Package: nwfscAgeingError (via r-universe)

May 16, 2026

Type Package

Title Source Code for Punt et al. (2008): Ageing Error and Imprecision

Version 1.3.3

Description Source code and examples for how to use double-read data to estimate ageing imprecision and bias from fishery otoliths. The code was developed for Punt et al. (2008) and updated in 2022.

URL http://github.com/pfmc-assessments/nwfscAgeingError

BugReports http://github.com/pfmc-assessments/nwfscAgeingError/issues

License GPL-2

Encoding UTF-8

RoxygenNote 7.2.1

Roxygen list(markdown = TRUE)

Repository https://jimianelli.r-universe.dev

Date/Publication 2022-10-26 21:41:15 UTC

RemoteUrl https://github.com/jimianelli/nwfscAgeingError

RemoteRef HEAD

RemoteSha f9a8caad080d3db25e19591e93e90b4e8a2ee7f5

Contents

`ageing_comparison` *Plot comparison of double age readings*

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

---

estgrowth.vb *Calculate von Bertanlaffy Growth Parameters from Lengths and Ages*

Description

Calculate von Bertanlaffy 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 Bertanlaffy 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 purpopses. 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$par, ReturnType = "Pred",
  sdFactor = 1,
  Ages = bio_dat[, "Age"],
  Lengths = bio_dat[, "Length_cm"])
plot(bio_dat$Age, predictions[, "Lhat_pred"],
  xlab = "Age (years)", ylab = "Predicted length (cm)")
exp(solve$par)

## End(Not run)
```

nwfscAgeingError *Ageing error software*

Description

Data input and stepwise model selection in R for the Punt et al. (2008) ageing error model James T. Thorson, Ian Stewart, and André E. Punt

Details

Function name: RunFn()

Background:

```
The Punt et al. (2008) model calculates the likelihood of model parameters given an observed dataset that
```

This summation across all possible values for a 'True' age for each otolith also requires a hyperdistribution representing the 'prior' probability that an otolith is any given age; this prior is parameterized using a set of hyperparameters in addition to the parameters that govern the SD and bias for each reader. Specifically, one hyperparameter is estimated for every age between (and including) a MinusAge and a PlusAge, which are defined exogenously for every model run. Ages above the PlusAge or below the MinusAge have a prior Proportion-at-Age defined as a loglinear deviation from the Proportion-at-Age for the PlusAge and MinusAge. The slope of these loglinear deviations thus constitutes an additional 1 or 2 fixed effect parameter to estimate. The 'True' proportion-at-age is then calculated from these fixed effect and log-linear slope parameters by normalizing the resulting distribution so that it sums to one.

Necessary Inputs:

```
Format data: Data should be formatted with unique reading records as rows and readers/labs as columns (ex
```

have the same reading error and bias. Any instance where a particular reader (or lab) provides multiple reads for a single otolith can be dealt with by creating a 2nd column for that reader, and configuring the model so that parameters for that 2nd column mirror the parameters for the 1st column for that reader. Select inputs: The call-function 'FnRun()' in R writes data in the necessary format and then calls the Punt (2008) model. This model requires several inputs, which are listed and explained below:

Data: This is the data set as previously formatted. If the data has multiple rows with identical reads, this will cause an error and the 'XXX.rep' file will have a properly formatted data matrix which can be cut-pasted into a 'XXX.dat' file for use.

SigOpt: This a vector with one entry for each reader (i.e. Ncol-1 entries). Each entry specifies the functional form of reading error as a function of true age. Possible entries include: '-1', '-2', '-3', etc: This will make this reader mirror the estimated SD from another reader to it's left. '-1' causes it to mirror the estimated SD for the first reader, etc. This number has to be lower than the current entry number.

'1' : Constant CV, i.e., a 1 parameter linear relationship of SD with true age.

'2': Curvilinear SD, i.e., a 3 parameter Hollings-form relationship of SD with true age

'3': Curvilinear with CV, i.e., a 3-parameter Hollings-form relationship of CV with true age

'4': No error (but potentially bias)

BiasOpt: This is a vector with one entry for each reader:

'-1', '-2', '-3': See SigOpt

'0': Unbiased

'1': Constant CV, 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

NDataSets: This is generally '1' and other values are not implemented in the current R-code.

MinAge: The minimum possible 'True' age

MaxAge: 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 this MinusAge have 'true' proportion-at-age (Pa) estimated as P_a=P_MinusAge?e^(?(MinusAge-a)), where ? is an estimated log-linear trend in the 'true' proportion-at-age. If MinusAge = MinAge, ? 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 'agemat.exe' is located and where all ADMB intermediate and output files should be located.

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: 'TRUE' indicates that ADMB output should be displayed in R; 'FALSE' does not.

Stepwise model selection in R

Function name: StepwiseFn()

Background:

```
Stepwise model selection allows many different model configurations to be explored: in this code, I have
```

Necessary Inputs:

```
Format data: Same as for a single-run
```

Select inputs: Most inputs are the same as for a single-run. However, the 'SigOpt' 'BiasOpt' and 'PlusAge' are now specified using a matrix called 'PossibleMat', which has 2*Nreaders+2 rows and as many columns as necessary. Row #1-#Nreaders specify the SigOpt for each reader; Next are the BiasOpt for each reader, followed by the PlusAge. The first entry in each row specifies the starting value for that parameter in the search algorithm; a value must be specified in the first column for

each parameter. Any parameter for which the search algorithm should search across possible values has other possible values in the 2nd, 3rd, and subsequent cells in that row. An example is given in Table 2.

Diagnostic figures in R

Function name: PlotOutputFn()

Background:

There are many ways to visualize the results that are provided by the Punt et al. (2008) model. Some of t

Diagnostic plots include:

Error and bias by reader: A panel graph where each panel shows the expected and standard deviation in age reads for that reader. This is displayed against a scatterplot of the 'Read' and 'Estimated' ages for each otolith that was read by that reader.

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.

Necessary Inputs:

The plotting function reads the 'XXX.rep' and 'XXX.par' files that are located in the directory that is provided. It also requires specifying the MaxAge and Data, as formatted and defined earlier.

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. Canadian Journal of Fisheries and Aquatic Sciences 65: 1991-2005.

Examples

