

# Package: AgeingError (via r-universe)

September 17, 2024

**Type** Package

**Title** Estimating ageing error with 'TMB' from double reads

**Version** 2.0.2

**Description** Implements Generalized Linear Mixed Effect Models (GLMMs) using 'TMB' to estimate ageing error from double reads of otoliths. The original analysis (Punt et al. 2008; <doi:10.1139/F08-111>) was written in `ADMB` and focused on estimating imprecision and bias. The newer version of the software allows for random effects using `TMB`.

**License** GPL-3

**URL** <https://pfmc-assessments.github.io/AgeingError/>

**BugReports** <https://github.com/pfmc-assessments/AgeingError/issues>

**Depends** R (>= 3.5.0)

**Imports** cli, fs, knitr, stats, TMB (>= 1.8.0)

**Suggests** spelling

**LinkingTo** RcppEigen, TMB

**VignetteBuilder** knitr

**ByteCompile** true

**Config/testthat/edition** 3

**Config/testthat/parallel** true

**Copyright** inst/COPYRIGHTS

**Encoding** UTF-8

**Language** en-US

**LazyData** true

**Roxygen** list(markdown = TRUE)

**RoxygenNote** 7.3.2

**SystemRequirements** GNU make, C++17

**Repository** <https://noaa-fisheries-integrated-toolbox.r-universe.dev>

**RemoteUrl** <https://github.com/pfmc-assessments/AgeingError>

**RemoteRef** HEAD

**RemoteSha** 967a6fa4b90e087095df42d862f82fdcb2548858

## Contents

ageing_comparison . . . . .	2
cMx . . . . .	3
CreateData . . . . .	4
CreateSpecs . . . . .	5
determine_n_sets . . . . .	5
DoApplyAgeError . . . . .	6
estgrowth.vb . . . . .	7
fit . . . . .	8
minimizer . . . . .	8
PlotOutputFn . . . . .	9
plot_output . . . . .	11
ProcessResults . . . . .	12
rMx . . . . .	13
run . . . . .	13
RunFn . . . . .	14
SimulatorFn . . . . .	19
StepwiseFn . . . . .	21
<b>Index</b>	<b>26</b>

---

ageing_comparison	<i>Plot comparison of double age readings</i>
-------------------	---

---

## Description

Plot with circles proportional to how many double readings fell in each pair of coordinates

## Usage

```
ageing_comparison(
  xvec,
  yvec,
  scale.pts = 2,
  col.pts = grDevices::grey(0.1, alpha = 0.5),
  col.hist = grDevices::rgb(0, 0, 0.5, alpha = 0.7),
  counts = TRUE,
  maxage = NULL,
  hist = TRUE,
  hist.frac = 0.1,
  xlab = "Age reader A",
```

```

    ylab = "Age reader B",
    title = NULL,
    png = FALSE,
    filename = "ageing_comparison.png",
    SaveFile = NULL,
    verbose = TRUE
)

```

### Arguments

xvec	vector of values from reader A
yvec	vector of values from reader B
scale.pts	Documentation needed.
col.pts	color for points
col.hist	color for histograms
counts	include text within each bubble showing count of values?
maxage	maximum age to include in the plot (doesn't yet work well)
hist	include a histogram along each axis?
hist.frac	maximum value of histograms as fraction of maxage
xlab	label for xvec
ylab	label for yvec
title	Optional title to add at top of plot
png	Save plot to PNG file?
filename	File name for PNG file.
SaveFile	directory where plot will be saved. NULL value will make it go to working directory.
verbose	Report messages as function runs.

### Author(s)

Ian G. Taylor

---

cMx

*Make a column matrix*

---

### Description

**The function is currently defined as:**

function (Input)

as.matrix(Input)

**Usage**

cMx(Input)

**Arguments**

Input           input to be converted to a matrix

**Author(s)**

James T. Thorson

---

CreateData           *Read the ageing error data*

---

**Description**

Read the ageing error data

**Usage**

```
CreateData(DataFile = "data.dat", NDataSet = 1, verbose = FALSE, EchoFile = "")
```

**Arguments**

DataFile       Filename for input data

NDataSet       Number of data sets within DataFile

verbose        Return messages to the console (in addition to any output to EchoFile)

EchoFile       A file path to a file that will be created or appended to if it already exists to store information about your data inputs. The default is '', which leads to output being printed to the screen rather than saved in a file. An example of a user-defined input would be 'EchoTMB.out'.

**Author(s)**

Andre E. Punt

---

CreateSpecs	<i>Read the ageing error specifications</i>
-------------	---

---

**Description**

Read the ageing error specifications

**Usage**

```
CreateSpecs(SpecsFile = "data.spc", DataSpecs, verbose = FALSE)
```

**Arguments**

SpecsFile	Filename for input specifications.
DataSpecs	The output from CreateData()
verbose	Return messages to the console (TRUE/FALSE)

**Author(s)**

Andre E. Punt

---

determine_n_sets	<i>Determine the number of data sets in a data file</i>
------------------	---

---

**Description**

Determine the number of data sets in a data file

**Usage**

```
determine_n_sets(file)
```

**Arguments**

file	A file path to a data file.
------	-----------------------------

**Value**

An integer giving the number of data sets in the file.

