles),
addressable	a real number that can be referenced but not estimated (i.e. addressable double),
addressable vector	a vector of real numbers that can be referenced but not estimated (i.e. vector of addressable doubles),
string	a categorical (string) value, or
string vector	a vector of categorical values.

Switches are parameters which are either true or false. Enter *true* as `true` or `t`, and *false* as `false` or `f`.

Integers must be entered as integers (i.e., if `year` is an integer then use 2008, not 2008.0)

Arguments of type integer vector, integer range, constant vector, estimable vector, addressable vector, or categorical vector must contain one or more entries on a row, separated by white space (tabs or spaces).

Note that parameters defined as addressable with the subcommand type `addressable` or `addressable vector` are usually derived IBM and are not directly estimable. As such, they can be acted upon by the model (e.g. called by various processes; have priors and/or penalties assigned to them), but they do not directly contribute to any estimation within the model

3.5.3. The command-block format

Each command-block consists of a single command (starting with the symbol `@`) and, for most commands, a unique label or an argument. Each command is then followed by its subcommands and their arguments, e.g.,

<code>@command, or</code>	<code>@command argument, or</code>	<code>@command label</code>
<code>subcommand argument</code>	<code>subcommand argument</code>	<code>subcommand argument</code>
<code>subcommand argument</code>	<code>subcommand argument</code>	<code>subcommand argument</code>
<code>.</code>	<code>.</code>	<code>.</code>
<code>.</code>	<code>.</code>	<code>.</code>
<code>etc.</code>	<code>etc.</code>	<code>etc.</code>

Blank lines are ignored, as is extra white space (i.e., tabs and spaces) between arguments. However, to start command block the `@` character must be the first character on the line and must not be preceded by any white space. Each input file must end with a carriage return.

There is no need to mark the end of a command block. This is automatically recognized by either

the end of the file, section, or the start of the next command block (which is marked by the @ on the first character of a line). Note, however, that the `!include` is the only exception to this rule (see Section 10 for details of the use of `!include`).

Commands, sub-commands and arguments in the input configuration files are not case sensitive. Labels and variable values are case sensitive. Also, on a Linux system, external calls to files are case sensitive (i.e., when using `!include file`, the argument `file` will be case sensitive).

3.5.4. Commenting out lines

Text on a line that follows an # is considered to be a comment and is ignored. To comment out a group of commands or subcommands, use a # at the beginning of each line to be ignored.

Alternatively, to comment out an entire block or section place a /* as the first character on the line to start the comment block, then end it with */. All lines (including line breaks) between /* and */ inclusive are ignored.

```
# This is a comment and will be ignored
@process NaturalMortality
m 0.2
/*
This block of code
is a comment and
will be ignored
*/
```

3.5.5. Determining THE IBM parameter names

When THE IBM processes an input configuration file it translates each command block and each subcommand block into a unique THE IBM object, each with a unique parameter name. For commands, this parameter name is simply the command label. For subcommands, the parameter name format is either:

`command[label].subcommand` if the command has a label, or

`command.subcommand` if the command has no label, or

`command[label].subcommand{i}` if the command has a label and the subcommand arguments are a vector, and we are accessing the *i*th element of that vector.

`command[label].subcommand{i:j}` if the command has a label, and the subcommand arguments are a vector, and we are accessing the elements from *i* to *j* (inclusive) of that vector.

The unique parameter name is used to reference that unique parameter when, for example, estimating, applying a penalty, projecting, time varying or applying a profile. For example, the parameter name of the Natural mortality rates subcommand `m` of the command `@process` with the label `NaturalMortality` is category related and so, the syntax to reference all `m` related categories is,

```
process[NaturalMortality].m
```

Or, the syntax to specify a single category for which to apply the natural mortality process is,

```
process[NaturalMortality].m{male}
```

All labels (parameter names) are user specified. As such, naming conventions are non-restrictive and can be model specific.

3.6. THE IBM exit status values

Whether THE IBM completes its task successfully or errors out gracefully, it returns a single exit status value 'completed' to the standard output. Error messages will be printed to the console. When configuration errors are found THE IBM will print error messages, along with the associated files and line numbers where the errors were identified, for example,

```
#1: At line 15 in Reports.ibm: Parameter '{' is not supported
```

4. The population section

4.1. Introduction

The population section specifies the model of the population dynamics. It describes the model structure (population structure), defines the population processes (e.g., recruitment, migration, and mortality), the selectivities, and associated parameters.

The population section consists of several components, including:

- The population structure;
- Model initialisation (i.e., the state of the partition at the start of the first year);
- The years over which the model runs (i.e., the start and end years of the model)
- The annual cycle (time-steps and processes that are applied in each time-step);
- The specification and parameters of the population processes (i.e., processes that add, remove individuals to or from the partition, or shift numbers between ages and categories in the partition);
- Selectivities;
- The spatial domain via defining layers
- Parameter values and their definitions; and
- Derived IBM, required as parameters for some processes (e.g. mature biomass to resolve any density dependent processes, such as the spawner-recruit relationship in a recruitment process).

4.2. Population structure

The basic structure of the population section of a THE IBM model is defined in terms of an annual cycle, time steps, states, and transitions.

The annual cycle defines what processes happen in each model year, and in what sequence. THE IBM runs on an annual cycle rather than, for example, a 6-monthly cycle.

Each year is split into one or more time steps, with at least one process occurring in each time step. Each time step can be thought of as representing a particular part of the calendar year, or time steps can be treated as an abstract sequence of events. In every time step, there exists a mortality block: a group of consecutive mortality-based processes, where individuals are removed from the partition (see Section 4.4.2).

The state is the current status of the population at any given time. The state can change one or more times in each time step of every year. The state object must contain sufficient information to figure out how the underlying population changes over time (given a model and a complete set of parameters).

The state can undergo a number of possible changes, called transitions. Transitions are accomplished by processes, including: recruitment, natural mortality, anthropogenic mortality, ageing, migration, tagging events, and maturation.

The division of the year into an arbitrary number of time steps allows the user to specify the exact order in which processes and observations occur throughout the year. The user needs to specify the time step in which each process occurs. If more than one process occurs in the same time step, the order in which to apply each process is specified in the `@time_step` block.

The key element of the state is the spatial world view, which holds all the entities.