```
## Not run:
  # File for Punt et al. (2008) model (pre-compiled in ADMB)
  SourceFile = paste0(
    system.file(package='AgeingErrorPackage'),'/executables/'
  )

  # This is where all runs will be located
  DateFile = paste(getwd(),'/',Sys.Date(),'/',sep='')
    dir.create(DateFile)

  ############################
  #
  # Generate and run with an artificial dataset
  #
  # SimulatorFn() generates the data and has the following settings
  #   Nreaders is the number of readers
  #   ReadsMat is a matrix 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)
#   SelexForm is the selectivity-at-age form
#     (logistic selex-at-age is the only one that is implemented)
#   SelexParams are standard to the logistic equation
#   BiasParams b in the following equation:
#     E[AgeRead] = b*TrueAge
#   ErrorParams are CV in the following equation:
#     Var[AgeRead] = (CV*TrueAge)^2
#   RecCV is the is the CV in recruitment
#     (and recruitment is assumed stationary over time)
#   RecAR1 is first-order autoregressive coefficient in recruitment
#   Amax is the maximum allowable age
#
##############################

##### 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 2 -- 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
Nreaders = 4
ReadsMat = cbind(NumberOfReads=rep(100,5),
  Reader1=c(1,1,0,0,1), Reader1_DoubleReads=c(0,1,0,0,0),
  Reader2=c(0,0,1,1,1), Reader2_DoubleReads=c(0,0,0,1,0)
)
rownames(ReadsMat) = c("Reader1_Only", "Reader1_DoubleReads",
  "Reader2_Only", "Reader2_DoubleReads", "Reader1_&_Reader2"
)

# Generate data
AgeReads = SimulatorFn(Nreaders = Nreaders, 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
)
utils::write.csv(AgeReads,
  file = paste0(DateFile,"Simulated_data_example.csv")
)

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

# 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
```

```
# KeepRow = ifelse(rowSums(ifelse(AgeReads==-999,0,1),na.rm=TRUE)<=1,
#   FALSE,TRUE
# )
# AgeReads = AgeReads[KeepRow,]
# Combine duplicate rows
AgeReads2 = rMx(c(1, AgeReads[1,])) # correctly formatted data object
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)) AgeReads2 = rbind(AgeReads2, c(1, AgeReads[RowI,]))
  # Add new row to AgeReads2
  if(!is.na(DupRow)) AgeReads2[DupRow,1] = AgeReads2[DupRow,1] + 1
  # Increment number of samples for the previous duplicate
}

######## Determine settings for ADMB

# Generate vector with settings for Bias
# One entry for each reader
# -X = Mirror the parameters for reader X
# 0 = Unbiased (at least 1 reader has to be)
# 1 = Linear bias
# 2 = Curvilinear bias (3 param)
BiasOpt = c(0,-1,2,-3)

# Generate vector with settings for SD
# One entry for each reader
# -X = Mirror the parameters for reader X
# 0 = No error
# 1 = Constant coefficient of variation
# 2 = Curvilinear standard deviation (3 param)
# 3 = Curvilinear coefficient of variation (3 param)
SigOpt = c(3,-1,3,-3)

# Define minimum and maximum ages for integral across unobserved ages
MinAge = 1
MaxAge = ceiling(max(AgeReads2)/10)*10

# Run the model
# Data=AgeReads2; SigOpt=SigOpt; BiasOpt=BiasOpt;
#   NDataSets=1; MinAge=MinAge; MaxAge=MaxAge;
#   RefAge=10; MinusAge=1; PlusAge=30; MaxSd=40;
#   MaxExpectedAge=MaxAge+10; SaveFile=DateFile;
#   AdmbFile=SourceFile; EffSampleSize=0; Intern=TRUE
RunFn(Data = AgeReads2, SigOpt = SigOpt, BiasOpt = BiasOpt,
  NDataSets = 1, MinAge = MinAge, MaxAge = MaxAge, RefAge = 10,
  MinusAge = 1, PlusAge = 30, SaveFile = DateFile, AdmbFile = SourceFile,
  EffSampleSize = 0, Intern = FALSE, JustWrite = FALSE
)

# Plot output
```

```
  # Data = AgeReads2; MaxAge = MaxAge; SaveFile = DateFile; PlotType = "PDF"
  PlotOutputFn(Data = AgeReads2, MaxAge = MaxAge, SaveFile = DateFile,
    PlotType = "PDF"
  )

## End(Not run)
```

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 ag,e 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 `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`.

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 ageing_comparison().

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 hte 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()

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

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 ag,e 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., `length(SigOpt) == NCOL(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 `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.
`AdmbFile`	An optional character entry that specifies the directory from which `agemat.exe` is to be copied from to `SaveFile`.
`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 `verbose = TRUE`.

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 = "nwfscAgeingError"), .Platform$file.sep),
  EffSampleSize = 0, Intern = FALSE, JustWrite = FALSE, CallType = "shell"
)

## End(Not run)
```

`SimulatorFn` *Simulate double reading data*

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: VarAgeRead = (CV*TrueAge)^2
`BiasParams`	Bias type b in the following equation: EAgeRead = b*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 shoudl 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 2 -- 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` 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.

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.

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 ag,e 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.
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