**Author(s)**

Kelli F. Johnson

---

DoApplyAgeError      *Run the ageing error optimization routine*

---

### Description

Run the ageing error optimization routine

### Usage

```
DoApplyAgeError(  
  Species = "AgeingError",  
  DataSpecs,  
  ModelSpecsInp,  
  AprobWght = 1e-06,  
  SlopeWght = 0.01,  
  SaveDir = getwd(),  
  verbose = FALSE  
)
```

### Arguments

Species	A string that will be used to create file names. Typically, users will use the common name for the species of interest, especially if you are saving files from multiple species in a single directory. Though, the default is "AgeingError".
DataSpecs	A data object returned from <a href="#">CreateData()</a> .
ModelSpecsInp	A specification object returned from <a href="#">CreateSpecs()</a> .
AprobWght, SlopeWght	Numeric values passed to the model. The defaults are 1e-06 and 0.01. Andre originally had these hard coded from his workspace. TODO: decide if they should be passed in the specifications or data files.
SaveDir	A path, relative or absolute, to a directory where the results will be saved. The directory need not exist currently as it will be created dynamically.
verbose	A logical specifying if messages should be printed. The default is to <b>NOT</b> print, i.e., verbose = FALSE.

### Author(s)

Andre E. Punt

estgrowth.vb

*Calculate von Bertalanffy Growth Parameters from Lengths and Ages***Description**

Calculate von Bertalanffy growth parameters from length and age data or predicted lengths given ages and input parameters.

**Usage**

```
estgrowth.vb(Par, Ages, Lengths, Returntype = c("NLL", "Pred"), sdFactor = 1)
```

**Arguments**

Par	A list of von Bertalanffy growth parameters in log space ordered as follows: K, Linf, L0, CV0, and CV1. Names will be assigned if they are not provided.
Ages	A vector of ages in years. Values of NA are accepted.
Lengths	A vector of Lengths in cm. Lengths can be NULL if Returntype == "Pred" because you are only predicting using ages, where the lengths are just needed for estimation purposes. If not NULL, ensure that there is one length measurement for every age measurement. Values of NA are accepted.
Returntype	A single character value with "NLL" being the default, which leads to the negative log-likelihood value being returned. If "Pred", then three values are returned for each combination of length and age, low, prediction, and high based on the input parameters and standard deviation factor, i.e., sdFactor.
sdFactor	The number of standard deviations to include in the low and high calculations. The default is 1.0.

**Value**

Depending on Returntype, either the negative log likelihood is returned based on fits to the data or a matrix of three columns with low, predicted, and high values for each combination of length and age. Distance of the low and high from the predicted value depends on the sdFactor, allowing confidence intervals based on normal theory or other theories to be created.

**Examples**

```
## Not run:
bio_dat <- data.frame(
  Age = rep(0:30, each = 20),
  Length_cm = rnorm(n = 31 * 20, mean = 50, sd = 5)
)
pars_in <- lapply(FUN = log, X = list(
  "K" = 0.13,
  "Linf" = 55,
  "L0" = 5,
  "CV0" = 0.1,
```

```

    "CV1" = 0.1
  ))
  solve <- optim(
    fn = estgrowth.vb, par = unlist(pars_in), hessian = FALSE,
    Ages = bio_dat[, "Age"],
    Lengths = bio_dat[, "Length_cm"]
  )
  predictions <- estgrowth.vb(
    Par = solve$par, Return Type = "Pred",
    sdFactor = 1,
    Ages = bio_dat[, "Age"],
    Lengths = bio_dat[, "Length_cm"]
  )
  plot(bio_dat$Age, predictions[, "Lhat_pred"],
    xlab = "Age (years)", ylab = "Predicted length (cm)"
  )
  exp(solve$par)

## End(Not run)

```

---

fit

*Fit an ageing error model using TMB*


---

### Description

Fit an ageing error model using TMB

### Usage

```
fit()
```

### Author(s)

Kelli F. Johnson

---

minimizer

*Minimize the negative log likelihood*


---

### Description

Minimize the negative log likelihood using "nlmimb" and/or "optim".



**Usage**

```

minimizer(
  model,
  method = c("optim", "nlmimb", "both"),
  lower,
  upper,
  verbose = FALSE
)

```

**Arguments**

model	A model to be optimized.
method	A string specifying the desired method to be used for the optimization routine. The options are listed in the function call, where the default is to use "optim". Using both routines is an option, via "both", and will lead to first optimizing the model using "nlminb" and then re-optimization of the model with "optim". Note that when using <code>stats::optim()</code> , the "L-BFGS-B" method is used rather than the default method of "Nelder-Mead".
lower, upper	Vectors of parameter bounds of the same length as the number of parameters in the model.
verbose	A logical specifying if messages should be printed. The default is to <b>NOT</b> print, i.e., verbose = FALSE.