An example, to specify a model with 2 categories (male and female) with ages 1-20 (with the last age a plus group) and an age-length relationship defined with the label `male_growth` and `female_growth`, then the `@model` block is specified as:

```
@model
start_year
final_year
min_age 1
max_age 20
age_plus_group True
initialisation_phases iphase
time_steps step1 step2 step3
```

4.3. The state object and the partition

4.4. Time sequences

The time sequence of the model is defined in the following parts;

- Annual cycle
- Mortality blocks
- Initialisation
- Model run years

4.4.1. Annual cycle

The annual cycle is implemented as a set of processes that occur, in a user-defined order, within each year. Time-steps are used to break the annual cycle into separate components, and allow observations to be associated with different time periods and processes. Any number of processes can occur within each time-step, in any order (although there are limitations around mortality based processes - see Section 4.4.2) and can occur multiple times within each time-step. Note that time-steps are not implemented during the initialisation phases (effectively, there is only one time-step), and that the annual cycle in the initialisation phases can, optionally, be different from that which is applied during the model years.

4.4.2. Mortality blocks

For every time step in an annual cycle there is an associated *mortality block*. Mortality blocks are a key concept in THE IBM.

Mortality blocks are used to define the ‘point’ in the model time sequence when observations (see Section 5) are evaluated, and derived quantities (see Section 4.6) are calculated.

A mortality block is defined as a consecutive sequence of mortality processes within a time step. The processes that are mortality processes are all pre-defined in THE IBM, and cannot be modified. These mortality processes are described in subsection 4.5.3.

THE IBM requires that each time step has exactly one mortality block. To achieve this, either all the mortality processes in a time step must be sequential (i.e., there can not be a non-mortality process between any two mortality processes within any one time step); or if no mortality processes occur in a time step then the mortality block is defined to occur at the end of the time step.

THE IBM will error out if more than one mortality block occurs in a single time step.

4.4.3. Initialisation

Initialisation is the process of determining the world's state just before `start_year`, whether it be equilibrium/steady state or some other initial state for the model (e.g exploited), prior to the start year of the model. This can be computationally expensive if a plus group is present in the partition.

Currently users can only initialise the partition via an iterative process. THE IBM does a few tricks to help speed up the initialisation. The first thing THE IBM does is gets the parameter `number_of_agents` and spits that number of agents uniformly, over the spatial domain, alternatively the user could supply a layer `layer_label` of proportions to seed the initial spatial distribution. This layer should sum to one so that the model initially seeds `number_of_agents`. When the THE IBM seeds the initial number of agents it also randomly assigns the agents an age based on an exponential distribution where the parameter λ of the exponential distribution is set by the command on the `@model, initialisation_z_seed`. We suggest setting this command to the assumed natural mortality of the model. To see what age structure this would look like you can quickly use **R** to visualise. by running the following code in an **R** terminal you could see.

```
Z_param = 0.2;
agents_per_cell = 1000;
hist(rexp(agents_per_cell,Z_param), breaks = 30, xlab = "age", ylab = "frequency", main = "
  Initial age structure in each cell")
```

Once THE IBM has seeded the agents, it iterates over the annual cycle to change an approximated initialisation state to one that is more like what would occur for your annual cycle, this is controlled by the `years` command in the `@initialisation` block. The number of iterations in the iterative initialisation can effect the model output, and these should be chosen to be large enough to allow the population state to fully converge. We recommend that a period of about two generations to ensure convergence. THE IBM can be requested to report a number of convergence statistics that can assist the user in determining the level of convergence.

In addition, the iterative initialisation phase can optionally be stopped early if some user defined convergence criteria is met. For a list of supplied years in the initialisation phase, convergence is defined as met if the proportional absolute summed difference between the state in year $t - 1$ and the state in year t ($\hat{\lambda}$) is less than a user defined λ where,

$$\hat{\lambda} = \frac{\sum_i \sum_j |\text{element}(i, j)_t - \text{element}(i, j)_{t-1}|}{\sum_i \sum_j \text{element}(i, j)_t} \quad (4.1)$$

Hence, for an iterative initialisation you need to define:

- The initialisation phases,

- The number of years in each phase, and
- the natural mortality process
- the growth process

Because the initialisation phase is responsible for seeding the initial agents, users must specify processes that the initialisation phase can seed parameters to agents. To see how parameters are set for each individual agent, users should see the individual processes in this Section 4.5.

An example of the syntax to implement this would be,

```
@model
...
initialisation_phases Iterative_initialisation

@initialisation_phase Iterative_initialisation
type iterative
years 50
lambda 0.0001
convergence_years 20 40
layer_label Base
growth_process_label von_bert
natural_mortality_process_label natural_mort
```

4.5. Population processes

4.5.1. Ageing

Ageing is an implicit process in the model, Each agent that is created or recruited gets assigned an birth year. This means that when ever we want to ask for the agent we just calculate `current_year - birth_year`, thus there is no explicit ageing process. Note that we do return the `max_age` of the model, if the agent is older than that age (so we only work with the truncated age distribution).

This means every fish automatically ages by one at the end of the year, or you could think of ageing a fish at the very beginning of the year (tomayto, tomahto), just be aware that you don't have control over that, but if you want to account for growth in between annual increments that is possible via time step proportion increments, see growth processes for more information.

4.5.2. Recruitment

4.5.2.1. Constant recruitment

4.5.2.2. Beverton-Holt recruitment

4.5.3. Mortality

Mortality processes remove individuals from the model domain (and also from memory).

4.5.3.1. Constant mortality rate

4.6. Derived Quantities

Derived quantities surround a mortality block and so the value of the derived quantity is calculated twice, once before the mortality block, and once after. The final value is an interpolation based on how much mortality the user wants to into account when calculating derived quantities.

Biomass

Mature Biomass

Abundance

4.7. Growth

4.8. Length-weight relationship

Growth currently means updating length and weight for an agent, if the Von Bertalanffy formula currently is used (currently the default in agent class for initialising length) it follows the following formula

$$\Delta L = (L_{\infty} - L)(1 - e^k) \quad (4.2)$$

If the basic length weight formulation is used, then an agents weight follows the following formula

$$\bar{w} = aL^b \quad (4.3)$$

Scaling the model agents up to the population biomass estimates is done via a global scalar.