**Author(s)**

Andre E. Punt

---

PlotOutputFn

*Plot output*

---

**Description**

Plots age comparisons and results from the fitted model. Comparisons must be conditioned on a true age that is not observed. And, in place of a true age, the diagnostic plots generally condition on an estimated age, which is fixed as the mode of the conditional probability at age for each otolith.

**Usage**

```

PlotOutputFn(
  Data,
  MaxAge,
  SaveFile,
  PlotType = c("PNG", "PDF"),
  subplot = 1:3,
  ReaderNames = NULL,
  ...
)

```

**Arguments**

Data	This is the data set with the first column being an integer providing the number of otoliths that are included in the row and the subsequent columns are the reader or lab estimated age where each reader/lab has a unique reading error and bias. The modeling framework allows for, at most, 15 readers, i.e., 16 columns. There should not be any identical rows in the data frame because otoliths that have the exact same read from every reader/lab should be combined into a single row with the count as the first column. If you failed to combine identical rows prior to running the model, you will be alerted with an error and the XXX.rep file will have a properly formatted data which can be cut-pasted into a XXX.dat file for use. Missing reads from a given reader/lab should be entered as -999. Order your reader/lab columns such that similar readers/labs are located next to one another because columns to the right can mirror columns to their immediate left in terms of parameter estimates.
MaxAge	An integer, specifying the maximum possible "true" age.
SaveFile	Directory where <code>agemat.exe</code> is located and where all ADMB intermediate and output files should be located. If <code>AdmbFile</code> is specified then <code>agemat.exe</code> is copied from that directory to <code>SaveFile</code> .
PlotType	A string specifying the type of saved plots that you desire. The default is to save .png files via an argument of "PNG". The other option is to save .pdf files via "PDF".
subplot	Vector of integers specifying which plots to create. The default is to create three plots.
ReaderNames	Vector with names of each reader, defaults to "Reader 1", "Reader 2", etc.
...	Additional arguments passed to <code>ageing_comparison()</code> .

**Details**

1. Error and bias by reader/lab: A panel graph is provided where each panel shows the expected and standard deviation in age reads for that reader/lab. This is displayed against a scatter plot of the read and estimated ages for each otolith that was read by that reader/lab.
2. Proportion-at-age histogram: The estimated proportion at age can be plotted as a histogram and is displayed against the observed distribution of read ages. This is useful to determine if the estimated proportion at age is generally plausible, e.g., whether it has too many ages where the estimated proportion at age approaches zero, which is unlikely in a composite sample with moderate effective sample sizes. This plot can also be used as a diagnostic to confirm that AIC has selected reasonable values for the `MinusAge` and `PlusAge` parameters.

The function will read in XXX.rep and XXX.par files that are located in `SaveFile`.

**Value**

Returns AIC, AICc, and BIC for fitted model.

**Author(s)**

James T. Thorson, Ian G. Taylor

**References**

Punt, A.E., Smith, D.C., KrusicGolub, K., and Robertson, S. 2008. Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australias southern and eastern scalefish and shark fishery. Can. J. Fish. Aquat. Sci. 65: 1991-2005.

**See Also**

- RunFn()
- StepwiseFn()

---

plot\_output

*Plot output*


---

**Description**

Plots age comparisons and results from the fitted Ageing Error model

**Usage**

```
plot_output(
  Data,
  IDataset,
  MaxAge,
  Report,
  subplot = 1:3,
  Nparameters = 0,
  LogLike = 0,
  ReaderNames = NULL,
  Species = "AgeingError",
  SaveDir = getwd(),
  verbose = FALSE,
  ...
)
```

**Arguments**

Data	Input data matrix
IDataset	Index of the data set used in creating the filename
MaxAge	Maximum estimated age
Report	Results from fitting the model
subplot	Vector of which plots to create.
Nparameters	Number of parameters
LogLike	Negative log likelihood from fitting the model

ReaderNames	Vector with names of each reader, defaults to 'Reader1', 'Reader2', etc. if left at the default argument of NULL. If you pass a vector of strings, the vector must be the same length as <code>NCOL(Data) - 1</code> .
Species	String used at beginning of the output files
SaveDir	Directory for fitted model
verbose	Report messages as function runs.
...	Additional arguments passed to <code>ageing_comparison()</code> .

**Value**

Returns AIC, AICc, and BIC for fitted model.

**Author(s)**

James T. Thorson, Ian G. Taylor

**References**

Punt, A.E., Smith, D.C., KrusicGolub, K., and Robertson, S. 2008. Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australias southern and eastern scalefish and shark fishery. *Can. J. Fish. Aquat. Sci.* 65: 1991-2005.

---

ProcessResults

*Process results of the ageing error estimation*

---

**Description**

Process results of the ageing error estimation

**Usage**

```
ProcessResults(
  Species = "AgeingError",
  SaveDir = getwd(),
  CalcEff = FALSE,
  verbose = FALSE
)
```

**Arguments**

Species	A string that will be used to create file names. Typically, users will use the common name for the species of interest, especially if you are saving files from multiple species in a single directory. Though, the default is "AgeingError".
SaveDir	A path, relative or absolute, to a directory where the results will be saved. The directory need not exist currently as it will be created dynamically.
CalcEff	Calculate effective sample sizes (TRUE/FALSE)
verbose	A logical specifying if messages should be printed. The default is to <b>NOT</b> print, i.e., <code>verbose = FALSE</code> .

**Author(s)**

Andre E. Punt

---

`rMx`*Make a row matrix*

---

**Description****The function is currently defined as:**

```
function (Input)
if (is.vector(Input)) Output <- t(as.matrix(Input))
if (!is.vector(Input)) Output <- as.matrix(Input)
Output
```

**Usage**`rMx(Input)`**Arguments**

Input	input to be converted into a row matrix
-------	---

**Author(s)**

James T. Thorson

---

`run`*Run ageing error routine*

---

**Description**

A wrapper for running a TMB model to estimate ageing error for a given data set and specification file.

**Usage**`run(file_data, file_specs, directory)`**Arguments**

<code>file_data</code>	A string specifying the file path to a data file.
<code>file_specs</code>	A string specifying the file path to the specifications file.
<code>directory</code>	A string specifying a file path to a directory where you would like to save the results.

**Value**

Invisibly return model output.

**Author(s)**

Kelli F. Johnson

---

RunFn

*Run ageing error model*

---

**Description**

Run the Punt et al. (2008) ADMB-based ageing error model from within R.

**Usage**

```
RunFn(
  Data,
  SigOpt,
  KnotAges,
  BiasOpt,
  NDataSets = 1,
  MinAge,
  MaxAge,
  RefAge,
  MinusAge,
  PlusAge,
  MaxSd,
  MaxExpectedAge,
  SaveFile,
  EffSampleSize = 0,
  Intern = TRUE,
  AdmbFile = NULL,
  JustWrite = FALSE,
  CallType = "system",
  ExtraArgs = " -est",
  verbose = TRUE
)
```

**Arguments**

**Data** This is the data set with the first column being an integer providing the number of otoliths that are included in the row and the subsequent columns are the reader or lab estimated age where each reader/lab has a unique reading error and bias. The modeling framework allows for, at most, 15 readers, i.e., 16 columns. There should not be any identical rows in the data frame because otoliths that have the exact same read from every reader/lab should be combined into a single row

with the count as the first column. If you failed to combine identical rows prior to running the model, you will be alerted with an error and the XXX.rep file will have a properly formatted data which can be cut-pasted into a XXX.dat file for use. Missing reads from a given reader/lab should be entered as -999. Order your reader/lab columns such that similar readers/labs are located next to one another because columns to the right can mirror columns to their immediate left in terms of parameter estimates.

**SigOpt** This a vector with one entry for each reader (i.e.,  $\text{length}(\text{SigOpt}) == \text{NCOL}(\text{Data}) - 1$ ). Each entry specifies the functional form of reading error as a function of true age. Possible entries include the following:

**-0-9+** Mirror the standard deviation of another reader, where the negative integer corresponds to the column of the reader/lab that is being mirrored minus one, e.g., -1 causes it to mirror reader/lab 1, for which data is stored in the second column of Data. This number must be lower than -1 times the current position in the vector.

**0** No error. But, there could be potential bias.

**1** Constant coefficient of variation, i.e., a 1-parameter linear relationship of the standard deviation with the true age.

**2** Curvilinear standard deviation, i.e., a 3-parameter Hollings-form relationship of standard deviation with true age.

**3** Curvilinear coefficient of variation, i.e., a 3-parameter Hollings-form relationship of coefficient of variation with true age.

**5** Spline with estimated slope at beginning and end where the number of parameters is 2 + number of knots.

**6** Linear interpolation with a first knot of 1 and a last knot of the maximum age, i.e., MaxAge.

**KnotAges** Ages associated with each knot. This is a necessary input for SigOpt = 5 or SigOpt = 6.

**BiasOpt** A vector with one entry for each reader/lab specifying the type of bias specific to each reader. Positive values lead to estimated parameters and negative values are used for shared parameters between readers, just like with SigOpt. Parameter sharing is common when there is more than one reader in a lab working together to refine their methods such that they have matching techniques. Possible entries include the following:

**-0-9+** Mirror the bias of another reader, where the negative integer corresponds to the column of the reader/lab that is being mirrored minus one, e.g., -1 causes it to mirror reader/lab 1, for which data is stored in the second column of Data. This number must be lower than -1 times the current position in the vector.