4.9. Selectivities

A selectivity is a function that can have a different value for each age class. Selectivities are used throughout THE IBM to interpret observations (Section ??) or to modify the effects of processes on each age class (Section 4). THE IBM implements a number of different parametric forms, including logistic, knife edge, and double normal selectivities. Selectivities are defined in there own command block (@selectivity), where the unique label is used by observations or processes to identify which selectivity to apply.

Selectivities are indexed by age, with indices from `min_age` to `max_age`. For example, for a logistic age-based selectivity with 50% selected at age 5 and 95% selected at age 7, would be defined by the `type=logistic` with parameters $a_{50} = 5$ and $a_{95} = (7 - 5) = 2$. The value of the selectivity at age $x = 7$ is 0.95, and the value at age $x = 3$ is 0.05. Note, while selectivities can be length based, use with caution as more testing is needed for this functionality.

The function values for some choices of parameters, for some selectivities, can result in a computer numeric overflow error (i.e., the number calculated from parameter values is either too large or too small to be represented in computer memory). THE IBM implements range checks on some parameters to test for a possible numeric overflow error before attempting to calculate function

values. For example, the logistic selectivity is implemented such that if $(a_{50} - x)/a_{t095} > 5$ then the value of the selectivity at $x = 0$, i.e., for $a_{50} = 5$, $a_{t095} = 0.1$, then the value of the selectivity at $x = 1$, without range checking would be 7.1×10^{-52} . With range checking, that value is 0 (as $(a_{50} - x)/a_{t095} = 40 > 5$).

The available selectivities are;

- Constant
- Knife-edge
- All values
- All values bounded
- Increasing
- Logistic
- Inverse logistic
- Logistic producing
- Double normal
- Double exponential

The available selectivities are described below.

4.9.1. constant

$$f(x) = C \tag{4.4}$$

The constant selectivity has the estimable parameter C.

4.9.2. knife_edge

$$f(x) = \begin{cases} 0, & \text{if } x < E \\ \alpha, & \text{if } x \geq E \end{cases} \tag{4.5}$$

The knife-edge ogive has the estimable parameter E and a scaling parameter α , where the default value of $\alpha = 1$.

4.9.3. all_values

$$f(x) = V_x \tag{4.6}$$

The all-values selectivity has estimable parameters $V_{low}, V_{low+1} \dots V_{high}$. Here, you need to provide the selectivity value for each age class.

4.9.4. all_values_bounded

$$f(x) = \begin{cases} 0, & \text{if } x < L \\ V_x, & \text{if } L \leq x \leq H \\ V_H, & \text{if } x > H \end{cases} \quad (4.7)$$

The all-values-bounded selectivity has non-estimable parameters L and H . The estimable parameters are $V_L, V_{L+1} \dots V_H$. Here, you need to provide an selectivity value for each age class from $L \dots H$.

4.9.5. increasing

$$f(x) = \begin{cases} 0, & \text{if } x < L \\ f(x-1) + \pi_x(\alpha - f(x-1)), & \text{if } L \leq x \leq H \\ f(\alpha), & \text{if } x \geq H \end{cases} \quad (4.8)$$

The increasing ogive has non-estimable parameters L and H . The estimable parameters are $\pi_L, \pi_{L+1} \dots \pi_H$ (but if these are estimated, they should always be constrained to be between 0 and 1). α is a scaling parameter, with default value of $\alpha = 1$. Note that the increasing ogive is similar to the all-values-bounded ogive, but is constrained to be non-decreasing.

4.9.6. logistic

$$f(x) = \alpha / [1 + 19^{(a_{50}-x)/a_{t095}}] \quad (4.9)$$

The logistic selectivity has estimable parameters a_{50} and a_{t095} . α is a scaling parameter, with default value of $\alpha = 1$. The logistic selectivity takes values 0.5α at $x = a_{50}$ and 0.95α at $x = a_{50} + a_{t095}$.

4.9.7. inverse_logistic

$$f(x) = \alpha - \alpha / [1 + 19^{(a_{50}-x)/a_{t095}}] \quad (4.10)$$

The inverse logistic selectivity has estimable parameters a_{50} and a_{t095} . α is a scaling parameter, with default value of $\alpha = 1$. The logistic selectivity takes values 0.5α at $x = a_{50}$ and 0.95α at $x = a_{50} - a_{t095}$.

4.9.8. logistic_producing

$$f(x) = \begin{cases} 0, & \text{if } x < L \\ \lambda(L), & \text{if } x = L \\ (\lambda(x) - \lambda(x-1)) / (1 - \lambda(x-1)), & \text{if } L < x < H \\ 1, & \text{if } x \geq H \end{cases} \quad (4.11)$$

The logistic-producing selectivity has the non-estimable parameters L and H , and has estimable parameters a_{50} and a_{t095} . α is a scaling parameter, with default value of $\alpha = 1$. For category transitions, $f(x)$ represents the proportion moving, not the proportion that have moved. This selectivity was designed for use in an age-based model to model maturity. In such a model, a logistic-producing maturation selectivity will (in the absence of other influences) make the proportions mature follow a logistic curve with parameters a_{50}, a_{t095} .

4.9.9. double_normal

$$f(x) = \begin{cases} \alpha 2^{-[(x-\mu)/\sigma_L]^2}, & \text{if } x \leq \mu \\ \alpha 2^{-[(x-\mu)/\sigma_R]^2}, & \text{if } x \geq \mu \end{cases} \quad (4.12)$$

The double-normal selectivity has estimable parameters a_1 , s_L , and s_R . α is a scaling parameter, with default value of $\alpha = 1$. It has values α at $x = a_1$, and 0.5α at $x = a_1 - s_L$ and $x = a_1 + s_R$.

4.9.10. double_exponential

$$f(x) = \begin{cases} \alpha y_0 (y_1/y_0)^{(x-x_0)/(x_1-x_0)}, & \text{if } x \leq x_0 \\ \alpha y_0 (y_2/y_0)^{(x-x_0)/(x_2-x_0)}, & \text{if } x > x_0 \end{cases} \quad (4.13)$$

The double-exponential selectivity has non-estimable parameters x_1 and x_2 , and estimable parameters x_0 , y_0 , y_1 , and y_2 . α is a scaling parameter, with default value of $\alpha = 1$. It can be ‘U-shaped’. Bounds for x_0 must be such that $x_1 < x_0 < x_2$. With $\alpha = 1$, the selectivity passes through the points (x_1, y_1) , (x_0, y_0) , and (x_2, y_2) . If both y_1 and y_2 are greater than y_0 the selectivity is ‘U-shaped’ with minimum at (x_0, y_0) .