**0** Unbiased, where at least one reader has to be unbiased.

**1** Constant coefficient of variation, i.e., a 1-parameter linear relationship of bias with true age.

**2** Curvilinear, i.e., a 2-parameter Hollings-form relationship of bias with true age.

An example entry for the situation where you have seven readers and you assume that the first reader is unbiased, readers 2-7 have a curvilinear bias, reader 3

shares parameters with reader 2, reader 5 shares parameters with reader 4, and reader 7 shares parameters with reader 6 would look like  $c(0, 2, -2, 2, -4, 2, -6)$ .

NDataSets	This is generally 1 and other values are not implemented.
MinAge	An integer, specifying the minimum possible "true" age.
MaxAge	An integer, specifying the maximum possible "true" age.
RefAge	An arbitrarily chosen age from which "true" age-composition fixed-effects are calculated as an offset. This has no effect on the answer but could potentially effect estimation speed.
MinusAge	The minimum age for which an age-specific age-composition is estimated. Ages below MinusAge have "true" proportion-at-age ( $P_a$ ) estimated as

$$P_a = P_{MinusAge} * exp(\beta * (MinusAge - a))$$

, where beta is an estimated log-linear trend in the "true" proportion-at-age. If MinusAge = MinAge, beta is not estimated.

PlusAge	Identical to MinusAge except defining the age above with age-specific age composition is not estimated.
MaxSd	An upper bound on possible values for the standard deviation of reading error.
MaxExpectedAge	Set to MaxAge.
SaveFile	Directory where <code>agemat.exe</code> is located and where all ADMB intermediate and output files should be located. If <code>AdmbFile</code> is specified then <code>agemat.exe</code> is copied from that directory to <code>SaveFile</code> .
EffSampleSize	Indicating whether effective sample size should be calculated. Missing values in the data matrix will cause this to be ineffective, in which case this should be set to 0.
Intern	A logical input that controls the amount of output displayed, where TRUE indicates that ADMB output should be displayed in R and FALSE leads to the suppression of this information.
AdmbFile	An optional character entry that specifies the directory from which <code>agemat.exe</code> is to be copied from to <code>SaveFile</code> .
JustWrite	A logical input that allows just the data files to be written without running ADMB executable.
CallType	Either "system" or "shell" depending on Operating System or how R is being run. The default is "system".
ExtraArgs	A string of characters providing extra arguments passed to ADMB. The default is "-est".
verbose	A logical input that controls the amount of feedback users receive from the program. The default is to provide the most output as possible with <code>verbose = TRUE</code> .



## Details

The premise of Punt *et al.* (2008) is to calculate the likelihood of model parameters given an observed data set of otolith age reads from multiple age readers. For each reader/lab, two parameters are defined, one for standard deviation and one for bias. The model calculates the expected age of each read and the standard deviation of a normally distributed reading error given the true age of an otolith. These relationships can be linear or curvilinear.

The true age is obviously an unobserved process and can be considered a random effect. Thus, the software computes the likelihood while summing across all possible discrete values for the true age of each otolith. This true age requires a hyperdistribution that represents the prior probability that an otolith is any given age. The hyperdistribution is controlled by a set of hyperparameters and the parameters that govern the standard deviation and bias of each age reader/lab. Specifically, one hyperparameter is estimated for every age between and including the MinusAge and PlusAge. Ages outside of this range have a prior proportion at age defined as a loglinear deviation from the proportion at age for the extreme ages, i.e., MinusAge and PlusAge. The slope of these loglinear deviations thus constitutes an additional 1 or 2 fixed effect parameters. The true proportion at age is then calculated from these fixed effects and loglinear slope parameters by normalizing the resulting distribution such that it sums to one.

## Author(s)

James T. Thorson, Ian J. Stewart, Andre E. Punt, Ian G. Taylor

## See Also

- `StepwiseFn()` will run multiple models.
- `PlotOutputFn()` will help summarize the output.

## Examples

```
example(SimulatorFn)
## Not run:
utils::write.csv(AgeReads,
  file = file.path(getwd(), "Simulated_data_example.csv")
)

## End(Not run)

##### Format data
Nreaders <- ncol(AgeReads)
# Change NA to -999 (which the Punt software considers missing data)
AgeReads <- ifelse(is.na(AgeReads), -999, AgeReads)

# Potentially eliminate rows that are only read once
# These rows have no information about reading error, but are potentially
# informative about latent age-structure. It is unknown whether eliminating
# these rows degrades estimation of error and bias, and is currently
# recommended to speed up computation
if (FALSE) {
  KeepRow <- ifelse(
    rowSums(ifelse(AgeReads == -999, 0, 1), na.rm = TRUE) <= 1,
```

```

    FALSE, TRUE
  )
  AgeReads <- AgeReads[KeepRow, ]
}

# AgeReads2 is the correctly formatted data object
AgeReads2 <- rMx(c(1, AgeReads[1, ]))

# Combine duplicate rows
for (RowI in 2:nrow(AgeReads)) {
  DupRow <- NA
  for (PreviousRowJ in 1:nrow(AgeReads2)) {
    if (all(
      AgeReads[RowI, 1:Nreaders] == AgeReads2[PreviousRowJ, 1:Nreaders + 1]
    )) {
      DupRow <- PreviousRowJ
    }
  }
  if (is.na(DupRow)) { # Add new row to AgeReads2
    AgeReads2 <- rbind(AgeReads2, c(1, AgeReads[RowI, ]))
  }
  if (!is.na(DupRow)) { # Increment number of samples for previous duplicate
    AgeReads2[DupRow, 1] <- AgeReads2[DupRow, 1] + 1
  }
}

##### Determine settings for ADMB
# Define minimum and maximum ages for integral across unobserved ages
MinAge <- 1
MaxAge <- ceiling(max(AgeReads2[, -1]) / 10) * 10
BiasOpt <- c(0, -1, 0, -3)
SigOpt <- c(1, -1, 6, -3)
# Necessary for SigOpt option 5 or 6
KnotAges <- list(NA, NA, c(1, 10, 20, MaxAge), NA)

##### Run the model (MAY TAKE 5-10 MINUTES)
## Not run:
fileloc <- file.path(tempdir(), "age")
dir.create(fileloc, showWarnings = FALSE)
RunFn(
  Data = AgeReads2, SigOpt = SigOpt, KnotAges = KnotAges,
  BiasOpt = BiasOpt,
  NDataSets = 1, MinAge = MinAge, MaxAge = MaxAge, RefAge = 10,
  MinusAge = 1, PlusAge = 30, SaveFile = fileloc,
  AdmbFile = file.path(system.file("executables",
    package = "nwfsAgeingError"
  )), .Platform$file.sep),
  EffSampleSize = 0, Intern = FALSE, JustWrite = FALSE, CallType = "shell"
)

## End(Not run)

```

---

SimulatorFn	<i>Simulate double reading data</i>
-------------	-------------------------------------

---

**Description**

A function to generate simulated double reading data with given properties

**Usage**

```

SimulatorFn(
  Nreaders,
  M,
  SelexForm,
  ErrorParams,
  BiasParams,
  SelexParams,
  ReadsMat,
  RecCv = 0.6,
  RecAr1 = 0.8,
  Amax = 100
)

```

**Arguments**

Nreaders	The number of ageing readers
M	True natural mortality
SelexForm	Form of selectivity-at-age (logistic selex-at-age is the only one that is implemented).
ErrorParams	Error type CV in the following equation: $\text{VarAgeRead} = (\text{CV} * \text{TrueAge})^2$
BiasParams	Bias type b in the following equation: $\text{EAgeRead} = b * \text{TrueAge}$
SelexParams	Selectivity parameters, which are standard to the logistic equation.
ReadsMat	Matrix describing number of reads per reader combination. Where each row specifies how many reads (in the first column) have a particular pattern of double reads (in the second through Nreaders+1 columns).
RecCv	CV of recruitment, and it should be noted that recruitment is assumed to be stationary over time.
RecAr1	First-order autoregressive coefficient for recruitment
Amax	True maximum age

**Value**

Returns a simulated double read matrix

**Author(s)**

James T. Thorson

**References**

Punt, A.E., Smith, D.C., KrusicGolub, K., and Robertson, S. 2008. Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australias southern and eastern scalefish and shark fishery. *Can. J. Fish. Aquat. Sci.* 65: 1991-2005.

**Examples**

```
# Parameters for generating data
# This represents 2 unique readers
# Row 1 -- Otoliths read only once by reader
# Row 2 -- Otoliths read twice by reader 1
# Row 3 -- Otoliths read only once by reader 2
# Row 4 -- Otoliths read twice by reader 2
# Row 5 -- Otoliths read once by reader 1 and once by reader 2
ReadsMat <- structure(matrix(
  nrow = 5, ncol = 5,
  c(
    rep(25, 5),
    1, 1, 0, 0, 1,
    0, 1, 0, 0, 0,
    0, 0, 1, 1, 1,
    0, 0, 0, 1, 0
  )
), dimnames = list(
  c(
    "Reader1_Only", "Reader1_DoubleReads",
    "Reader2_Only", "Reader2_DoubleReads",
    "Reader1_&_Reader2"
  ),
  c(
    "NumberOfReads",
    "Reader1", "Reader1_DoubleReads",
    "Reader2", "Reader2_DoubleReads"
  )
))

# Generate data
set.seed(2)
AgeReads <- SimulatorFn(
  Nreaders = 4, M = 0.2,
  SelexForm = "Logistic",
  SelexParams = c(5, 0.2), BiasParams = c(1, 1, 1.1, 1.1),
  ErrorParams = c(0.2, 0.2, 0.2, 0.2), ReadsMat = ReadsMat,
  RecCv = 0.6, RecAr1 = 0.8, Amax = 100
)
```

StepwiseFn

*Step-wise model selection***Description**

Run step-wise model selection to facilitate the exploration of several modelling configurations using Akaike information criterion (AIC).