Selectivities `all_values` and `all_values_bounded` can be addressed in additional priors using the following syntax,

```
@selectivity maturity
type all_values
v 0.001 0.1 0.2 0.3 0.4 0.3 0.2 0.1

## encourage ages 3-8 to be smooth.
@additional_prior smooth_maturity
type vector_smooth
parameter selectivity[maturity].values{3:8}
```

5. The observation section

5.1. Observations

6. The report section

The report section specifies the printouts and other outputs from the model. THE IBM does not, in general, produce any output unless requested by a valid `@report` block.

Reports from THE IBM can be defined to print partition and states objects at a particular point in time, observation summaries, estimated parameters and objective function values. See below for a more extensive list of report types, and an example of an observation report.

7. Population command and subcommand syntax

7.1. Model structure

@model *label* Define an object of type *model*

start_year Define the first year of the model, immediately following initialisation

Type: non-negative integer

Default: No Default

Value: Defines the first year of the model, ≥ 1 , e.g. 1990

final_year Define the final year of the model, excluding years in the projection period

Type: non-negative integer

Default: No Default

Value: Defines the last year of the model, i.e., the model is run from start_year to final_year

min_age Minimum age of individuals in the population

Type: non-negative integer

Default: 0

Value: $0 \leq \text{age}_{\min} \leq \text{age}_{\max}$

max_age Maximum age of individuals in the population

Type: non-negative integer

Default: 0

Value: $0 \leq \text{age}_{\min} \leq \text{age}_{\max}$

age_plus Define the oldest age or extra length midpoint (plus group size) as a plus group

Type: boolean

Default: false

Value: true, false

initialisation_phase_labels Define the labels of the phases of the initialisation

Type: string vector

Default: true

Value: A list of valid labels defined by @initialisation_phase

time_steps Define the labels of the time steps, in the order that they are applied, to form the annual cycle

Type: string vector

Default: No Default

Value: A list of valid labels defined by @time_step

length_bins

Type: non-negative integer vector

Default: true

length_plus Is the last bin a plus group

Type: boolean

Default: false

base_layer_label Label for the base layer

Type: string

Default: No Default

latitude_layer_label Label for the latitude layer

Type: string

Default: ""

longitude_layer_label Label for the longitude layer

Type: string

Default: ""

nrows number of rows in spatial domain

Type: non-negative integer

Default: No Default

Lower Bound: 1 (inclusive)

ncols number of columns in spatial domain

Type: non-negative integer

Default: No Default

Lower Bound: 1 (inclusive)

sexed Is sex an attribute of you agent?

Type: boolean

Default: false

proportion_male what proportion of the generated agents should be male?

Type: float

Default: 1.0

Lower Bound: 0.0 (inclusive)

Upper Bound: 1.0 (inclusive)

maturity_ogive_label Maturity ogive label for each sex

Type: string vector

Default: false

growth_process_label Label for the growth process in the annual cycle

Type: string

Default: No Default

natural_mortality_process_label Label for the natural mortality process in the annual

cycle
Type: string
Default: No Default

7.2. Initialisation

@initialisationphase *label* Define an object of type *initialisationphase*

label The label of the initialisation phase
Type: string
Default: No Default

type The type of initialisation
Type: string
Default: iterative

7.2.1. @initialisation_phase[label].type=iterative

years The number of iterations (years) over which to execute this initialisation phase
Type: non-negative integer
Default: No Default

number_of_agents The number of agents to initially seed in the partition
Type: non-negative integer
Default: No Default

layer_label The label of a layer that you want to seed a distribution by.
Type: string
Default: ""

7.3. Time-steps

@timestep *label* Define an object of type *timestep*

label The label of the timestep
Type: string
Default: No Default

processes The labels of the processes for this time step in the order that they occur
Type: string vector
Default: No Default

7.4. Processes

@process *label* Define an object of type *process*

label The label of the process

Type: string

Default: No Default

type The type of process

Type: string

Default: ""

7.4.1. @process[label].type=growth_von_bertalanffy_with_basic_growth

distribution the distribution to allocate the parameters to the agents

Type: string

Default: No Default

cv The cv of the distribution

Type: float

Default: No Default

7.4.2. @process[label].type=maturity

7.4.3. @process[label].type=mortality_constant_rate_mortality

distribution the distribution to allocate the parameters to the agents

Type: string

Default: No Default

cv The cv of the distribution

Type: float

Default: No Default

7.4.4. @process[label].type=nop

7.4.5. @process[label].type=recruitment_beverton_holt

b0 B0

Type: float

Default: false

steepness Steepness

Type: float

Default: 1.0

`ycs_values` YCS (Year-Class-Strength) Values

Type: constant-float vector

Default: No Default

`recruitment_layer_label` A label for the recruitment layer

Type: string

Default: No Default

`ssb` A label for the SSB derived quantity

Type: string

Default: No Default

7.4.6. `@process[label].type=recruitment_constant`

`b0` B0

Type: float

Default: false

`recruitment_layer_label` A label for the recruitment layer

Type: string

Default: No Default

`ssb` A label for the SSB derived quantity

Type: string

Default: No Default

7.5. Time varying parameters

`@timevarying` *label* Define an object of type *timevarying*

`label` The time-varying label

Type: string

Default: No Default

`type` The time-varying type

Type: string

Default: ""

`years` Years in which to vary the values

Type: non-negative integer vector

Default: No Default

`parameter` The name of the parameter to time vary

Type: string
Default: No Default

7.5.1. `@time_varying[label].type=annual_shift`

values

Type: constant vector
Default: No Default

a

Type: constant
Default: No Default

b

Type: constant
Default: No Default

c

Type: constant
Default: No Default

scaling_years

Type: non-negative integer vector
Default: true

7.5.2. `@time_varying[label].type=constant`

values Value to assign to addressable

Type: estimable vector
Default: No Default

7.5.3. `@time_varying[label].type=exogenous`

a Shift parameter

Type: estimable
Default: No Default

exogeneous_variable Values of exogeneous variable for each year

Type: constant vector
Default: No Default

7.5.4. @time_varying[label].type=linear

slope The slope of the linear trend (additive unit per year)