**Usage**

```
StepwiseFn(
  SearchMat,
  Data,
  NDataSets,
  KnotAges,
  MinAge,
  MaxAge,
  RefAge,
  MaxSd,
  MaxExpectedAge,
  SaveFile,
  EffSampleSize = 0,
  Intern = TRUE,
  InformationCriterion = c("AIC", "AICc", "BIC"),
  SelectAges = TRUE
)
```

**Arguments**

SearchMat	<p>A matrix explaining stepwise model selection options. One row for each readers error and one row for each readers bias + 2 rows, one for MinusAge, i.e., the age where the proportion at age begins to decrease exponentially with decreasing age, and one for PlusAge, i.e., the age where the proportion-at-age begins to decrease exponentially with increasing age.</p> <p>Each element of a given row is a possible value to search across for that reader. So, the number of columns of SearchMat will be the maximum number of options that you want to include. Think of it as several vectors stacked row-wise where shorter rows are filled in with NA values. If reader two only has two options that the analyst wants to search over the remainder of the columns should be filled with NA values for that row.</p>
Data	<p>This is the data set with the first column being an integer providing the number of otoliths that are included in the row and the subsequent columns are the reader or lab estimated age where each reader/lab has a unique reading error and bias. The modeling framework allows for, at most, 15 readers, i.e., 16 columns. There should not be any identical rows in the data frame because otoliths that have the exact same read from every reader/lab should be combined into a single row</p>

with the count as the first column. If you failed to combine identical rows prior to running the model, you will be alerted with an error and the XXX.rep file will have a properly formatted data which can be cut-pasted into a XXX.dat file for use. Missing reads from a given reader/lab should be entered as -999. Order your reader/lab columns such that similar readers/labs are located next to one another because columns to the right can mirror columns to their immediate left in terms of parameter estimates.

NDataSets	This is generally 1 and other values are not implemented.
KnotAges	Ages associated with each knot. This is a necessary input for SigOpt = 5 or SigOpt = 6.
MinAge	An integer, specifying the minimum possible "true" age.
MaxAge	An integer, specifying the maximum possible "true" age.
RefAge	An arbitrarily chosen age from which "true" age-composition fixed-effects are calculated as an offset. This has no effect on the answer but could potentially effect estimation speed.
MaxSd	An upper bound on possible values for the standard deviation of reading error.
MaxExpectedAge	Set to MaxAge.
SaveFile	Directory where agemat.exe is located and where all ADMB intermediate and output files should be located. If AdmbFile is specified then agemat.exe is copied from that directory to SaveFile.
EffSampleSize	Indicating whether effective sample size should be calculated. Missing values in the data matrix will cause this to be ineffective, in which case this should be set to 0.
Intern	A logical input that controls the amount of output displayed, where TRUE indicates that ADMB output should be displayed in R and FALSE leads to the suppression of this information.
InformationCriterion	A string specifying the type of information criterion that should be used to choose the best model. The default is to use AIC, though AIC corrected for small sample sizes and BIC are also available.
SelectAges	A logical input specifying if the boundaries should be based on MinusAge and PlusAge. The default is TRUE.

### Details

AIC seems like an appropriate method to select among possible values for PlusAge, i.e., the last row of SearchMat, because PlusAge determines the number of estimated fixed-effect hyperparameters that are used to define the true proportion-at-age hyperdistribution. This hyperdistribution is in turn used as a prior when integrating across a true age associated with each otolith. This true age, which is a latent effect, can be interpreted as a random effect with one for each observation. So, the use of AIC to select among parameterizations of the fixed effects defining this hyperdistribution is customary (Pinheiro and Bates, 2009). This was tested for sablefish, where AIC lead to a true proportion at age that was biologically plausible.

### Author(s)

James T. Thorson

## References

Punt, A.E., Smith, D.C., KrusicGolub, K., and Robertson, S. 2008. Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australia's southern and eastern scalefish and shark fishery. *Can. J. Fish. Aquat. Sci.* 65: 1991-2005.

Pinheiro, J.C., and Bates, D. 2009. *Mixed-Effects Models in S and S-PLUS*. Springer, Germany.

## See Also

- RunFn() will run a single model, where this function runs multiple models.
- PlotOutputFn() will help summarize the output from RunFn().