Type: estimable

Default: No Default

intercept The intercept of the linear trend value for the first year

Type: estimable

Default: No Default

7.5.5. @time_varying[label].type=random_draw

mean Mean

Type: estimable

Default: 0

sigma Standard deviation

Type: estimable

Default: 1

distribution distribution

Type: string

Default: normal

Allowed Values: normal, lognormal

7.5.6. @time_varying[label].type=random_walk

mean Mean

Type: estimable

Default: 0

sigma Standard deviation

Type: estimable

Default: 1

upper_bound Upper bound for the random walk

Type: constant

Default: 1

lower_bound Lower bound for the random walk

Type: constant

Default: 1

rho Auto Correlation parameter

Type: constant
Default: 1

distribution distribution
Type: string
Default: normal

7.6. Derived quantities

@derivedquantity *label* Define an object of type *derivedquantity*

label Label of the derived quantity
Type: string
Default: No Default

type Type of derived quantity
Type: string
Default: No Default

time_step The time step in which to calculate the derived quantity after
Type: string
Default: No Default

proportion_through_mortality_block Proportion through the mortality block of the time
step when calculated
Type: float
Default: 0.5
Lower Bound: 0.0 (inclusive)
Upper Bound: 1.0 (inclusive)

7.6.1. @derived_quantity[label].type=abundance

selectivity A label for the selectivity
Type: string
Default: No Default

layer_label A label for the layer that indicates which cells to calculate abundance in over
Type: string
Default: No Default

7.6.2. @derived_quantity[label].type=biomass

selectivity A label for the selectivity

Type: string

Default: No Default

biomass_layer_label A label for the layer that indicates which cells to calculate biomass over

Type: string

Default: No Default

7.6.3. @derived_quantity[label].type=mature_biomass

biomass_layer_label A label for the layer that indicates which cells to calculate biomass over

Type: string

Default: No Default

7.7. Selectivities

@selectivity label Define an object of type *selectivity*

label The label for this selectivity

Type: string

Default: No Default

type The type of selectivity

Type: string

Default: No Default

length_based Is the selectivity length based

Type: boolean

Default: false

values

Type: Addressable vector

Default: No Default

length_values

Type: Addressable vector

Default: No Default

7.7.1. @selectivity[label].type=all_values

v V
Type: estimable vector
Default: No Default

7.7.2. @selectivity[label].type=all_values_bounded

l L
Type: non-negative integer
Default: No Default

h H
Type: non-negative integer
Default: No Default

v V
Type: estimable vector
Default: No Default

7.7.3. @selectivity[label].type=constant

c C
Type: estimable
Default: No Default

7.7.4. @selectivity[label].type=double_exponential

x0 X0
Type: estimable
Default: No Default

x1 X1
Type: constant
Default: No Default

x2 X2
Type: constant
Default: No Default

y0 Y0
Type: estimable
Default: No Default

y1 Y1
Type: estimable
Default: No Default

y2 Y2
Type: estimable
Default: No Default

alpha Alpha
Type: estimable
Default: 1.0

7.7.5. @selectivity[label].type=double normal

mu Mu
Type: estimable
Default: No Default

sigma_l Sigma L
Type: estimable
Default: No Default

sigma_r Sigma R
Type: estimable
Default: No Default

alpha Alpha
Type: estimable
Default: 1.0

7.7.6. @selectivity[label].type=increasing

l Low
Type: non-negative integer
Default: No Default

h High
Type: non-negative integer
Default: No Default

v V
Type: estimable vector
Default: No Default

alpha Alpha
Type: constant
Default: 1.0

7.7.7. @selectivity[label].type=inverse_logistic

a50 A50
Type: estimable
Default: No Default

ato95 aTo95
Type: estimable
Default: No Default

alpha Alpha
Type: estimable
Default: 1.0

7.7.8. @selectivity[label].type=knife_edge

e Edge
Type: estimable
Default: No Default

alpha Alpha
Type: estimable
Default: 1.0

7.7.9. @selectivity[label].type=logistic

a50 A50
Type: estimable
Default: No Default

ato95 Ato95
Type: estimable
Default: No Default

alpha Alpha
Type: estimable
Default: 1.0

7.7.10. @selectivity[label].type=logistic_producing

l Low
Type: non-negative integer
Default: No Default

h High
Type: non-negative integer
Default: No Default

a50 A50
Type: estimable
Default: No Default

ato95 Ato95
Type: estimable
Default: No Default

alpha Alpha
Type: estimable
Default: 1.0

7.8. Layers

@layer *label* Define an object of type *layer*

label Label of the Layer
Type: string
Default: No Default

type Type of Layer
Type: string
Default: No Default

7.8.1. @layer[label].type=int_layer**7.8.2. @layer[label].type=numeric_layer**

An example of how to specify a table in the layers.

```
table layer
1
end_table
```

8. Observation command and subcommand syntax

8.1. Observation types

The observation types available are,

Observations of proportions of individuals by age class

Observations of proportions of individuals between categories within each age class

Relative and absolute abundance observations

Relative and absolute biomass observations

Each type of observation requires a set of subcommands and arguments specific to that process.

@observation *label* Define an object of type *observation*

label Label

Type: string

Default: No Default

type Type of observation

Type: string

Default: No Default

likelihood Type of likelihood to use

Type: string

Default: No Default

categories Category labels to use

Type: string vector

Default: true

delta Robustification value (delta) for the likelihood

Type: constant

Default: DELTA

simulation_likelihood Simulation likelihood to use

Type: string

Default: ""

likelihood_multiplier Likelihood score multiplier

Type: constant

Default: Double(1.0)

error_value_multiplier Error value multiplier for likelihood

Type: constant

Default: Double(1.0)

8.2. Likelihoods

@likelihood *label* Define an object of type *likelihood*

8.2.1. @likelihood[label].type=binomial

8.2.2. @likelihood[label].type=binomial approx

8.2.3. @likelihood[label].type=dirichlet

8.2.4. @likelihood[label].type=log_normal

8.2.5. @likelihood[label].type=log_normal_with_q

8.2.6. @likelihood[label].type=multinomial

8.2.7. @likelihood[label].type=normal

8.2.8. @likelihood[label].type=pseudo

9. Report command and subcommand syntax

9.1. Report commands and subcommands

@report *label* Define an object of type *report*

label The label for the report

Type: string

Default: No Default

type The type of report

Type: string

Default: No Default

file_name The File Name if you want this report to be in a seperate file

Type: string

Default: ""

write_mode The write mode

Type: string

Default: overwrite

Allowed Values: overwrite, append, incremental_suffix

9.1.1. `@report[label].type=derived_quantity`

9.1.2. `@report[label].type=initialisation_partition`

9.1.3. `@report[label].type=process`

`process` Process label that is reported
 Type: string
 Default: ""

9.1.4. `@report[label].type=standard_header`

9.1.5. `@report[label].type=summarise_agents`

`years` Years
 Type: non-negative integer vector
 Default: true

`time_step` Time Step label
 Type: string
 Default: No Default

`number_of_agents` Number of agents to summarise
 Type: non-negative integer
 Default: No Default
 Lower Bound: 1 (inclusive)

9.1.6. `@report[label].type=world_age_frequency`

`years` Years
 Type: non-negative integer vector
 Default: true

`time_step` Time Step label
 Type: string
 Default: ""

10. Including commands from other files

`@include file` Include an external file

`file` The name of the external file to include

Type: string

Default: No default

Value: A valid external file

Condition: The file name must be enclosed in double quotes

Example: `!include "my_file.ibm"`

Note: `!include` does not denote the end of the previous command block as is the case for all other commands

11. Syntax conventions, examples and niceties

11.1. Input File Specification

The file format used for THE IBM is based on the formats used for Casal2, CASAL and SPM. It's a standard text file that contains definitions organised into blocks.

Without exception, every object specified in a configuration file is part of a block. At the top level blocks have a one-to-one relationships with components in the system.

Some general notes about writing configuration files:

1. Whitespace can be used freely. Tabs and spaces are both accepted
2. A block ends only at the beginning of a new block or end of final configuration file
3. You can include another configuration file from anywhere
4. Included files are placed inline, so you can continue a block in a new file
5. The configuration files support inline declarations of objects

11.1.1. Keywords And Reserved Characters

In order to allow efficient creation of input files CASAL2's file format contains special keywords and characters that cannot be used for labels etc.

@Block Definitions

Every new block in the configuration file must start with a block definition character. The reserved character for this is the @ character

Example:

```
@block1 <label>  
type <type>
```

```
@block2 <label>  
type <type>
```

'type' Keyword

The 'type' keyword is used for declaring the sub-type of a defined block. Any block object that has multiple sub-types will use the type keyword.

Example:

```
@block1 <label>  
type <sub_type>
```

```
@block2 <label>  
type <sub_type>
```

(Single-Line Comment)

Comments are supported in the configuration file in either single-line (to end-of-line) or multi-line
Example:

```
@block <label>
type <sub_type> #Descriptive comment
#parameter <value_1> This whole line is commented out
parameter <value_1> #<value_2>(value_2 is commented out)
```

/* */ (Multi-Line Comment)

Multiple line comments are supported by surrounding the comments in /* and */
Example:

```
@block <label>
type <sub_type>
parameter <value_1>
parameter <value_1> <value_2>

\*
Do not load this process
@block <label>
type <sub_type>
parameter <value_1>
parameter <value_1> <value_2>
*\
```

{ } (Indexing Parameters)

Users can reference individual elements of a map using the { } syntax, for example when estimating `ycs_values` you may only want to estimate a block of YCS not all of them say between 1975 and 2012. Example:

```
@estimate YCS
parameter process[Recruitment].ycs_values{1975:2012}
type uniform
lower_bound
upper_bound
```

':' (Range Specifier)

The range specifier allows you to specify a range of values at once instead of having to input them manually. Ranges can be either incremental or decremental.
Example:

```
@process my_recruitment_process
type constant_recruitment
years_to_run 1999:2009 #With range specifier

@process my_mortality_process
type natural_mortality
years_to_run 2000 2001 2002 2003 2004 2005 2006 2007 #Without range specifier
```

'table' and 'end_table' Keyword

The table keyword is used to define a table of information used as a parameter. The line following the table declaration must contain a list of columns to be used. Following lines are rows of the table. Each row must have the same number of values as the number of columns specified. The table definition must end with the 'end_table' keyword on it's own line. The first row of a table will be the name of the columns if required.

Example:

```
@block <label>
type <sub_type>
parameter <value_1>
table <table_label>
<column_1> <column_2> <column_n>
<row1_value1> <row1_value2> <row1_valueN>
<row2_value1> <row2_value2> <row2_valueN>
end_table
```

[] (Inline Declarations)

When an object takes the label of a target object as a parameter this can be replaced with an inline declaration. An inline declaration is a complete declaration of an object one line. This is designed to allow the configuration writer to simplify the configuration writing process.

Example:

```
#With inline declaration with label specified for time step
@model
time_steps step_one=[type=iterative; processes=recruitment ageing]

#With inline declaration with default label (model.1)
@model
time_steps [type=iterative; processes=recruitment ageing]

#Without inline declaration
@model
time_steps step_one

@time_step step_one
processes recruitment ageing
```

Parameters

THE IBM also allows parameters that are of type vector or map to be referenced and estimated partially. An example of a parameter that is type vector is `yces_values` in a recruitment process. Let say a recruitment block was specified as follows,

```
@process WestRecruitment
type recruitment_beverton_holt
r0 400000
years
yses_values 1 1 1 1 1 1 1
yses_years 1975:1983
```

An alternative specification to the sequence of values you can use an astrix to shorthand repeating integers e.g.

```
yces_values 1*8
```

```
steepness 0.9
```

```
age 1
```

Lets say we wanted to only estimate the last four years of the parameter `process[WestRecruitment].yces_values`. This can be done as specified in the following `@estimate` block,

```
@estimate
parameter process[WestRecruitment].yces_values{1979:1983}
type uniform
lower_bound 0.1 0.1 0.1 0.1
upper_bound 10 10 10 10
```

Note the first element of a vector is indexed by 1. This syntax can be applied to parameters that are of type `map` as well, for information on what type a parameter is see the syntax section. An example of a parameter that is of type `map` is `@time_varying[label].type=constant`. For the following `@time_varying` block,

```
@time_varying q_step1
type constant
parameter catchability[Fishq].q
years 1992 1993 1994 1995
value 0.2 0.2 0.2 0.2
```