## Examples

```
example(RunFn)
## Not run:
##### Run the model (MAY TAKE 5-10 MINUTES)
fileloc <- file.path(tempdir(), "age")
dir.create(fileloc, showWarnings = FALSE)
RunFn(
  Data = AgeReads2, SigOpt = SigOpt, KnotAges = KnotAges,
  BiasOpt = BiasOpt,
  NDataSets = 1, MinAge = MinAge, MaxAge = MaxAge, RefAge = 10,
  MinusAge = 1, PlusAge = 30, SaveFile = fileloc,
  AdmbFile = file.path(system.file("executables",
    package = "nwfscAgeingError"
  )), .Platform$file.sep),
  EffSampleSize = 0, Intern = FALSE, JustWrite = FALSE, CallType = "shell"
)
# Plot output
PlotOutputFn(
  Data = AgeReads2, MaxAge = MaxAge,
  SaveFile = fileloc, PlotType = "PDF"
)

## End(Not run)

##### Stepwise selection

# Parameters
MaxAge <- ceiling(max(AgeReads2) / 10) * 10
MinAge <- 1

##### Stepwise selection
StartMinusAge <- 1
StartPlusAge <- 30

# Define matrix explaining stepwise model selection options
# One row for each reader + 2 rows for
# PlusAge (age where the proportion-at-age begins to
# decrease exponentially with increasing age) and
# MinusAge (the age where the proportion-at-age begins to
```

```

# decrease exponentially with decreasing age)
# Each element of a given row is a possible value to search
# across for that reader
SearchMat <- array(NA,
  dim = c(Nreaders * 2 + 2, 7),
  dimnames = list(
    c(
      paste("Error_Reader", 1:Nreaders),
      paste("Bias_Reader", 1:Nreaders), "MinusAge", "PlusAge"
    ),
    paste("Option", 1:7)
  )
)
# Readers 1 and 3 search across options 1-3 for ERROR
SearchMat[c(1, 3), 1:3] <- rep(1, 2) %o% c(1, 2, 3)
# Reader 2 mirrors reader 1
SearchMat[2, 1] <- -1
# Reader 4 mirrors reader 3
SearchMat[4, 1] <- -3
# Reader 1 has no BIAS
SearchMat[5, 1] <- 0
# Reader 2 mirrors reader 1
SearchMat[6, 1] <- -1
# Reader 3 search across options 0-2 for BIAS
SearchMat[7, 1:3] <- c(1, 2, 0)
# Reader 4 mirrors reader 3
SearchMat[8, 1] <- -3
# MinusAge searches with a search kernal of -10,-4,-1,+0,+1,+4,+10
SearchMat[9, 1:7] <- c(
  StartMinusAge,
  StartMinusAge - 10,
  StartMinusAge - 4,
  StartMinusAge - 1,
  StartMinusAge + 1,
  StartMinusAge + 4,
  StartMinusAge + 10
)
SearchMat[9, 1:7] <- ifelse(SearchMat[9, 1:7] < MinAge,
  NA, SearchMat[9, 1:7]
)
# PlusAge searches with a search kernal of -10,-4,-1,+0,+1,+4,+10
SearchMat[10, 1:7] <- c(
  StartPlusAge,
  StartPlusAge - 10,
  StartPlusAge - 4,
  StartPlusAge - 1,
  StartPlusAge + 1,
  StartPlusAge + 4,
  StartPlusAge + 10
)
SearchMat[10, 1:7] <- ifelse(SearchMat[10, 1:7] > MaxAge,
  NA, SearchMat[10, 1:7]
)

```



```
# Run model selection
# This outputs a series of files
# 1. "Stepwise - Model loop X.txt" --
# Shows the AIC/BIC/AICc value for all different combinations
# of parameters arising from changing one parameter at a time
# according to SearchMat during loop X
# 2. "Stepwise - Record.txt" --
# The Xth row of IcRecord shows the record of the
# Information Criterion for all trials in loop X,
# while the Xth row of StateRecord shows the current selected values
# for all parameters at the end of loop X
# 3. Standard plots for each loop
# WARNING: One run of this stepwise model building example can take
# 8+ hours, and should be run overnight
## Not run:
StepwiseFn(
  SearchMat = SearchMat, Data = AgeReads2,
  NDataSets = 1, MinAge = MinAge, MaxAge = MaxAge,
  RefAge = 10, MaxSd = 40, MaxExpectedAge = MaxAge + 10,
  SaveFile = fileloc, InformationCriterion = c("AIC", "AICc", "BIC")[3]
)

## End(Not run)
```

# Index

0-9, [15](#)

ageing\_comparison, [2](#)  
ageing\_comparison(), [12](#)  
AgeRead, [19](#)

cMx, [3](#)  
CreateData, [4](#)  
CreateData(), [6](#)  
CreateSpecs, [5](#)  
CreateSpecs(), [6](#)

determine\_n\_sets, [5](#)  
DoApplyAgeError, [6](#)

estgrowth.vb, [7](#)

fit, [8](#)

minimizer, [8](#)

plot\_output, [11](#)  
PlotOutputFn, [9](#)  
ProcessResults, [12](#)

rMx, [13](#)  
run, [13](#)  
RunFn, [14](#)

SimulatorFn, [19](#)  
stats::optim(), [9](#)  
StepwiseFn, [21](#)