In this example a user may want to estimate only one element of the map (say 1992), but force all other years to be the same as the one estimate. This can be done in an estimate block as follows,

```
@estimate
parameter time_varying[q_step1].value{1992}
same time_varying[q_step1].value{1993:1995}
type uniform
lower_bound 0.1 0.1 0.1 0.1
upper_bound 10 10 10 10
```

In line declaration

In line declarations can help shorten models by passing `@` blocks, for example

```
@observation chatCPUE
type biomass
catchability [q=6.52606e-005]
time_step one
categories male+female
selectivities chatFselMale chatFselFemale
likelihood lognormal
years 1992:2001
time_step_proportion 1.0
obs 1.50 1.10 0.93 1.33 1.53 0.90 0.68 0.75 0.57 1.23
error_value 0.35

@estimate
```

```
parameter catchability[chatTANbiomass.one].q
type uniform_log
lower_bound 1e-2
upper_bound 1
In line declaration tips
```

In the above code we are defining and estimating catchability without explicitly creating an `@catchability` block.

When you do an inline declaration the new object will be created with the name of the creator's `label.index` where `index` will be the word if it's one-nine and the number if it's 10+, for example,

```
@mortality halfm
selectivities [type=constant; c=1]
```

```
would create
@selectivity halfm.one
```

if there were 10 categories all with there own selectivity the 10th selectivity would be labelled,

```
@selectivity halfm.10
```

11.2. Processes

Processes are special in how they can be defined, all throughout this document we have been referring to specifying a process as follows,

```
@process Recruitment
type recruitment_beverton_holt
```

However for convenience and for file clarity you could equally specify this block as follows,

```
@recruitment Recruitment
type beverton_holt
```

The trick is that you can replace the keyword `process` with the first word of the process type, in the example above this is the `recruitment` this can be away of creating more reader friendly/lay term configuration scripts. More examples follow;

```
@mortality Fishing_and_M
type instantaneous
```

```
@transition Migration
type category
```


12. Post processing output using R

13. Troubleshooting

13.1. Introduction

This section is to aid users in debugging models, if you cannot resolve an issue using these guidelines then don't hesitate to contact the development team. To report an issue please follow the format described in Section 13.3. We are hoping that most user errors will be well documented and that THE IBM will produce informative error messages. In the case where this doesn't happen, there are some quick and easy tactics that users can do to attempt to resolve or at least isolate an error/bug. Using THE IBM's internal logging out system, this is invoked at the command line with by the `--loglevel` parameter followed by one of these arguments; `trace`, `finest`, `fine`, `medium`. An example of implementing logging with trace level at the command line is,

```
ibm -r --loglevel trace > output.log 2> log.out
```

The above command will output THE IBM normal reports into the file "output.log" where as the `2>` syntax will print the error logged out information into the file "log.out". You should be able to see where THE IBM is exiting by going to the end of the "log.out" file.

```
LOG_FINE() << "Model: State change to Execute";
```

taken from line 379.

13.2. Reporting errors

If you find a bug or problem in THE IBM, please email the IBM Development Team submit an issue on the github repository found at <https://github.com/Craig44/IBM/issues>. The latter is preferred as it will automatically document the issue which is better than depending on the development team, who may be forgetful. Please follow the guidelines below, as they will enhance the debugging process which can be quite time consuming.

13.3. Guidelines for reporting a problem with THE IBM

1. Check to ensure you are using the most recent version of THE IBM. Its possible that the error or problem you are having may have ready been resolved.
2. Describe the version of THE IBM are you using? e.g., "THE IBM v2018-07-17 (rev. cd12b61) Microsoft Windows executable". The version is provided by THE IBM with the following command `ibm -v`.
3. What operating system or environment are you using? e.g., "IBM-PC Intel CPU running Microsoft Windows 10 Enterprise".
4. Give a brief one-line description of the problem, e.g., "a segmentation fault was reported".
5. If the problem is reproducible, please list the exact steps required to cause it, remembering to include the relevant THE IBM configuration file, other input files, and any out generated. Specify the *exact* command line arguments that were used, e.g., "Using the command `***. -*` reports a segmentation fault. The input configuration files are attached."
6. If the problem is not reproducible (only happened once, or occasionally for no apparent reason), please describe the circumstances in which it occurred and the symptoms observed (but note it is much harder to reproduce and hence fix non-reproducible bugs, but if several

reports are made over time that relate to the same thing, then this may help to track down the problem), e.g., “THE IBM crashed, but I cannot reproduce how I did it. It seemed to be related to a local network crash but I cannot be sure.”

7. If the problem causes any error messages to appear, please give the *exact* text displayed, e.g.,
segmentation fault (core dumped).
8. Remember to attach all relevant input and output files so that the problem can be reproduced (it can be helpful to compress these into a single file e.g. zip file). Without these, it is usually not possible to determine the cause of the problem, and we are unlikely to provide any assistance. Note that it is helpful to be as specific as possible when describing the problem.

14. THE IBM software license

GNU GENERAL PUBLIC LICENSE

Copyright © 1989, 1991 Free Software Foundation, Inc.

51 Franklin Street, Fifth Floor, Boston, MA 02110-1301, USA

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

Preamble

The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change free software—to make sure the software is free for all its users. This General Public License applies to most of the Free Software Foundation's software and to any other program whose authors commit to using it. (Some other Free Software Foundation software is covered by the GNU Library General Public License instead.) You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things.

To protect your rights, we need to make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the software, or if you modify it.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must give the recipients all the rights that you have. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

We protect your rights with two steps: (1) copyright the software, and (2) offer you this license which gives you legal permission to copy, distribute and/or modify the software.

Also, for each author's protection and ours, we want to make certain that everyone understands that there is no warranty for this free software. If the software is modified by someone else and passed on, we want its recipients to know that what they have is not the original, so that any problems introduced by others will not reflect on the original authors' reputations.

Finally, any free program is threatened constantly by software patents. We wish to avoid the danger that redistributors of a free program will individually obtain patent licenses, in effect making the program proprietary. To prevent this, we have made it clear that any patent must be licensed for everyone's free use or not licensed at all.

The precise terms and conditions for copying, distribution and modification follow.

TERMS AND CONDITIONS FOR COPYING, DISTRIBUTION AND MODIFICATION

0. This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License.

The “Program”, below, refers to any such program or work, and a “work based on the Program” means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term “modification”.) Each licensee is addressed as “you”.

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running the Program is not restricted, and the output from the Program is covered only if its contents constitute a work based on the Program (independent of having been made by running the Program). Whether that is true depends on what the Program does.

1. You may copy and distribute verbatim copies of the Program’s source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and give any other recipients of the Program a copy of this License along with the Program.

You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2. You may modify your copy or copies of the Program or any portion of it, thus forming a work based on the Program, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:
 - a) You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change.
 - b) You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.
 - c) If the modified program normally reads commands interactively when run, you must cause it, when started running for such interactive use in the most ordinary way, to print or display an announcement including an appropriate copyright notice and a notice that there is no warranty (or else, saying that you provide a warranty) and that users may redistribute the program under these conditions, and telling the user how to view a copy of this License. (Exception: if the Program itself is interactive but does not normally print such an announcement, your work based on the Program is not required to print an announcement.)

These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Program.

In addition, mere aggregation of another work not based on the Program with the Program (or with a work based on the Program) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3. You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:
 - a) Accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,
 - b) Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine-readable copy of the corresponding source code, to be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,
 - c) Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable. However, as a special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

If distribution of executable or object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place counts as distribution of the source code, even though third parties are not compelled to copy the source along with the object code.

4. You may not copy, modify, sublicense, or distribute the Program except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense or distribute the Program is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.
5. You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Program or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Program (or any work based on the Program), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Program or works based on it.
6. Each time you redistribute the Program (or any work based on the Program), the recipient automatically receives a license from the original licensor to copy, distribute or modify the Program subject to these terms and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.
7. If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you

from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Program at all. For example, if a patent license would not permit royalty-free redistribution of the Program by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Program.

If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply and the section as a whole is intended to apply in other circumstances.

It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system, which is implemented by public license practices. Many people have made generous contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to distribute software through any other system and a licensee cannot impose that choice.

This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

8. If the distribution and/or use of the Program is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Program under this License may add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.
9. The Free Software Foundation may publish revised and/or new versions of the General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Program specifies a version number of this License which applies to it and “any later version”, you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of this License, you may choose any version ever published by the Free Software Foundation.

10. If you wish to incorporate parts of the Program into other free programs whose distribution conditions are different, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

NO WARRANTY

11. BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE

PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

12. IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MAY MODIFY AND/OR REDISTRIBUTE THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

15. Acknowledgements

We thank the developers of Casal2 (Rasmussen et al., 2016), CASAL (Bull et al., 2012) and (Dunn et al., 2015) for their ideas that led to the development of THE IBM.

Much of the structure of THE IBM, equations, and documentation in this manual draw heavily on similar components of the fisheries population model Casal2 (Rasmussen et al., 2016), CASAL (Bull et al., 2012) and the spatial model SPM (Dunn et al., 2015). We thank the authors of Casal2, CASAL and SPM for their permission to use their work as the basis for parts of THE IBM and allow the use of the definitions, concepts, and documentation.

16. References

- B Bull, R I C C Francis, A. Dunn, A McKenzie, D J Gilbert, M H Smith, R Bian, and D Fu. CASAL C++ Algorithmic Stock Assessment Laboratory): CASAL user manual v2.30-2012/03/21. Technical Report 135, National Institute of Water and Atmospheric Research Ltd (NIWA), 2012.
- A. Dunn, S. Rasmussen, and S. Mormede. Spatial population model user manual, spm v1.1-2016-03-04 (rev. 1278). Technical Report 138, National Institute of Water and Atmospheric Research Ltd (NIWA), 2015.
- Makoto Matsumoto and Takuji Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation: Special Issue on Uniform Random Number Generation*, 8(1):3–30, January 1998.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014. URL <http://www.R-project.org/>.
- S. Rasmussen, I. Doonan, A. Dunn, C. Marsh, K. Large, and S. Mormede. Casal2 user manual. Technical Report 139, National Institute of Water and Atmospheric Research Ltd (NIWA), 2016.

Appendices

A. An Appendix

17. Index

- Ageing, 14
- Annual cycle, 11, 12
- Beverton-Holt recruitment, 14
- BOOST C++ library, 2
- Command
 - Include files, 9
- Command block format, 8
- Command line arguments, 6
- Commands, 7
- Commands
 - Subcommands, 7
- Commenting out lines, 9
- Comments, 9
- Constant recruitment, 14
- Derived Quantities, 15
- Derived quantities, 3
- Determining parameter names, 9
- Disk space, 1
- Exit status value, 10
- gcc, 2
- Getting help, 2
- github, 1
- Growth, 15
- In line declaration, 26
- Including external files, 5
- Initialisation, 11–13
- Input configuration file, 5
- Input configuration file syntax, 7
- Length-weight relationship, 15
- Linux, 1, 2, 5
- Microsoft Windows, 1, 2, 5
- Mingw, 2
- Model
 - derived quantities, 3
 - initialisation, 11
 - partition, 3
 - processes, 3
 - state, 3
 - time-steps, 3
- Model overview, 3
- Model run years, 12
- Mortality, 14
- Mortality blocks, 12
- Necessary files, 1
- Notifying errors, 2
- Observations, 19
- Optional command line arguments, 6
- Output header information, 6
- Parameter names, 9
- Parameters, 25
- Partition, 3
- Population processes, 14
- Population section, 11
- Population structure, 11
- Post processing, 29
- Post processing output using **R**, 29
- Post processing section, 29
- Processes, 3
- Random number generator, 2
- Recruitment, 14
- Redirecting standard error, 5
- Redirecting standard out, 5
- Redirecting standard output, 5
- Reports, 21
- Reports section, 21
- Running THE IBM, 5
- Selectivities, 15
 - All-values, 16
 - All-values-bounded, 17
 - Double-exponential, 18
 - Double-normal, 18
 - Increasing, 17
 - Inverse-logistic, 17
 - Logistic, 17
 - Logistic-producing, 17
- standard error, 5
- standard output, 5
- State, 3
- Subcommand argument type, 8
- Syntax conventions, examples and niceties, 23
- System requirements, 1
- Technical specifications, 2
- The estimation section, 3
- The observation section, 3
- The population section, 3, 11

The report section, 3, 21
The state object and the partition, 12
Time sequences, 12
time-steps, 3

User assistance, 2
Using THE IBM, 5

Where to get THE IBM, 1

